

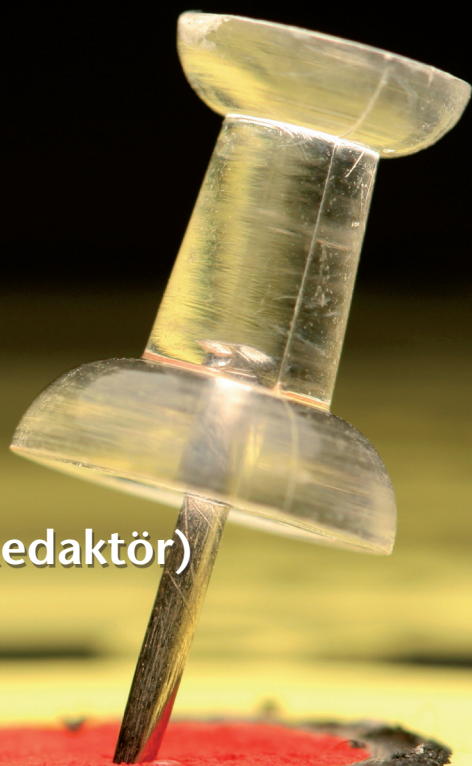


Statistiska centralbyrån Statistics Sweden

Frågor och svar

om frågekonstruktion i enkät- och intervjuundersökningar

Andreas Persson (Redaktör)



Frågor och svar

– om frågekonstruktion i enkät-
och intervjuundersökningar

Statistiska centralbyrån
2016

Frågor och svar

– om frågekonstruktion i enkät- och intervjuundersökningar

Questions and answers

– question design in self-administered- and interview questionnaires

Statistics Sweden

2016

Producent
Producer SCB
Statistics Sweden
Box 24300, SE-104 51 Stockholm
+46 10 479 40 00

Förfrågningar
Enquiries Andreas Persson +46 10 479 62 48
andreas.persson@scb.se

Lina Fjelkegård +46 10 479 43 59
lina.fjelkegard@scb.se

ISBN 978-91-618-1653-8 (Print)
URN:NBN:SE:SCB-2016-X08BR1601_pdf
Printed in Sweden
SCB-Tryck, Örebro 2016.12

Förord

”Som man frågar får man svar” är ett talesätt som i högsta grad är relevant för statistiska undersökningar. Att ställa frågor där svaren syftar till att skapa underlag för statistik ställer höga krav på den som konstruerar frågorna.

Under 1980-talet gav Statistiska centralbyrån ut boken ”Att fråga – Om frågekonstruktion vid intervjuundersökningar och postenkäter” (Wärneryd m.fl.). Den senaste versionen kom 1990.

Sedan dess har förutsättningarna för datainsamling förändrats och forskningen om frågekonstruktion gått framåt. ”Frågor och svar” tar till viss del avstamp i ”Att fråga” men är en fristående bok, anpassad till dagens förutsättningar. Boken ger både en introduktion till frågekonstruktion och en fördjupning i ämnet. Den ger praktiska tips men innehåller även teoretiska resonemang utifrån forskning och praktiska erfarenheter.

Huvudförfattare och redaktör är Andreas Persson vid Statistiska centralbyrån och medförfattare är Lina Fjelkegård, Statistiska centralbyrån, Pia Hartwig, Statistiska centralbyrån och Anna Sundström, Umeå universitet. Utöver dessa har ett flertal kunniga personer bidragit med synpunkter.

Åsikter och uppfattningar i denna publikation är författarnas egna och överensstämmer inte nödvändigtvis med SCB:s.

Statistiska centralbyrån i december 2016

Stefan Lundgren

Maj Eriksson Gothe

SCB tackar

Tack vare våra uppgiftslämnare – privatpersoner, företag, myndigheter och organisationer – kan SCB tillhandahålla tillförlitlig och aktuell statistik som tillgodoser samhällets informationsbehov.

Innehåll

Förord	1
Om boken	7
Bakgrund	7
Målsättning och målgrupp	8
1 Att skriva frågor Av A.Persson	11
Vad är en fråga?	12
Frågornas mål	17
Att skriva frågor i praktiken	20
Att tänka på respondenterna	24
Praktiska språktips	26
Sammanfattning	30
2 Frågeteknik Av A.Persson	33
Övergripande mål för frågekonstruktion	33
Två frågor i samma fråga	33
Inbyggda förutsättningar	34
Definiera frågan i tid och rum	37
Ledande frågor	39
Filter- och följdfrågor.....	43
Frågeformuleringens innehåll	45
Att tänka på med frågor som mätinstrument	56
Frågekonstruktion och intervjuer	63
Svarsalternativen.....	67
Känsliga frågor	70
Sammanfattning	75
3 Svarsalternativ Av A.Persson	81
Svarsalternativ	81
Öppna svarsalternativ.....	83

Öppna frågor kontra slutna svarsalternativ	89
Slutna svarsalternativ	90
Slutna svarsalternativ utan given inbördes ordning	110
Svarsalternativ i intervjufrågor	116
Svarsalternativ och tolkningar	119
Sammanfattning	127
4 Enkätmetodik <i>Av L. Fjelkegård och A. Persson</i>	134
Formulärkonstruktion	134
Annat utskicksmaterial.....	145
Sammanfattning	158
5 Layout <i>Av P. Hartwig och A. Persson</i>	162
Att kommunicera med grafik och symboler	162
Perception – hur vi förstår information från våra sinnen	163
Allmänt om design	168
Layout i enkäter (pappers- och webbenkät)	169
Layout i webbenkäter	188
Sammanfattning	199
6 Datainsamlingsmetod <i>Av A. Persson</i>	203
Olika datainsamlingsmetoder	203
Likheter och skillnader mellan olika datainsamlingsmetoder	205
Frågetyper och datainsamlingsmetod	228
Mixed-mode (kombinerad insamling)	231
Sammanfattning	254
7 Språket och kommunikation <i>Av A. Persson</i>	260
Språket som kommunikationsmedel	260
Vad är språket?	261
Konversationens logik och praktik.....	264
Några övergripande språkprinciper	271

Sammanfattning	281
8 Att mäta med frågor <i>Av L. Fjelkegård</i>	284
Varför undersöka med hjälp av frågor?	284
Från problem till fråga.....	288
Sammanfattning	305
9 Svarsprocessen och felkällor <i>Av A. Persson</i>	306
Inledning	306
Forskning om frågekonstruktion – historik och teoriutveckling.....	306
Kognition och kommunikation i svarsprocessen	308
Svarsprocessen i företag och organisationer	312
Felkällor och mätfel	314
Sammanfattning	321
10 Att fråga efter kunskap <i>Av A. Sundström</i>	324
Kunskapsfrågor i enkätundersökningar	324
Att mäta olika typer av kunskap	325
Konstruktion av kunskapsfrågor	327
Sammanfattning	337
11 Att fråga om det förflutna <i>Av A. Persson</i>	339
Hur minnet fungerar.....	339
Allmänt om källor till minnesfel	341
Att minnas <i>när</i> något hände (datering).....	345
Att minnas <i>antalet gånger</i> något hände (frekvensbestämning).....	351
Att främja återerinring och minska minnesfel	356
Sammanfattning	360
12 Att fråga om attityder <i>Av A. Persson</i>	364
Attityder kan vara svåra att mäta.....	364
Vad är en attityd?	364
Svårigheten med attitydfrågor	366

Abstrakt och konkret nivå – konsekvenser för frågekonstruktion.....	369
Hur går det till när attitydfrågor besvaras?	373
Att konstruera attitydfrågor	375
Sammanfattning	378
13 Ordningsföljder och kontext <i>Av A. Persson</i>	381
Ordningsföljder och kontext.....	381
Frågors ordningsföljd.....	382
Ordningsföljd <i>inom</i> frågor	386
Andra typer av kontext.....	389
Positionseffekter	392
Sammanfattning	398
14 Att utvärdera frågor <i>Av A. Persson</i>	402
Varför utvärdera?.....	402
Olika metoder för att utvärdera frågor	403
Kombinationer av testmetoder.....	431
Sammanfattning	431
Ordlista	436
Referenser	439
Tabell- och figurförteckning	466

Om boken

Andreas Persson, SCB

Bakgrund

”Att fråga – Om frågekonstruktion vid intervjuundersökningar och post-enkäter” av Bo Wärneryd m.fl. utgavs 1980 av SCB. Från början var den tänkt för internt bruk men den fick god spridning även utanför myndigheten. Därför släpptes en omarbetad upplaga 1986 som sedan följdes av en uppdatering 1990. ”Att fråga” har använts inom många olika områden, till exempel som kurslitteratur på universitet och högskolor. Den har fyllt en viktig roll på ett område där det inte funnits så mycket litteratur på svenska.

Arbetet med den här boken – ”Frågor och svar” – inleddes som en uppdatering av ”Att fråga”. Ganska tidigt togs emellertid beslutet att använda ett nytt format.

”Frågor och svar” är alltså både en fristående bok och ett slags uppföljare till ”Att fråga”. ”Att fråga” har dock inte gått förlorad utan tvärtom nyligen digitaliserats. Den finns att ladda ner utan kostnad som pdf på [SCB:s webbplats](#).

”Frågor och svar” bygger också på resultat från den internationella forskningen på området samt på praktiska erfarenheter från framförallt SCB:s mättekniska enhet. Även om boken delvis bygger på erfarenheter från SCB är den dock mer allmän i ansatsen och inte en beskrivning av myndighetens verksamhet och tillvägagångssätt på området.

Boken som helhet har ett fokus på individ- och hushållsundersökningar men mycket av innehållet är även tillämpligt på undersökningar till företag och organisationer.

Frågekonstruktion är ett område där många saker kan variera och där det inte alltid är givet vilken lösning som är bäst. Den här boken ska därför inte ses som ett facit för frågekonstruktion utan som ett kunskapsunderlag.

Målsättning och målgrupp

På det stora hela är frågekonstruktion ett ganska litet forskningsområde. Det finns en internationell litteratur men inte så många svenska motsvarigheter. Inte sällan utgörs den svenska litteraturen om frågekonstruktion av enskilda kapitel i bredare böcker om undersökningsmetodik ("Att fråga" utgör ett av ganska få undantag; se även Wenemark, 2017). Denna bok har dock fokus på just frågekonstruktionen. Av den anledningen riktar den sig inte till någon specifik, avgränsad målgrupp. Tvärtom strävar den efter att kunna tillgodose många olika behov som kan tänkas finnas på området.

Boken innehåller praktiska tips för den som vill ha omedelbart stöd till formulerandet eller redigerandet av enkät- eller intervjufrågor. Den innehåller också teoretiska idéer om olika fenomen som kan påverka svaren på frågorna. Dessa teoretiska idéer är inte lika enkla att omedelbart överföra till en konkret fråga i ett visst frågeformulär, men de ökar förhoppningsvis den allmänna förståelsen för viktiga faktorer inom området, vilken sedan kan komma till nytta i en mängd olika situationer. Boken innehåller även information som ligger på en bredare nivå men som är relevant för frågekonstruktion (till exempel hur språket fungerar).

Boken har alltså olika typer av innehåll. Lite förenklat skulle man kanske kunna använda distinktioner som "teori – praktik" eller "grunder – fördjupning" för att beskriva dessa. En viss del av innehållet ligger på en teoretisk nivå och en annan på en praktisk nivå. Kanske intresserar dessa olika läsare men det kommer säkert att vara vanligt att samma läsare söker olika typ av innehåll i olika situationer. Det är således viktigt att kunna navigera i boken utifrån det behov man har. Det har inte varit något självändamål med boken att den ska läsas från pärm till pärm.

Hur boken kan läsas

Boken innehåller alltså olika typer av material. Det var inte givet hur det bäst skulle organiseras. Ett alternativ skulle kunna ha varit att använda en kronologisk modell som följer processen att skapa ett frågeformulär från ax till limpa. Nackdelen med en sådan uppdelning är att de bitar som flest läsare förmodligen kommer att återvända till (till exempel tips om frågeteknik och svarsalternativ) då skulle ha hamnat utspritt och relativt sent i boken. Med en sådan ansats skulle dessutom inledningen ha blivit ganska teoritung.

Istället har materialet delats upp i grunder och fördjupningar, enligt distinktionen i avsnittet ovan. Först kommer de mer grundläggande kapitlen:

- 1) Att skriva frågor
- 2) Frågeteknik
- 3) Svartalternativ
- 4) Enkätmetodik
- 5) Layout
- 6) Datainsamlingsmetod

Det första kapitlet är en introduktion till att skriva frågor i praktiken. Det tar upp förutsättningar och mål. Det andra kapitlet är en samling tumregler att tillämpa, och fallgropar att undvika, vid frågekonstruktion. Det tredje kapitlet innehåller tips och saker att tänka på vid utformandet av svartalternativ. Det fjärde kapitlet innehåller metoder om hur frågor sätts samman i ett frågeformulär och för kommunikationen med urvals-personerna. Det femte kapitlet handlar om den visuella designen, enkätens layout och det sjätte om hur olika datainsamlingsmetoder kan påverka mätningen.

Resterande kapitel i boken innehåller mer teori och fördjupningar:

- 7) Språket och kommunikation
- 8) Att mäta med frågor
- 9) Svartsprocessen och felkällor
- 10) Att fråga efter kunskap
- 11) Att fråga om det förflutna
- 12) Att fråga om attityder
- 13) Ordningsföljder och kontext
- 14) Att utvärdera frågor

Det inledande kapitlet i fördjupningsdelen, kapitel sju, beskriver språket som kommunikationsmedel. Kapitel åtta handlar om det förarbete som bör göras innan man börjar formulera frågor. Det nionde kapitlet innehåller olika teoretiska modeller för besvarandet, den så kallade svartsprocessen. De ger en förståelse för vad som kan gå fel och vad som kan vara svårt när det gäller frågor. Alla dessa tre kapitel är högst relevanta för frågekonstruktion men innehåller inte så många praktiska tips som direkt kan överföras till ett frågeformulär. Därefter följer

ytterligare tre kapitel som handlar om risker och möjligheter med tre typer av frågor. Kapitel tretton handlar om hur relationen mellan frågorna kan påverka svaren. Det fjortonde och avslutande kapitlet beskriver hur frågorna kan utvärderas för att se om de fungerar som tänkt.

Den här uppdelningen är dock inte fullkomlig. Fördjupningskapitlen innehåller också en del praktiska tips och de grundläggande kapitlen innehåller också en del djupare förklaringar. För att hjälpa läsaren avslutas alla kapitel med punktlistor med kapitlets viktiga punkter och saker att tänka på vid frågekonstruktion. De är tänkta som genvägar för dem som är ute efter "snabba fakta" utan att behöva läsa de bakomliggande resonemangen. De kan också fungera som en vägledande innehållsförteckning för läsaren. De som vill ha konkreta tips på frågenivå är dock generellt bäst hjälpta av de inledande huvudkapitlen. De som vill veta mer om de bakomliggande teorierna har mest att hämta i fördjupningskapitlen.

Målsättningen har varit att boken ska kunna förstås utan stora krav på förkunskaper men viss erfarenhet av enkät- eller intervjufrågor och undersökningar underlättar förståelsen, speciellt av kapitlen i bokens andra hälft. I slutet av boken finns en ordlista som förklarar de vanligaste begreppen och hur de används i denna bok.

SCB tackar

Tack vare våra uppgiftslämnare – privatpersoner, företag, myndigheter och organisationer – kan SCB tillhandahålla tillförlitlig och aktuell statistik som tillgodoser samhällets informationsbehov.

Tack till Pernilla Andersson, Erika Bergentz, Linda Björneskog, Anette Björnram, Karolina Bock, Helena Bäckström, Eva Elvers, Johan Erikson, Maj Eriksson Gothe, Lina Fjelkegård, Eva Furubjelke, Karin Hansson, Pia Hartwig, Marcus Hugosson, Boris Lorenc, Per-Erik Lyrén, Joakim Malmdin, Isabel Pettersson, Martin Ribe, Fredrik Scheffer, Tova Stenlund, Anna Sundström, Jennica Wallenborg, Marika Wenemark, Klas Wibell, Maria Öman och mätteknikerna på Enheten för mätteknik som lämnat synpunkter på olika delar av bokens innehåll. De eventuella felaktigheter som återstår är författarnas egna.

1 Att skriva frågor

Författare: Andreas Persson (SCB)

I media kan vi läsa att de politiska partierna tappar eller vinner sympatier, att en viss kost kan ha konsekvenser för hälsan och att andelen arbetslösa förändrats. Detta är bara några axplock ur nyhetsflödet en helt vanlig dag. Något dessa exempel har gemensamt är att de bygger på undersökningar som ställt frågor till människor. Att fråga är ett naturligt och bra sätt att få veta hur saker och ting ligger till. Det är en central del i vår kommunikation med andra människor och företeelsen att ställa frågor är förmodligen lika gammal som språket självt. Att ställa frågor har alltså ett viktigt vardagligt syfte. Men syftet kan också vara utredande och vetenskapligt, där vi har ett ämne vi vill veta mer om. Vissa ämnen, speciellt inom samhällsvetenskapen, kan främst undersökas genom att ställa frågor till människor. Frågeställandet kan ske genom olika former av kommunikation. Det kan ske fritt och på djupet, till ett fåtal personer (så kallade kvalitativa intervjuer) eller mer standardiserat till en större grupp personer, till exempel via en enkät. Det är den sistnämnda formen av frågeställande som den här boken handlar om.

När man ställer frågor är målsättningen naturligtvis att motparten ska förstå frågorna på rätt sätt och kunna besvara dem. Sker det genom en enkät eller standardiserad intervju ska svaren dessutom gärna passa något av de svarsalternativ som erbjuds. Detta kan tyckas vara enkla förutsättningar. Många forskningsrapporter och praktiska erfarenheter pekar emellertid på att det inte alltid är så enkelt utan tvärtom kan vara svårt att uppfylla den grundläggande målsättningen till fullo. Till viss del beror det på att frågeställandet är beroende av språket, som i sig är ett flexibelt verktyg. Språket kan till exempel liknas vid en kameleont – innebörden kan skifta mellan olika sammanhang, vilket förstås kan få konsekvenser för hur frågorna uppfattas av motparten. Det finns även andra aspekter som kan komplicera frågeställandet. Följden kan bli att undersökningen får ett delvis missvisande eller i värsta fall ett felaktigt resultat. Frågorna är alltså av yttersta vikt.

Vad är en fråga?

Det kan vara på sin plats att i en bok om frågekonstruktion redan i inledningen resonera kring och definiera vad en fråga faktiskt är för något. Vad menas med en fråga? På ett övergripande plan kan en fråga definieras som *en språkhandling riktad till någon med målet att få information som respons*. Frågor kan dock vara formulerade på olika sätt grammatiskt (läs mer i Tourangeau, Rips & Rasinski, 2000).

Frågesats

Det mest tydliga exemplet på en fråga är frågesatsen som i sin tur kan delas upp i:

- ja-nej-frågor

EXEMPEL: *Har du fast anställning?*

- frågeordsfrågor

EXEMPEL: *Vilken form av anställning har du?*

Det som kännetecknar frågesatsen är att den har en lite annan ordföljd jämfört med motsvarande påståendeform – subjektet och verbet/hjälperbet byter plats. Jämför påståendeformen "Du har fast anställning" med frågeformen "Har du fast anställning?". Frågesatsen avslutas också med ett frågetecken. I ja-nej-frågor är det generellt sett den enda skillnaden mellan frågeform och påståendeform. Det är därför enkelt att omvandla ett påstående till en ja-nej-fråga och tvärtom (se figur nedan). Det kan vara bra att känna till eftersom båda formerna är vanliga i enkäter.

Figur 1.1

Frågesats

Har du fast anställning?

Påståendeform

Du har fast anställning

Frågeordsfrågor ("Vilken form av anställning har du?") avslutas också med frågetecken. Subjektet och verbet/hjälperbet byter plats men satsen

inleds dessutom med ett frågeord (interrogativt pronomen eller adverb), till exempel "hur", "vem", "vad", "vilken" eller "när". Frågeordsfrågor kan inte lika enkelt omvandlas till påståenden. Till skillnad från påståenden och ja-nej-frågor ska respondenten i frågeordsfrågor inte bara ta ställning till innehållet i frågan utan också komma fram till ett svar, som i sig inte ingår i frågeformuleringen (Tourangeau, Rips & Rasinski, 2000). Exempel frågan ovan ("Vilken form av anställning har du?") kan alltså inte besvaras med ett ställningstagande (till exempel ja eller nej) utan svaret ligger utanför frågan. Detta är en väsentlig skillnad mellan frågeordsfrågor, ja-nej-frågor och påståenden.

Andra former av frågor

Frågesatsen är kanske det vanligaste sättet att ställa frågor. Men det finns andra sätt och alla frågesatser är inte frågor i den betydelsen att de syftar till att få information från mottagaren (som en fråga definierades ovan).

EXEMPEL:

Kan du sluta skrika?

Frågan ovan är formulerad som en ja-nej-fråga men den är inte en språkhandling avsedd att få information från en mottagare. Syftet är inte heller att den ska besvaras med ett "Ja" eller "Nej". Snarare är det en uppmaning till någon att sluta skrika. På motsvarande sätt finns det språkhandlingar som har som syfte att ge information men som inte är en frågesats, till exempel:

EXEMPEL:

Ange din anställningsform i fältet nedan.

Frågor i frågeformulär

Att fastställa vad en fråga har för grammatisk form är alltså inte helt enkelt eftersom de kan uttryckas på olika sätt språkligt. Frågor i frågeformulär syftar till att få information från en mottagare som respons och i boken kallas de hädanefter för frågor, oavsett deras grammatiska form. De vanligaste formerna är:

- frågesats (ja-nej-frågor och frågeordsfrågor)

EXEMPEL: *Vilken är din huvudsakliga sysselsättning?*

- uppmaning
EXEMPEL: *Markera i listan nedan vilka fritidsaktiviteter du ägnar dig åt.*
- påstående
EXEMPEL: *Jag tycker om att ta risker.*
- nyckelord
EXEMPEL: *Födelseår.*

Alla dessa varianter är vanligt förekommande i frågeformulär. Frågesatsen är redan beskriven ovanför. En uppmaning ber någon, i detta fall respondenten, att utföra en handling (till exempel att skriva något i enkäten). Ett påstående deklarerar att något förhåller sig på ett visst sätt. Respondentens uppgift är att avgöra i vilken grad det stämmer eller om påståendet är sant eller falskt. Nyckelordet kan ses som en förkortad variant av en fråga eller kanske en uppmaning. Ett enskilt ord representerar då den underliggande frågeformuleringen. Utgångspunkten är att det är underförstått vad respondenten förväntas göra och svara på.

Det finns ingen statistik för hur vanliga de olika typerna är. Vår erfarenhet från Enheten för mätteknik på SCB är att frågesatsen är den vanligaste formen i undersökningar riktade till individer och hushåll och att påståendeformen är den näst vanligaste. I undersökningar riktade till företag och organisationer är uppmaningen vanligare som frågeform och påståenden mer ovanliga.

Fördelar och nackdelar med de olika varianterna

Frågesatsen bäst

Vår erfarenhet på Enheten för mätteknik på SCB är att frågesatsen generellt sett är det bästa sättet att ställa frågor på. Frågesatsen är nämligen en så pass central del i människors kommunikation i allmänhet – när man får en fråga förväntas man svara. Detta förekommer så ofta i vardagen att det nästan är en inlärd reflex hos människor. Uppmaningar kan vara tydliga av samma skäl som frågor. En uppmaning förväntas man följa, men den har en mer styrande ton som det kan finnas goda skäl att undvika.

Nyckelord bör överhuvudtaget inte förkomma i frågeformulär. I vissa väldigt tydliga fall kan de fylla sin funktion utan större problem, men de

ger ett mindre seriöst intryck än frågesatser. Frågeformuläret kan framstå mer som ett protokoll eller en checklista.

Påståenden – en debatterad frågeform

Påståenden har en lång tradition som frågeform. Särskilt vanliga är de vid mjukare mätningar, som till exempel av attityder. Vår uppfattning är att de ibland används per automatik utan att det har övervägts om en frågesats skulle ha passat bättre.

EXEMPEL:

Jag trivs med mitt jobb.

Alternativ frågesats: Trivs du med ditt jobb?

Till skillnad från att besvara frågor är det sällan man behöver ta ställning till påståenden i vardagen. Frågeformulär är nästan det enda tillfälle när personer tvingas ta ställning till listor med olika påståenden. Därför kan det vara ett ovant moment för många och kognitiva tester visar att en del respondenter (speciellt unga och äldre) kan ha problem med det.

Det finns även en del andra problem med påståenden. Ett problem är att det kan vara lättare att hålla med om än att ta avstånd från ett påstående. Det gäller speciellt om respondenten är okoncentrerad eller inte så insatt i ämnet. Andra problem grundar sig inte främst på påståendeformatet i sig utan på svarsskalan (som ofta är baserad på "instämmer" eller "stämmer"). Eftersom alla påståenden kan ha samma svarsskala kan det vara frestande att skapa långa tabeller av påståendefrågor. Det kan leda till att respondenterna börjar slarva med svaren och markera samma alternativ rakt igenom (så kallad "straight-lining"). Det kan också bli problem för respondenten att matcha sitt svar mot svarsskalan (Revilla, Saris & Krosnick, 2014).

Vissa forskare är alltså kritiska till påståendeformen (Converse & Presser, 1986; Schaeffer & Presser, 2003). Helt klart är att vid raka faktafrågor bör man överväga frågesatser istället för påståendeformen (läs mer i kapitel 3).

Sammanfattande punkter:

- Siffrorna i en undersökning baseras ofta på språkliga förfrågningar till respondenter – frågorna är viktiga!

- En fråga definieras som *en språkhandling riktad till någon med målet att få information som respons*.
- Frågesatsen är den vanligaste formen att ställa frågor på.
- Frågesatsen kan delas upp i ja-nej-frågor och frågeordsfrågor (inleds till exempel med hur, var, när, varför).
- Frågor kan också ställas som uppmaningar, påståenden eller med hjälp av nyckelord.
- Frågesatsen är den vanligaste (och bästa) formen att ställa frågor på. Påståenden är ganska vanliga i undersökningar riktade till individer och hushåll. Uppmaningar är vanligare i undersökningar till företag och organisationer.
- Det kan vara lättare att hålla med än att ta avstånd från påståenden.
- Vid faktabetonade frågor kan man överväga om påståendena kan omvandlas till frågesatser.

Andra frågetyper

Frågor kan alltså delas upp utifrån deras grammatiska form men frågor kan också delas in i typer baserat på vad de syftar till att mäta. De vanligaste typerna är:

- Faktafrågor: frågor som frågar efter en faktauppgift, exempelvis vilket år respondenten är född.
- Beteendefrågor: frågor som frågar efter respondentens beteende, vanligtvis om eller hur ofta respondenten utfört en viss aktivitet.
- Kunskapsfrågor: frågor som syftar till att mäta respondentens kunskap om något.
- Attitydfrågor: frågor som syftar till att mäta respondentens åsikt om eller attityd till något.
- Upplevelsefrågor: den här typen av frågor har ingen given beteckning utan olika varianter kan säkert förekomma. Typen syftar till att mäta hur respondenterna tänker och känner, till exempel hur de mår.
- Hypotetiska frågor: frågor som ber respondenten ta ställning till något hypotetiskt, alltså något som inte existerar men som skulle kunna göra det. Exempelvis vad respondenten skulle göra om han eller hon vann på lotto.

Lite förenklat kan man säga att de tre översta typerna är objektiva frågetyper i den betydelsen att det för dem finns ett sant svar som, åtminstone i teorin, respondentens svar skulle kunna kontrolleras mot. De tre nedre kan på motsatt sätt kallas för subjektiva frågetyper eftersom de berör respondentens egen uppfattning, bedömning eller upplevelse. Ingen av frågetyperna ovan är entydig utan de överlappar varandra. En del frågor kan dessutom vara svåra att placera i endast en av typerna och kan passa in på flera. Likväl kan frågetyperna vara en bra utgångspunkt för diskussion om frågekonstruktion och används även i denna bok (t.ex. har attitydfrågor och kunskapsfrågor egna kapitel).

- Frågor kan också delas in i typer baserat på vad de syftar till att mäta, t.ex. faktafrågor, beteendefrågor och attitydfrågor.

Frågornas mål

Frågekonstruktören

”Frågekonstruktören” används i denna bok som beteckning på den person som konstruerar frågorna i ett frågeformulär och ansvarar för frågetekniken (att frågorna skrivs enligt ”konstens alla regler”, se kapitel 2). Beteckningen kan ge sken av att det är ett eget yrke att vara frågekonstruktör men så är det sällan. Snarare syftar beteckningen på en roll med en viss kompetens. Det kan vara en vanlig roll för vissa och en ovanlig roll för andra, beroende på hur mycket man arbetar med frågor och undersökningar. Samma person kan också ha flera roller i en undersökning, till exempel både ansvara för frågekonstruktionen och för det praktiska genomförandet av undersökningen. Det kan även vara så att flera personer delar samma roll i en undersökning. Exempelvis att det finns flera frågekonstruktörer som tillsammans konstruerar frågorna eller att de ansvarar för olika avsnitt i frågeformuläret. Alltså, med ”frågekonstruktör” avses den roll som ansvarar för att frågorna konstrueras enligt frågeteknikens regler.

Grundregler

Senare kapitel i denna bok kommer att beskriva principer för frågekonstruktion mer djupgående, men redan i inledningen av boken kan det vara bra att fastslå ett antal grundregler. Dessa kan ses som

självklarheter som de flesta frågekonstruktörer håller med om, men de är likväl viktiga och bra att ha i bakhuvudet:

- Frågan ska kunna förstås av alla respondenter.
- Frågan ska förstås på samma sätt av alla respondenter.
- Frågan ska mäta det som avses.
- Frågan ska vara så kort som möjligt (men inte förkortas på förståelsens bekostnad).
- Respondenten ska inte reagera negativt på frågan.
- Frågan ska ha ett syfte (annars bör den inte ställas och ta upp respondenternas tid).
- Det ska finnas ett svarsalternativ för alla respondenter, men inte fler än ett per respondent om inte frågan tillåter det.
- Det ska vara tydligt hur frågan ska besvaras.
- Svarsalternativens format ska vara förenligt med de analyser man planerar att göra.

Dessa regler bör alltså frågekonstruktören ha i åtanke. Regelbrott kan få tråkiga konsekvenser – oftast i form av lägre kvalitet på de data som samlas in.

Illusionen om den perfekta frågan

De flesta frågekonstruktörer har nog som mål att skriva den perfekta frågan – en fråga som inbegriper alla grundregler ovan. Det är en riktig utmaning! Det kan därför vara på sin plats att utveckla grundreglerna en aning och avslöja att den perfekta frågan är nästintill en omöjlighet. Det gör inte andemeningen i grundreglerna mindre sanna men det är svårt att uppfylla dem alla i praktiken. Detta beror på att det finns så många variationer att det kan vara knivigt att anpassa frågan efter dem alla.

Den största variationen står nog respondentgruppen för. De är människor och människor är olika, både på gruppnivå och som individer. De kan till exempel variera vad gäller språket genom att ha olika ordförråd, tolka samma ord på olika sätt och vara olika i hur de läser och lyssnar. Vissa kan till exempel behöva få ett ord eller en term i en fråga förklarad medan andra, som redan kände till betydelsen, uppfattar förklaringen som störande.

Respondenterna varierar förstås också vad gäller erfarenheter, åsikter och situation. En fråga kan vara som skraddarsydd för vissa respondenter men passa andra respondenters situation eller erfarenheter mindre bra.

En lösning kan vara att tillföra en beskrivning så att frågan är relevant för alla, men då kan den istället bli krånglig och för lång för vissa.

Frågor kan också variera i hur de förhåller sig till ämnet de ska mäta. En fråga kan vara utmärkt för att mäta en viss aspekt av till exempel levnadsförhållanden eller sysselsättning, men mindre bra för andra aspekter. Hur perfekt en fråga är beror alltså på vad som ska mätas. En fråga som passar utmärkt för att mäta en viss aspekt av något är alltså inte nödvändigtvis lika bra i ett annat syfte.

Frågor kan också ha ett bäst före-datum. En fråga som fungerade bra för tio år sedan är kanske inte lika lämplig idag. Respondenterna kan ha förändrats så att de uppfattar den annorlunda. Språket kan ha utvecklats så termer och ord har lite andra betydelser än förr. Oftast går den här typen av förändringar långsamt men de kan också inträffa plötsligt. Till exempel kan en händelse få så stor uppmärksamhet att den gör så att en fråga uppfattas som annorlunda än innan.

Ovanstående är några exempel på olika variationer. De visar att "den perfekta frågan" är en orimlig förväntning. Detta ska varken ses som något nedslående eller invecklat utan visar bara på de förutsättningar som frågekonstruktören arbetar efter. Snarare än att söka "den perfekta frågan" kan målet istället vara att göra så goda bedömningar och bra kompromisser som möjligt.

Goda och dåliga exempel

En ganska vanlig uppfattning som man stöter på är att böcker och kurser om frågekonstruktion ägnar mycket mer tid åt att beskriva vad som är fel än vad som är rätt. Det är alltså vanligare att utgå från dåliga exempel än från goda. Denna bok lider väl i viss mån också av denna "negativa pedagogik" även om ambitionen har varit att i sådana fall alltid förklara vad som utgör ett problem och hur det kan åtgärdas. Men det finns skäl till att material om frågekonstruktion oftare utgår från fel än rätt. De viktigaste är:

1. **Det är enklare att avgöra att något inte fungerar än att avgöra att det faktiskt fungerar.** Ibland behövs bara lite enkel språklogik för att avgöra att en fråga inte kommer att fungera som tänkt. Om en fråga är obegriplig kommer inte respondenterna att kunna förstå eller besvara den korrekt. Detta är ofta förhållandevis entydigt och enkelt att avgöra. För att avgöra om en fråga faktiskt fungerar enligt

grundreglerna ovan behövs oftast djupare utredningar, till exempel provintervjuer eller analyser av svaren på frågan. Detta beror på att många av faktorerna i grundreglerna är okända på förhand (till exempel hur respondenterna kommer att tolka frågan). Utifrån forskningsresultat och vår erfarenhet kan vi göra goda bedömningar om att en fråga kommer att fungera bra, men det är inte lika enkelt och entydigt som att avgöra om något *inte* kommer att fungera.

2. **Dåligt är alltid dåligt men bra är oftast bara bra ibland.** De dåliga exemplen innehåller oftast problem som gör att de alltid är dåliga, oavsett sammanhang eller förutsättningar. Bra exempel är däremot oftast bara bra i ett givet sammanhang och inte nödvändigtvis i ett annat. Detta hänger ihop med resonemanget om "den perfekta frågan" under förra rubriken. En fråga kan till exempel vara utmärkt för att mäta en viss aspekt av något i en viss typ av undersökning för en viss respondentgrupp, men kanske inte lika bra under andra förutsättningar. Det kan alltså finnas risker med att lansera en fråga som bra och felfri.

- Det finns ett antal grundregler för frågekonstruktion, bland annat att frågornas ska förstås på samma sätt av alla respondenter.
- Det finns en "illusion om den perfekta frågan" men den kan vara svår att uppnå i praktiken på grund av att "det varierar".

Att skriva frågor i praktiken

Vetenskap eller hantverk?

En av de första böckerna om frågekonstruktion är "The art of asking questions" av Stanley Payne (Payne, 1951). Titeln signalerar tydligt att författaren anser att frågekonstruktion är en konst eller ett hantverk. Drygt femtio år senare släppte Schaeffer och Presser (2003) en artikel med namnet "The science of asking questions". Titeln var en passning till Paynes bok men med förändringen att frågekonstruktion nu är en vetenskap. Så vad är frågekonstruktion – en vetenskap eller ett hantverk? Sanningen är nog att det är bådadera.

Till skillnad från när Payne skrev sin bok finns det idag ganska mycket forskning om frågekonstruktion. De resultaten är förstås ett bra stöd när

man skriver sina frågor. De flesta frågekonstruktörer kommer dock sannolikt att hamna i situationer där de forskningsresultat som finns är svåra att tillämpa eller går stick i stäv. I sådana situationer måste frågekonstruktören förlita sig på sin erfarenhet, sina språkfärdigheter eller sin känsla av hur respondenterna kommer att tolka innebörden av frågan. Den här typen av bedömningar är inre processer som är svårare att observera och beskriva vetenskapligt utan mer har formen av konst eller ett hantverk.

Konflikter och kompromisser

Frågekonstruktion är oftast ett grupparbete mellan flera personer som dessutom har skilda kompetenser eller roller. Exakt vilka kompetenser som ingår varierar från fall till fall och mellan olika organisationer. Vanligt är att arbetet med frågorna innehåller

- uppdragsgivare eller beställare av undersökningen
- ämneskompetens
- statistisk metodkompetens
- administrativ kompetens.

Dessa kompetenser kan representeras av en person vardera men oftare är fler personer inblandade. Exempelvis kan en undersökning innehålla mer än ett ämne och metodkompetensen kan vara uppdelad i statistiker och frågekonstruktörer.

I samarbetet kan det förstås uppstå konflikter mellan olika infallsvinklar. Kanske kan den bästa layoutlösningen för en fråga inte lösas produktionsmässigt? Kanske håller inte frågekonstruktören med om hur ämnesspecialisten vill introducera ämnet i frågeformuläret? Detta är något som de flesta frågekonstruktörer säkert redan har erfarenhet av, om inte bör man vara beredd på det. Frågorna i ett frågeformulär är viktiga men de är inte allt. Fråge- och formulärkonstruktion är ett samarbete där det kan uppstå konflikter och krävas kompromisser, precis som vid all annan typ av samarbete.

Problem i praktiken

Fråge- och formulärkonstruktion är alltså ett samarbete. Det finns ett antal situationer som kan uppstå och som kan vara bra att tänka på i förhand. Därmed inte sagt att de är vanliga men de kan vara bra att förebygga så de inte uppstår:

Fel person skriver frågorna

Det händer att någon annan än frågekonstruktören, till exempel ämnes-specialisten, skriver frågorna. Eftersom specialisten kan ämnet som frågorna ska handla om kan det tyckas naturligt att denne också skriver frågorna. Men om inte specialisten har erfarenhet av och kunskap i frågekonstruktion bör förstås frågeskrivandet kompletteras med den kompetensen. Bara för att man kan ämnet skriver man inte nödvändigtvis bra frågor. Båda kompetenserna – ämne och frågekonstruktion – är nödvändiga för att få fram bra frågor.

Frågekonstruktionen underskattas

Eftersom de flesta kan läsa och skriva är det lätt hänt att en del även ser frågekonstruktion som något naturligt och självklart. Precis som den här boken är ett bevis på finns det tvärtom mycket att lära sig och tänka på vid frågekonstruktion. Därmed inte sagt att personer utan utbildning eller erfarenhet nödvändigtvis skriver dåliga frågor. Personer kan ha goda färdigheter för frågor (till exempel bra språkkänsla) utan att ha läst en enda bok i ämnet. Men sannolikheten för bra frågor är högre om personen dessutom har erfarenhet och kunskap om frågekonstruktion. Det viktiga är att momentet inte bagatelliseras.

Frågorna börjar skrivas för tidigt

Undersökningars syften är oftast brett fastställda inledningsvis, till exempel att en undersökning ska handla om "trivseln på arbetsplatsen". Det är en ganska lång resa från sådana breda beskrivningar till specifika frågor i frågeformuläret. Vad menas med trivsel och vilka delar består det av? En risk är att gå för snabbt igenom den här processen och börja för tidigt med att skriva frågor. I diskussioner om frågorna kan det senare framgå att man egentligen inte förtydligat tillräckligt exakt vad frågorna ska syfta till. Då måste man gå tillbaka och det kanske det inte finns tid för i det läget. Det är alltså mycket bättre att göra allt förarbete på en gång och inte börja skriva frågorna förrän man bestämt vad de ska syfta till.

Frågorna utvärderas inte eller för sent

Denna punkt hänger delvis ihop med punkt två ovan. Om man tycker att frågorna "känns bra" kanske det inte framstår som så nödvändigt att testa dem innan de tas i bruk. Då tar man en risk som kan vara ganska stor. En rekommendation är att alltid testa nya frågor. Även småskaliga, informella prov på kollegor eller familjemedlemmar kan ge värdefull information. En annan risk är att man visserligen testat frågorna men så

pass sent att det i praktiken inte finns utrymme att göra förändringar. För att kunna ta hand om testresultatet på ett bra sätt krävs därför god planering.

Färdiga frågor ges stor betydelse

Att använda redan färdiga frågor har många fördelar. Man slipper arbetet med att formulera nya frågor, man kan jämföra undersökningens resultat med andra undersökningar och om frågorna har använts tidigare har man förmodligen fått indikationer på hur de fungerat. Men man bör inte vara okritisk mot färdiga frågor. Deras bäst före-datum kan ha gått ut eller de kan passa mindre bra för den undersökning man själv tänkt göra.

Ifrågasättande av frågorna

Att sätta samman ett frågeformulär är även det ett grupparbete. Det är många som kan ha synpunkter på hur frågorna är skrivna. Sådana diskussioner kan ofta vara roliga, intressanta och givande. Frågekonstruktören bör likväl vara beredd med argument och fakta för sina val. Det är också bra om diskussionerna inte fastnar för mycket i sådant som handlar mer om tycke och smak än det meningsbärande i frågorna. Ska man till exempel skriva "till exempel" eller "t ex"? Det går att diskutera men sannolikt har det inte så stor betydelse för själva mätningen. Det är ofta mer givande om diskussioner kring frågorna fokuserar på det substantiella snarare än det formella.

Målkonflikter för mätningen

Även vad gäller själva mätningen kan det uppstå konflikter mellan olika mål vilket i sin tur påverkar frågekonstruktionen. Kanske riskerar den bästa formuleringen att leda till för hög respondentbörda? Kanske den lämpligaste formuleringen avviker för mycket från frågorna i andra undersökningar, eller tidigare insamlingar av samma undersökning, som man vill jämföra med? Det finns inte alltid självklara val vad gäller den typen av målkonflikter utan det handlar om hur man bedömer och värderar de olika målen. Det kan dock vara bra att fastställa från början om till exempel jämförbarhet med andra undersökningar är en prioritet.

Sammanfattande punkter:

- Att skriva frågor är både ett hantverk och en vetenskap.
- I praktiken är frågeskrivandet ofta ett samarbete mellan olika kompetenser.

- I samarbetet kan det uppstå konflikter vad gäller mål, frågeformuleringar, om frågorna ska utvärderas eller i synen på tidigare använda frågor.

Att tänka på respondenterna

Vilka är respondenterna?

Respondenterna är de som svarar på frågorna i frågeformuläret. När man läser böcker och artiklar om frågekonstruktion används beteckningarna "respondent" och "respondenterna" ofta. Det kan vara lätt hänt att omedvetet få intrycket att respondenterna skulle vara en speciell grupp men det är naturligtvis ingen grupp det handlar om utan en *roll*. Vem som helst kan få den rollen, till exempel beroende på hur undersökningars urval dras. Som frågekonstruktör kan det vara bra att påminna sig själv om detta, bland annat för att motverka en mekaniserad syn på respondentrollen.

Respondenters förhållningssätt till frågorna

Respondenter tolkar och besvarar inte alltid frågorna på det sätt som frågekonstruktören föreställt sig "vid skrivbordet". Detta är viktigt att känna till när man arbetar med att konstruera frågor. Nedanstående punkter diskuteras mer utförligt på andra ställen i boken och är endast sammanfattade här:

- Att svara på en enkät är oftast **ingen prioritet för respondenten**. De som ansvarar för enkäten tycker naturligtvis att den är viktig och att ämnet är angeläget men det tycker inte nödvändigtvis respondenterna i samma grad.
- Det finns alltid respondenter som försöker **avsluta enkäten med minsta möjliga ansträngning**. Generellt sett kan man dock utgå från att de flesta respondenter försöker förstå och besvara frågorna korrekt.
- Om en fråga innehåller vaga ord kan visserligen vissa respondenter reagera med "det beror på" men många kommer nog att, utan att reflektera över saken, **utgå från den tolkning som ligger nära till hands** för dem där och då (Clark & Schober, 1992).
- **Svarsalternativ hjälper till att avgränsa vad som efterfrågas** i frågan och respondenterna använder den informationen, ibland till och med för att bättre förstå vad frågan syftar på eller för att få "förslag" på

möjliga svar. Det är en av anledningarna till att öppna och slutna frågor kan ge så olika svar (Clark & Schober, 1992).

- **Respondenterna svarar generellt på frågorna** även om de egentligen kanske inte har informationen som efterfrågas. Att man fått svar på en fråga betyder alltså inte att man frågat något som respondenterna faktiskt kan svara på. Respondenterna kan istället för att besvara frågan på det sätt som egentligen krävs till exempel utgå från hur det oftast är, vad som är normalt, en enskild händelse snarare än alla relevanta tillfällen eller göra en uppskattning (Clark & Schober, 1992).
- Respondenterna utgår från att frågekonstruktören har ett syfte bakom att en fråga är ställd på just det sätt den är ställd. Det **inkluderar sådant som frågekonstruktören kanske egentligen inte haft någon djupare tanke bakom**, till exempel att olika beteckningar används för samma sak i olika frågor.
- Respondenterna förstår inte frågorna var och en för sig som om de vore isolerade från varandra. Att besvara ett frågeformulär kan tvärtom ses som en "kumulativ" process – **föregående fråga påverkar hur nästa förstås** (Clark & Schober, 1992).
Frågekonstruktören kan delvis avhjälpa detta genom att visa när det är dags att "börja på nytt", till exempel genom att introducera ett nytt avsnitt.

Att värna om respondenterna

Även om mycket av forskningen om frågekonstruktion har haft ett mätperspektiv är det naturligtvis också viktigt att tänka på respondenterna. Frågekonstruktören bör ha ett respondentperspektiv (Wenemark, 2010; Wenemark, 2017). Frågor ska inte bidra med mer börda än nödvändigt för respondenten. Respondenten ska heller inte komma till skada och helst inte bli irriterad eller upprörd över frågornas formulering eller innehåll. Allra bäst är det om respondenterna själva tycker att besvarandet är meningsfullt och att de kan känna att de har fått ge en rättvisande bild. Detta är mycket viktigt – dels av etiska skäl men det ligger också i frågekonstruktörens egenintresse. Att personer vill delta som respondenter är en nödvändighet för surveyundersökningar. Annars blir det inga svar på frågorna. Respondentrollen är alltså en resurs som man inte ska ta för given utan tvärtom värna om. Ett sätt att praktisera detta perspektiv är att tänka på respondenterna, inte som en otydlig och osynlig grupp, utan som individer.

Sammanfattande punkter:

- Att vara respondent är en roll.
- Frågekonstruktören bör reflektera över respondenters förhållningssätt till frågor i praktiken (se punktlista ovan).
- Det är viktigt att tänka på och värna om respondenterna. Bördan ska vara låg och respondenterna ska inte komma till skada och helst inte bli irriterande.

Praktiska språktips

Språket är det verktyg vi använder för att förmedla vad det är vi vill att respondenterna ska ge för information. Kapitel 7 i denna bok handlar om språket men syftar främst till att ge förståelse för språket som kommunikationsverktyg. Som frågekonstruktör är man förstas också behjälpt av mer handfasta och praktiskt inriktade språktips. Det finns många källor med värdefulla tips om hur man kan skriva (t.ex. Myndigheternas skrivregler (Språkrådet, 2014)). En del av källorna fokuserar på längre texter och innehåller en del som inte direkt är relevant för frågekonstruktion (till exempel om styckeindelning). Men ofta finns det många tips som även är tillämpbara i frågeformulär.

Nedanstående tips är inspirerade av Langlet & Wärneryd (1990) , Myndigheternas skrivregler (Språkrådet, 2014) och SCB:s interna riktlinjer.

Vem är mottagare?

Man bör alltid fundera på vem som är mottagaren till texten. Vem ska läsa och begripa den? I frågeformulär är förstas respondenterna mottagare men vilka är de och, framförallt, vad kännetecknar deras språkbruk? Här är svaret allt som oftast att "det varierar".

Respondenterna varierar naturligtvis, som alla andra, vad gäller till exempel ordförråd. Ett mål kan vara att använda ord som är "minsta gemensamma nämnare" i betydelsen att alla respondenter förstår dem. Man bör sträva efter att förenkla frågor så långt som möjligt. Dock bör man vara uppmärksam på att inte innebörden ändras om man ersätter ord med enklare synonymer.

Om frågorna innehåller vardagliga ämnen kan kanske ens egna erfarenheter av vad som är enkla och svåra ord vara tillräcklig för att skriva ett första utkast på frågor. Men vid mer specifika ämnesområden

kan man behöva undersöka vad som är lämpliga ord och termer att använda, speciellt om undersökningen har en specifik målgrupp.

Svåra ord

Ett generellt tips är att svåra ord bör ersättas med enklare synonymer. Vad är då ett svårt ord? Det kan vara ord som är ålderdomliga eller ovanliga så att alla respondenter därför inte känner till dem. Svarta listan (Regeringskansliet, 2011) är en bra förteckning över sådana ord och innehåller även förslag på alternativ. Synonymböcker och synonymfunktionen i ordbehandlingsprogram kan också vara till stor hjälp för att ersätta svåra ord.

Facktermer kan också vara svåra att förstå för de som inte är insatta i ämnesområdet. Om undersökningar riktar sig till en specifik grupp som kan förväntas känna till dessa termer finns det naturligtvis inga hinder för att använda dem. Men om man vänder sig till allmänheten bör facktermer undvikas om det går.

I vissa fall kan det dock vara besvärligt att ersätta ett svårt ord utan att till exempel förståelsen försvåras för andra respondenter som kände till ordet. Det kan även finnas andra skäl till att det ibland kan vara svårt att byta ut ett ord eller en term. Det man då bör göra är att tydligt definiera betydelsen och se till att den är lättåtkomlig.

Undvik förkortningar

Förkortningar gör texten kortare men oftast svårare att läsa. Detta gäller speciellt för förkortningar som används inom fackområden men som inte är kända för allmänheten. Även om förkortningen definieras gör sådana förkortningar texten mer svårläst och det stör förståelsen om respondenten måste påminna sig om vad förkortningen betyder.

Även mer språkliga förkortningar bör undvikas. Vissa av dem är välkända och konsekvensen av att använda blir då inte nödvändigtvis så stora (till exempel "d.v.s." eller "m.m.") men det säkraste är att skriva ut dem.

Enda undantaget till att undvika förkortningar är de fall där förkortningen är mer känd än det som förkortningen står för. Ett sådant exempel kan vara LO och Landsorganisationen.

Tänk på längden på meningarna

Långa meningar är ofta mer komplicerade. Det blir mer information för läsaren att hålla reda på. Risken för att begå något språkligt misstag som frågekonstruktör ökar också vid mer komplicerade satser. Långa, komplicerade meningar kan med fördel delas upp i flera kortare. Att bara ha korta meningar kan dock ge ett styligt intryck. Det bästa är oftast att blanda långa och korta meningar.

Undvik mångtydiga ord

Vissa ord har flera betydelser. Sådana ord bör undvikas eftersom det försvårar förståelsen. Respondenten måste själv avgöra vilken av flera möjliga tolkningar som är mest rimlig. Ibland kan det till och med vara omöjligt att avgöra eftersom flera tolkningar kan vara lika giltiga. Ett exempel på ett mångtydigt ord är "då" som kan betyda både "när" och "eftersom".

EXEMPEL:

Redovisning behövs inte då fokusgrupper ska genomföras

Det går i ovanstående exempel inte att avgöra vilken av de två tolkningarna ("då" eller "eftersom") som är rätt utan man måste tolka med hjälp av sammanhanget eller sina förkunskaper, vilket gör texten svårförståelig. Mångtydiga ord bör därför undvikas.

Intervjufrågor

Intervjufrågor läses upp av en intervjuare men är nästan alltid författade i skrift från början. Här kan skillnaderna mellan tal- och skriftspråk bli tydliga. Intervjufrågor är ofta inte tillräckligt anpassade till talspråk från början utan det är något man måste arbeta med. Saker som framstått som naturliga och problemfria på papper kan upplevas som för formella, krystade eller tjatiga när man läser högt.

Intervjufrågor bör därför alltid läsas högt på prov. Ett enkelt sätt är att läsa dem högt på egen hand för sig själv men allra bäst är att genomföra hela intervjuer.

Testa din text

Det är svårt att bedöma sin egen text, bland annat eftersom man naturligtvis vet vad man själv menar och syftar på i texten. Därför bör man testa sin text på andra innan den används. Allra bäst är det om man

kan testa texten på personer inom den tänkta målgruppen. Att testa sina enkätfrågor på en kollega kan ge värdefull information men om kollegan är insatt i undersökningen och enkäten är det inte jämförbart med att låta en potentiell respondent läsa frågorna.

I ordbehandlingsprogram kan man också testa texter och få ut statistik om till exempel antal ord per mening eller antalet meningar per stycke. Microsoft Word redovisar också läsbarhetsstatistik som kan ge en indikation på svårighetsgraden på texten. Riktmärken kan vara att skönlitterär text brukar ligga på 30-40. Facktexter ligger högre men värden över 60 brukar anses som svåra, avancerade texter.

Dela upp istället för uppdel

Att dela upp sammansatta ord förenklar generellt texten eftersom orden blir kortare och enklare.

Substantivsjukan

Böcker om språktips brukar varna för "substantivsjukan". Den innebär att man i texten omvandlar verb (till exempel "förbättra") till substantiv (till exempel "en förbättring"). Resultatet blir oftast att texten inte bara blir längre, eftersom man måste lägga till fler ord, utan dessutom svårare.

EXEMPEL:

Vi ska förbättra verksamheten (verb)

Vi ska genomföra en förbättring av verksamheten (substantiv)

Detta är alltså så pass vanligt att det fått sin egen beteckning (substantivsjukan) och man bör ha det i åtanke när man skriver och granskar texter. För att inte försvåra texten – undvik att göra om verb till substantiv.

Undvik passiv form på verbet

I den passiva formen ("en undersökning genomfördes") framgår det inte vem som gjort något. Det blir både otydligt och tyngre att läsa. Det är bättre att använda den aktiva formen som "SCB genomförde en undersökning".

De flesta av tipsen i detta avsnitt gäller texter i allmänhet och är inte specifika för just frågor i frågeformulär. Men avsnittet är en naturlig

brygga till nästa kapitel om frågeteknik, som också innehåller en del språktips men enbart ur ett frågeperspektiv.

Sammanfattande punkter:

- Tänk på vem som är mottagare.
- Undvik svåra ord.
- Undvik förkortningar.
- Tänk på längden på meningarna och blanda gärna.
- Undvik mångtydiga ord.
- Prova intervjufrågor genom att läsa dem högt.
- Testa din text på andra och i ordbehandlingsprogram.
- Att dela upp sammansatta ord förenklar generellt texten.
- Undvik att göra om verb till substantiv.
- Undvik passiv form på verbet.

Sammanfattning

Att skriva frågor – sammanfattning**Vad är en fråga?**

- Siffrorna i en undersökning baseras ofta på språkliga förfrågningar till respondenter – frågorna är viktiga!
- En fråga definieras som *en språkhandling riktad till någon med målet att få information som respons*.
- Frågesatsen är den vanligaste formen att ställa frågor på.
- Frågesatsen kan delas upp i ja-nej-frågor och frågeordsfrågor (inleds till exempel med hur, var, när, varför).
- Frågor kan också ställas som uppmaningar, påståenden eller med hjälp av nyckelord.
- Frågesatsen är den vanligaste (och bästa) formen att ställa frågor på. Påståenden är ganska vanliga i undersökningar riktade till individer och hushåll. Uppmaningar är vanligare i undersökningar till företag och organisationer.

- Det kan vara lättare att hålla med än att ta avstånd från påståenden.
- Vid faktabetonade frågor kan man överväga om påståendena kan omvandlas till frågesatser.
- Frågor kan också delas in i typer baserat på vad de syftar till att mäta, t.ex. faktafrågor, beteendefrågor och attitydfrågor.

Frågornas mål:

- Det finns ett antal grundregler för frågekonstruktion, bland annat att frågornas ska förstås på samma sätt av alla respondenter.
- Det finns en "illusion om den perfekta frågan" men den kan vara svår att uppnå i praktiken på grund av att "det varierar".

Att skriva frågor i praktiken:

- Att skriva frågor är både ett hantverk och en vetenskap.
- I praktiken är frågeskrivandet ofta ett samarbete mellan olika kompetenser.
- I samarbetet kan det uppstå konflikter vad gäller mål, frågeformuleringar, om frågorna ska utvärderas eller i synen på tidigare använda frågor.

Att tänka på respondenterna:

- Att vara respondent är en roll.
- Frågekonstruktören bör reflektera över respondenters förhållningssätt till frågor i praktiken (se punktlista ovan).
- Det är viktigt att tänka på och värna om respondenterna. Bördan ska vara låg och respondenterna ska inte komma till skada och helst inte bli irriterande.

Praktiska språktips:

- Tänk på vem som är mottagare.
- Undvik svåra ord.
- Undvik förkortningar.
- Tänk på längden på meningarna och blanda gärna.
- Undvik mångtydiga ord.
- Prova intervjufrågor genom att läsa dem högt.

- Testa din text på andra och i ordbehandlingsprogram.
- Att dela upp sammansatta ord förenklar generellt texten.
- Undvik att göra om verb till substantiv.
- Undvik passiv form på verbet.

2 Frågeteknik

Författare: Andreas Persson (SCB)

Övergripande mål för frågekonstruktion

Detta kapitel handlar om frågekonstruktion och ligger på en praktisk nivå med verkliga exempel som utgångspunkt. Det redogör för tips som är gemensamma och kan vara aktuella för många olika frågetyper. Kapitlet ger exempel på fallgropar man bör undvika och tumregler som är bra att följa. Vissa av tipsen bygger på forskning, andra på erfarenhet. Flera av dem återfinns i många böcker som inkluderar avsnitt om frågekonstruktion (till exempel Trost, 1994; Dahmström, 2005; Hansagi & Allebeck, 1994; Wärneryd, 1990). En del av tipsen kan tyckas uppenbara men det är ett delvis förrädisk intryck. Alla är vanliga i frågeformulär och det krävs uppmärksamhet för att upptäcka dem och kunskap för att korrigera dem.

Två frågor i samma fråga

En grundregel i frågekonstruktion är att en fråga endast ska fråga om ett ämne. Om denna regel överträds i frågekonstruktionen kan det vålla problem både för respondenterna och för dem som ska analysera svaren. En klassificering av problem i 2 000 frågor visade att denna typ av problem förekom i 10 procent av fallen (Belson, 1981). Feltypen måste således betraktas som väldigt vanlig (Davidsson & Wärneryd, 1990).

EXEMPEL:

Tycker du att lärare och elever ska ha längre vinterlov?

Tror du att mjölk- och äggpriserna nästa år kommer att bli högre eller lägre än nu?

Exemplen ovan (Davidsson & Wärneryd, 1990) innehåller två frågor i respektive frågeformulering. Den första frågar om både lärare och elever. Den andra om både mjölk- och äggpriser. Detta kan vålla problem för respondenter som kanske håller med om eller uppfyller den ena delen av frågan (till exempel att elever ska ha längre lov) men inte den andra (att lärarna ska ha det). Det kan leda till att de hoppar över frågan, vilket resulterar i partiellt bortfall. Det kan även leda till att de markerar något

som inte stämmer (till exempel att de svarar "Ja" trots att det bara stämmer för eleverna i frågan, inte för lärarna), vilket ger ett mätfel. Frågornas svar blir därmed också svåranalyserade. Vad betyder "Ja" och "Nej"? Det är svårt att veta eftersom frågan innehåller mer än en fråga.

Det kan ibland vara svårt att avgöra om en fråga verkligen innehåller flera frågor eller inte. Man kan ju tycka att frågan "Tycker du att elever och lärare ska ha längre vinterlov?" endast är *en* fråga och syftar på kombinationen av elever och lärare. Det kan likväl bli problematiskt för de respondenter som bara håller med om det ena och inte det andra.

Hur ska detta problem åtgärdas? Den enklaste lösningen är att dela upp frågan i två frågor – en fråga för åsikten om lärares vinterlov och en annan fråga för åsikten om elevernas. Ett alternativ är att lägga in svarsalternativ som täcker in frågeformuleringens två frågor, till exempel "Ja, men bara för eleverna". Ibland kan det dock vara så att det är just kombinationen av flera saker som är syftet med frågan, till exempel inställning till *både* x och y. I de fallen måste frågekonstruktören tydligt kommunicera att det är just kombinationen som avses genom att betona det i språket, till exempel lägga till "både" i frågan.

Sammanfattande punkter:

- Ställ inte mer än en fråga åt gången.
- Problemet med flera frågor i samma fråga kan enklast lösas genom att dela upp det på flera frågor.
- Om det är kombinationen av flera saker som åsyftas i frågan måste det tydligt framgå i frågeformuleringen, till exempel genom att skriva "både x och y".

Inbyggda förutsättningar

Dold förutsättning i frågan

En annan grundregel i frågekonstruktion är att frågan inte ska ta några förutsättningar för givna. Det kan vara frestande att gå rakt på sak i sin frågeformulering utan att först ha utrett om respondenterna uppfyller de grundläggande förutsättningarna.

EXEMPEL:

Vilka arbetstider har du?

Ovanstående fråga har minst ett inbyggt antagande. Den förutsätter att respondenten har ett arbete och kanske hyfsat regelbundna arbetstider. Om inte dessa antaganden stämmer blir det problem för respondenten. Det kan leda till att respondenten hoppar över frågan (partiellt bortfall) eller svarar som om antagandet är giltigt trots att det inte är det (mätfel).

Problemet löses enklast genom att tillföra ett svarsalternativ som korrigerar antagandet och erkänner att det kanske inte gäller, till exempel "Jag har inget arbete". Om det dolda antagandet endast gäller en fråga är det sannolikt den bästa lösningen. Om antagandet gäller fler frågor, till exempel ett helt frågebatteri, kan det vara bättre att ha en inledande filterfråga (till exempel "Har du något arbete?") med hoppanvisningar. Det separerar de respondenter som de följande frågorna är aktuella för från dem som de inte är aktuella för (läs mer om filterfrågor nedan).

Explicit förutsättning i frågan

En annan variant är när antagandet är explicit uttryckt i frågeformuleringen. Det betyder att frågeformuleringen ger en förutsättning och sedan ställer en fråga utifrån den förutsättningen. Det här problemet påminner mycket om att ha två frågor i samma fråga (se föregående avsnitt). Det kan ibland vara svårt att veta om respondenten svarat på frågan eller på förutsättningen. På motsvarande vis kan det vara svårt för respondenterna att ta ställning till frågan om till exempel den beskrivna förutsättningen inte gäller.

EXEMPEL:

Har du varit i kontakt med oss senaste året – hur upplevde du servicen?

Mycket bra
Bra
Varken eller
Dålig
Mycket dålig

Om man inte haft någon kontakt (förutsättningen i exemplet ovan) kan man inte bedöma servicen.

En annan typ av exempel med förutsättningar:

EXEMPEL:

Har det under det senaste året hänt att du avstått från att ge dig ut på kvällen av oro för att bli överfallen, rånad eller på annat sätt ofredad?

Ja

Nej

Ovanstående fråga har flera villkor inbyggda och uttryckta i frågan – att respondenten avstått från att ge sig ut, att det skett på kvällen, att det är på grund av oro och att oron baseras på risken att bli överfallen, rånad eller ofredad. Det är viktigt att vara uppmärksam på att respondenten kan ta olika vägar vid dessa villkor. Vissa av villkoren kan stämma för respondenten, andra inte. Respondenten kan till exempel ha avstått från att ge sig ut men av andra skäl än oro. På motsvarande sätt kan respondenten vara orolig men ändå ge sig ut. Man får anta att de som svarar "Ja" uppfyller alla förutsättningar. Men man bör också vara uppmärksam på att "Nej" kan innehålla olika typer av svar som kan uppfylla vissa av frågans villkor men inte alla, till exempel att de är oroliga men ändå gett sig ut. Ett alternativt sätt att ställa denna fråga är att först fråga vilka som avstått och sedan fråga om skälen (Davidsson & Wärneryd, 1990).

Ett klassiskt exempel på hur fel det kan bli med explicita förutsättningar i frågeformuleringen kommer från opinionsinstitutet The Harris Poll (Fråga rätt, 2001).

EXEMPEL:

Har du ofta, ibland, nästan aldrig eller aldrig haft skuld känslor därför att du varit otrogen mot din fru?

På frågan svarade 1 procent "Ofta", 14 procent "Ibland" eller "Nästan aldrig" och 85 procent att de "Aldrig" hade skuld känslor på grund av sin otrohet. Eller vad var det de svarade på? Frågan förutsätter otrohet och utifrån det frågar den om skuld känslor. Förutom att frågor av den här typen kan vålla huvudbry och irritation går det inte heller att använda resultatet eftersom man inte vet vad respondenterna svarat på.

Inbyggda förutsättningar kan alltså orsaka problem i en fråga. Problemet kan lösas genom en filterfråga eller med hjälp av svarsalternativen.

Sammanfattande punkter:

- Om frågan tar vissa förutsättningar för givna ställer det till problem för de respondenter som inte uppfyller dem (kan leda till både bortfall och mätfel).
- De inbyggda förutsättningarna kan vara dolda i den betydelsen att de inte uttrycks direkt i frågeformuleringen.
- De inbyggda förutsättningarna kan också vara explicita genom att frågan först konstaterar något (som kanske inte stämmer för respondenten) och sedan frågar något utifrån det.
- Problemet löses genom att antingen erbjuda ett svarsalternativ för dem som förutsättningarna inte gäller för eller inleda med en filterfråga.

Definiera frågan i tid och rum

Tidsmässiga utgångspunkter i frågan

Vissa frågor är tidslösa i den bemärkelsen att svaret inte påverkas av när de ställs. "I vilket land är du född?" är ett exempel på en sådan fråga. I andra typer av frågor är kanske inte tidpunkten relevant i sig. Frågans syfte kan vara att mäta *om* något hänt, inte när det hände. "Har du någon gång legat inlagd på sjukhus i minst en vecka?" är ett exempel på en sådan fråga. I många fall är dock inte frågorna helt fria utan behöver avgränsas i tid och rum. Annars kommer respondenterna att göra det på egen hand, medvetet eller omedvetet, men troligtvis olika sinsemellan och inte nödvändigtvis på samma sätt som frågekonstruktören tänkt sig.

Att mäta ett "normalläge"

Hur ska man då definiera frågan i tid? Detta behandlas utförligt i kapitel 11, Att fråga om det förflutna, och nämns här endast översiktligt. Ibland kan tidsbestämningen vara avsiktligt obestämd. Detta kan vara aktuellt när det inte är någon specifik mätperiod som avses utan frågan syftar till att mäta något slags "normalläge" – hur det oftast är. Vanligt är då att frågan inkluderar specificeringar som "vanligtvis", "brukar" eller "generellt" för att förtydliga avsikten med frågan. Dessa specificeringar är viktiga eftersom respondenten annars kan bli förbryllad om frågan avser någon särskild tidsperiod eller inte. Respondenten uppmanas alltså att göra ett överslag. Ibland kan detta vara fullt tillräckligt för det syfte man har med frågorna. Men det kan finnas risker om det som ska mätas

varierar mycket (till exempel mellan årstider eller mellan veckodagar och helger). I sådana fall finns det kanske inte något "normalläge" som frågan gör anspråk på.

Att mäta en viss period – en "referensperiod"

I andra fall krävs det att tidsperioden är bestämd och specifik. Respondenten förväntas då svara för en begränsad och avgränsad period, en så kallad referensperiod. Det finns framförallt två typer av problem som kan uppstå vad gäller en frågas referensperiod. Det ena är om referensperioden är tillräckligt tydlig. Vad menas till exempel med "under senaste året"? Är det kalenderår eller de senaste tolv månaderna som avses? Detta är ett exempel på en referensperiod som kan tyckas självklar vid första anblicken men som alltså kan vara ottydlig för respondenterna.

Den andra typen av problem har att göra med om referensperiodens längd är rimlig, givet det som ska mätas. Generellt sett kan man säga att händelser som är framträdande, betydelsefulla och ovanliga är lättare att minnas längre tillbaka i tiden. Exempel på sådana händelser är större inköp, sjukhusvistelser och centrala livshändelser (studenten, giftermål, nytt jobb etc.). Den här typen av händelser kan alltså personer minnas på en bra nivå flera år tillbaka och frågor om sådana saker kan därför ha en längre referensperiod.

För mer vardagliga och triviala händelser (till exempel veckoinköp eller motionerande) behövs en kortare referensperiod på högst någon eller några veckor tillbaka.

"Rumsliga" utgångspunkter i frågan

Med att definiera frågorna "i rum" menas att det ska framgå vilka andra utgångspunkter än tidsmässiga som respondenten ska ha. Ska de till exempel svara för sig själva eller för hela hushållet? Detta är också viktigt att definiera på en tillräckligt specifik nivå.

EXEMPEL:

Hur länge har du bott här? (Fråga rätt, 2001)

Ovanstående fråga saknar rumslig förankring eftersom "här" kan betyda många saker. Ska respondenterna till exempel svara för sin bostad eller

för grannskapet, staden eller kommunen? Det måste förtydligas för att svaren ska bli meningsfulla.

Sammanfattande punkter:

- De tidsmässiga utgångspunkterna i frågan ska framgå för respondenten.
- Om det är ett "normalläge" som åsyftas i frågan bör det tydliggöras genom tillägg som till exempel "vanligtvis", "oftast" eller "generellt sett".
- Om frågan syftar på en specifik period (en så kallad referensperiod) ska den vara tydlig och rimlig, givet det som ska mätas.
- Framträdande händelser kan ha en längre referensperiod (sträcka sig flera år bakåt i tiden).
- Triviala händelser bör ha en kort referensperiod (vecka/veckor).

Ledande frågor

Frågeformuleringen i sig kan påverka resultatet i en viss riktning och mot en viss typ av svar snarare än andra. Detta är naturligtvis inte önskvärt. Respondenterna ska reagera på innehållet i frågan, inte på formuleringen i sig (Davidsson & Wärneryd, 1990). Detta kallas för att frågan är ledande och kan orsaka missvisande svar. Om resultatet används som opinionsbildning eller beslutsunderlag kan konsekvenserna bli allvarliga. Problemet är vanligast vid attityd- och åsiktsfrågor men kan även förekomma vid andra typer av frågor. Respondenterna kan också reagera negativt på att besvara ledande frågor, speciellt om de har en annan åsikt än den som frågan leder mot. Att en fråga är ledande behöver dock inte vara ett utslag av att frågekonstruktören har en missledande agenda. Det är tvärtom lätt hänt att frågor är oavsiktligt ledande.

Värdeladdat innehåll

En form av ledande frågor är när frågeformuleringen innehåller beskrivande ord som är icke-neutrala och som tar ställning i något avseende.

EXEMPEL:

Vad tycker du om de nya kränkande reglerna mot rökning på arbetsplatsen?

”Kränkande” är frågekonstruktörens egen värdering av reglerna. Det är en partisk beskrivning. Respondenterna kan påverkas av och svara i enlighet med den. Denna risk borde vara störst om det är ämnen som respondenten inte är så insatt i. Genom värderingen avslöjar också frågekonstruktören sin egen åsikt i ämnet. Respondenterna kan instämma i den, oavsett vad de egentligen tycker, för att vara ”goda respondenter” eller för att undvika en ansträngd situation. Det sistnämnda gäller framförallt i intervjuer där svaret ges till en annan människa och där det dessutom kan framstå som om det är intervjuarens egen åsikt som uttrycks.

Ett annat sätt att ladda frågan är att plantera en norm, till exempel hur de flesta människor känner, tänker eller gör.

EXEMPEL:

De flesta föräldrar tycker det är viktigt att barnen håller på med någon idrott, tycker du att det är viktigt eller mindre viktigt?

Exemplet ovan sätter en norm för vad de flesta föräldrar tycker och vad som kan uppfattas som ”normalt”. För att svara ”mindre viktigt” måste respondenterna avvika från den normen och svara att de är annorlunda, vilket en del personer drar sig för.

Ett annat sätt är att anspela på auktoriteter, till exempel ”Framstående forskare anser numera...”. Konsekvensen är densamma som normen i det förra exemplet – frågan ger tyngd åt ena riktningen och för att svara i andra riktningen måste respondenterna gå emot vad ”framstående forskare anser” (Fråga rätt, 2001).

Alla dessa frågor är exempel på värdeladdat innehåll i frågeformuleringen som man bör undvika.

Frågans riktning och balans

En annan form av ledande fråga är när frågeformuleringen lutar åt ett håll snarare än ett annat. Exempelvis kan bara den ena sidan av sakfrågan tas upp i frågeformuleringen medan den andra sidan inte nämns och därför blir avvikande. Detta är ganska vanligt förekommande i frågeformulär. Exempelvis frågor av typen ”Tycker du att statsministern, generellt sett, sköter sin uppgift på ett bra sätt?” eller ”Är du nöjd med...”. Det kan vara lättare att svara ”ja” än ”nej” på sådana frågor. Respondenterna kan också känna att de måste motivera ett ”nej”

(Davidsson & Wärneryd, 1990). Risken är störst vid frågor där respondenten kanske inte har ett givet svar eller är okoncentrerad. En lösning är att ha med båda perspektiven i frågeformuleringen, till exempel "Är du nöjd eller missnöjd med..." eller "Tycker du staten borde en införa en lag mot... eller är du motståndare till en sådan lag?". Dessa formuleringar leder inte respondenten i någon speciell riktning utan respondenten måste ta ställning på egen hand. De formuleringarna främjar också ett byte av svarsskalan från ja-och-nej till en svarsskala med fler skalsteg. Det kan vara en fördel eftersom åsikter ofta är mer nyanserade än vad en ja-och-nej-skala tillåter.

En lite svagare variant av ledande frågor är när båda riktningarna visserligen är inkluderade i frågeformuleringen, men inte i samma omfattning.

EXEMPEL:

Tycker du att Sverige ska införa euron som ny valuta eller inte?

Frågan ovan inkluderar båda riktningarna men den ena riktningen ("eller inte") består av ett påhäng på slutet. I litteraturen kallas detta för minimal balansering – jämför med full balansering: "Tycker du att Sverige ska införa euron som ny valuta eller behålla den svenska kronan?" Är då minimal balansering tillräckligt för att undvika problemet med ledande frågeformuleringar? En minimalt balanserad fråga blir kortare men å andra sidan blir det ena perspektivet inte lika tydligt framträdande. Det viktiga är om det blir skillnader i svar mellan helt och minimalt balanserade frågor. Enligt en studie (Schaeffer, Krosnick, Langer & Merkle, 2005) blir det inga skillnader mellan de två frågetyperna. Även om det behövs mer forskning indikerar detta att minimalt balanserade frågor inte nödvändigtvis är någon halvmesyr mellan balanserade och obalanserade frågor utan kan vara en acceptabel lösning.

En annan variant är frågor som visserligen redovisar båda riktningarna, men där beskrivningen av den ena riktningen är mer utförlig och inkluderar skäl eller motiveringar.

EXEMPEL:

Ska vi stoppa kärnkraften för att hindra framtida katastrofer eller ska vi bygga ut? (Davidsson & Wärneryd, 1990)

Ovanstående fråga inkluderar ett skäl ("hindra framtida katastrofer", vilket i sig är värdeladdat, se tidigare i avsnittet) för den ena riktningen (stoppa kärnkraften), men inte för den andra riktningen (bygga ut). Frågeformuleringen blir då obalanserad. Om skäl ska ingå i frågeformuleringen bör alla sidorna av sakfrågan vara representerade i samma grad (Davidsson & Wärneryd, 1990).

Pseudofrågor

Ledande frågor kan vara ett sätt att, medvetet eller omedvetet, styra svaren. Även om frågeformuleringen är balanserad kan det faktum att frågan ställs på ett förenklat sett leda till ett missvisande resultat. Det brukar kallas pseudobudskap när ett budskap endast innehåller självklarheter som ingen kan tycka annorlunda om. På motsvarande sätt kan frågeformuleringar ta upp ämnen på ett så ytligt sätt att få personer kan tycka annorlunda, givet att konkurrerande alternativ, baksidor, nackdelar med mera inte täcks av undersökningens frågor.

EXEMPEL:

Tycker du att regeringen ska satsa mer eller mindre på välfärden?

Ovanstående fråga är inget problem formuleringsmässigt. Men utan sammanhang och detaljnivå, till exempel flera frågor eller beskrivningar om resursfördelningar och prioriteringar, kan det vara enkelt för en respondent att bara hålla med – det är bra om det görs mer. Verkligheten är dock ofta komplex, fylld av kompromissande situationer och det finns inte sällan konflikter om till exempel resurser. Om frågeformuläret inte erkänner detta kan resultatet ge en förenklad och felaktig bild av respondenternas åsikter. Om frågeformuläret bara innehåller frågor om vissa aspekter och inte heller mäter hur viktiga respondenten tycker att dessa är, kan resultatet alltså ge en skev bild (läs mer i kapitel 12).

Ledande svarsalternativ

Svarsalternativen är en del av frågan och ska, precis som frågeformuleringen, vara balanserade och neutrala. De ska inte innehålla värde-laddade ord eller vara fler för den ena riktningen än för den andra.

Ett problem vid åsiktsfrågor är att det kan vara lättare att hålla med om än att ta avstånd från något. Det gäller särskilt om respondenten är okoncentrerad, omotiverad eller inte har någon stark åsikt eller kunskap

om frågeområdet. Denna tendens är framförallt identifierad vid påståenden som använder svarsskalor baserade på "Instämmer helt-Instämmer inte alls" eller motsvarande (kallas för "acquiescence" på engelska) (läs mer i kapitel 3).

Sammanfattande punkter:

- En ledande fråga är när frågeformuleringen i sig (snarare än innehållet i frågan) leder respondenterna till ett visst snarare än ett annat svar.
- En fråga kan vara ledande genom att använda värdeladdade och icke-neutrale ord.
- Att anspela på auktoriteter (till exempel forskare eller läkare) är också ledande.
- Att plantera en norm för vad andra tycker eller gör är också ledande (till exempel "de allra flesta tycker si, vad tycker du?").
- Att bara redovisa ena sidan av en sakfråga kan vara ledande.
- En s.k. "minimal balansering", att ha "eller inte" i slutet av frågan, kan vara ett fungerande sätt att göra en ledande fråga neutral.
- Att ställa frågor om komplicerade sakfrågor på ett sådant förenklat sätt att få kan tycka annorlunda är också ledande, även om frågeformuleringen i sig kan vara neutral.
- Svartalternativen kan vara ledande på samma sätt som frågeformuleringen.
- Svartalternativen kan också vara ledande genom att den ena sidan representeras av fler svartalternativ än den andra sidan.

Filter- och följdfrågor

Ibland kan det vara nödvändigt att dela upp respondenterna i olika grupper (så kallade filter) som besvarar olika delar av frågeformuläret. Det kan förekomma i två typer:

Filterfrågor

Filterfrågor fungerar som vattendelare i frågeformuläret. Beroende på hur respondenterna svarar på filterfrågan slussas de vidare till olika frågor i frågeformuläret.

EXEMPEL:**Har du arbetat övertid den senaste veckan?**

Ja → Gå till fråga 22

Nej → Gå till fråga 44

Följdfrågor

Följdfrågor följer på andra frågor men ska endast besvaras om vissa villkor uppfylls. Exempel: Fråga 2 nedan besvaras bara om respondenten svarat "Ja" på fråga 1.

EXEMPEL:**1. Står du som ägare till någon bil?**

Ja

Nej

2. Om ja:**Hur många bilar står du som ägare till?**

Filter fyller en viktig funktion i ett frågeformulär. De gör det möjligt att ställa frågor som inte är relevanta för alla respondenter. Om man inte har filter riskerar man att de som frågan inte är relevant för ändå svarar på den (mätfel), eller till och med blir så konfunderade eller irriterade att de avbryter sitt deltagande (objektbortfall). Filter bör föregå de frågor som inte är relevanta för alla respondenter.

Filter är dock inte helt oproblematiska i pappersenkäter. De medför att respondenten måste bryta mönstret att svara på frågeformuläret kontinuerligt fråga-för-fråga för att istället hitta rätt väg genom frågeformuläret på egen hand. I datorstödda insamlingsätt (telefonintervjuer eller webbenkäter) är detta inget problem. I webbenkäter kan filtren programmeras och i telefonintervjuer är det intervjuaren som med hjälp av ett datorstöd administrerar frågorna. I en pappersenkät finns det dock risk för fel vid filterfrågor och det finns därför skäl att försöka begränsa antalet filter.

Ska filterfrågor eller följdfrågor användas? Om det är fler än en följdfråga som följer på en huvudfråga bör filterfrågor användas, eftersom det annars blir otydligt vad till exempel "Om ja" syftar på. Om det är endast en följdfråga kan dock andra exemplet ovan användas, speciellt om

frågeformuläret har många filter i övrigt (läs mer om layout vid filterfrågor i kapitel 5).

I ett frågeformulär som har flera filter finns det två möjliga sätt att strukturera uppföljningsfrågorna. Det är antingen direkt efter respektive filterfråga eller samlat på slutet efter att alla filterfrågor ställts. Det förstnämnda kan kännas mest naturligt och passar ofta bäst ämnesmässigt, men forskningen visar att det kan ge sämre svarskvalitet. Anledningen verkar vara att vissa respondenter uppmärksammar att filterfrågorna ger uppföljningsfrågor och därför kan svara nekande på filterfrågorna i syfte att förkorta svarstiden (Eckman, Kreuter, Kirchner, Jäckle, Tourangeau & Presser, 2014). Det kan alltså vara värt att fundera på hur filtren i formuläret bäst ska struktureras.

Sammanfattande punkter:

- Ett frågeformulär innehåller ibland frågor som inte är relevanta för alla respondenter utan bara för dem som svarat på ett visst sätt.
- En lösning är filterfrågor – ett slags förgreningsfråga – som separerar respondentgrupperna och leder dem vidare till respektive aktuell plats i frågeformuläret.
- Ett annat sätt är följdfrågor som har inbyggda villkor – att de bara ska besvaras om respondenten svarade "Ja" på frågan innan.
- Följdfrågor fungerar bra om det endast är en fråga, annars är filterfrågor bäst. I pappersenkäter, där respondenten sköter hoppen själv, finns det en del risker med filterfrågor. De bör därför inte vara för många i en enkät.

Frågeformuleringens innehåll

Negationer i frågan

Att undvika negationer (som t.ex. "ej", "icke" eller "inte") i frågor är en gammal tumregel inom frågekonstruktion. Detta gäller särskilt i frågor vars svarsskalor också har negationer. Respondenten måste då nämligen ta ställning till dubbla negationer när de ska svara, vilket är en ganska krävande uppgift.

EXEMPEL:

Vilka av följande försäkringstyper hade du inte under förra året?

Oavsett svarsskalan finns det ett värde i att undvika negationer. Negationer har nämligen en kraftfull funktion i språket – de vänder upp och ned på betydelsen i satsen eller på betydelsen i frågan. Om man tillför en negation så får frågan motsatt betydelse jämfört med om den stått utan negation. Frågor med negationer är därför ofta mer krävande att läsa, åtminstone i de fall där negationen kommer sent i frågan. Den största risken är dock att betydelsen av satsen står och faller med ett ord. Om respondenten läser lite slarvigt kan de missa negationen och tolka frågan som att den betyder motsatsen mot vad den faktiskt gör.

Det finns alltså flera skäl till att försöka undvika negationer.

Ords ordningsföljd och förståelse av frågan

Grammatiken styr var ord ska vara placerade i en sats eller i en fråga. Men man kan även spjälka upp innehållet i en fråga i lite grövre beståndsdelar, vilka i sig kan ha olika ordningsföljd. Som frågekonstruktör är det värt att tänka på detta.

EXEMPEL:

Hur nöjd eller missnöjd är du med de kurser du gick hos oss under 2015 vad gäller undervisningen?

Ovanstående fråga kan sägas bestå av följande delar:

- Frågeansats ("Hur nöjd eller missnöjd"): Detta är inriktningen på frågan och svarsdimensionen i vilken svaret ges.
- Frågeobjekt ("kurser du gick hos oss"): Det som efterfrågas och som respondenten ska utvärdera.
- Tidsmässig avgränsning ("under 2015"): Den tidsperiod som frågan gäller.
- Frågeobjektets avgränsning ("vad gäller undervisningen"): En avgränsning och precisering av vilka delar av frågeobjektet som just denna fråga avser.

Det finns ingen etablerad nomenklatur för hur man bäst ska dela upp ett frågeinnehåll. Frågan i exemplet kan spjälkas upp på andra sätt än ovan men det finns helt klart en poäng i att tänka kring frågor på det här sättet. Om man gör det upptäcker man att frågeinnehållet kan struktureras på

olika sätt inom frågan och ha olika ordningsföljder. Nedan är några exempel med endast små språkliga korrigeringar:

- *Under 2015, hur nöjd eller missnöjd är du med undervisningen i de kurser du gick hos oss?*
- *Hur nöjd eller missnöjd är du med undervisningen i de kurser du gick hos oss under 2015?*
- *I de kurser du gick hos oss 2015, hur nöjd eller missnöjd är du vad gäller undervisningen?*
- *Vad gäller undervisningen på de kurser du gick hos oss under 2015, hur nöjd eller missnöjd är du?*

Förutsatt att man inte gör något grammatiskt misstag så ska betydelsen vara densamma oavsett hur vi disponerar frågeinnehållet. Men olika ordningsföljder kan påverka hur lätt eller svår frågan är att förstå. Som frågekonstruktör handlar det om att försöka avgränsa respondentens tänkande i rätt ordning – att strukturera innehållet i den följd som är lättast för respondenten att förstå. Även om det inte finns någon generell lösning för detta kan vissa sekvenser av innehållet vara lättare att förstå än andra. Exempelvis är det ofta bra att introducera frågeobjektet innan man introducerar avgränsningar för frågeobjektet. Trattprincipen, att gå från brett till smalt, är alltså ett bra rättesnöre.

Ovanstående är dock breda tumregler. Vad som är den lämpligaste ordningen kan variera mellan olika frågor och frågeformulär kanske främst på grund av respondentens förkunskap och förväntningar, baserade på till exempel tidigare besvarade frågor. Om respondenten exempelvis redan besvarat flera frågor om hur nöjd eller missnöjd han eller hon var med olika aspekter av kurserna under 2015, så bör kanske avgränsningen lyftas fram i frågan eftersom det är det unika med just den här frågan, jämfört med tidigare frågor.

Det finns ofta mycket att vinna på att ha ett bredare perspektiv än de enskilda frågorna. Om exempelvis hela frågeformuläret handlar om 2015 så behöver inte den tidsmässiga avgränsningen ha en särskilt framträdande roll i frågan eller kanske inte ens vara med. Om man har flera frågor som delar mycket av frågeinnehållet är det ofta en god idé att skapa en gemensam lösning för dem och lyfta fram det som gör respektive fråga unik:

EXEMPEL:

I de kurser du gick hos oss under 2015, hur nöjd eller missnöjd är du med...

... undervisningen?

... schemat?

... biblioteket?

Sammanfattande punkter:

- En fråga består av olika typer av innehåll, t.ex. olika typer av avgränsningar såsom referensperiod och specificeringar.
- Det är ofta inte givet i vilken ordning de olika innehållen bäst ska komma men det är värt att reflektera över.

Exempel i frågor

Det är ganska vanligt att frågor innehåller preciseringar i syfte att avgränsa frågans innehåll och ge frågan en mer specifik inriktning.

EXEMPEL:

I de kurser du gick hos oss 2015, hur nöjd eller missnöjd är du vad gäller undervisningen?

”Vad gäller undervisningen” i frågan ovan är en sådan precisering och styr respondenten till att tänka på och svara för just det, snarare än för kurserna i allmänhet.

Att inkludera exempel i en frågeformulering kan ha en liknande men ändå tydligt annorlunda funktion. Exempel är inte en uttömmande lista över allt som ska ingå i frågan utan snarare förslag på vad som kan ingå i frågan.

För att förklara lite djupare vad ett ”exempel” är kan man utgå från ett kategorisystem där man kan sortera föremål eller företeelser från generell till specifik (läs mer i kapitel 7). Ett exempel på ett sådant system är:

Djur → Däggdjur → Katt

”Däggdjur” är mer specifikt än ”djur”, och ”katt” är mer specifik än ”däggdjur”. Alla däggdjur är dock inte katter utan ”katt” är endast en underkategori till den högre kategorin ”däggdjur”. Katt kan sägas vara

ett exempel på ett däggdjur – en underkategori av många möjliga till en högre och mer generell kategori.

Vad är då syftet med exempel i frågeformuleringar? Det är alltså *inte* att respondenterna bara ska svara för exemplen. Det är inte heller att exemplen ska bli för styrande över svaren. Men detta är precis vad som kan hända och är en risk när man använder exempel.

Syftena med exempel kan vara

- att påminna respondenterna om fall som de annars förmodligen inte skulle ha tänkt på (se nedan). När detta är syftet kan exemplen utgöras av ovanliga underkategorier som respondenten troligtvis inte skulle ha tänkt på annars
- att förtydliga eller illustrera den högre kategorin. Kanske blir "motion" lite för vagt i frågeformuleringen och man exemplifierar därför med "cykling", "löpning" och "gym" – några av underkategorierna till den högre kategorin "motion"
- att sätta fart på minnet. Genom att ge exempel kan man få respondenten att börja tänka och förhoppningen är att tänkandet "sprids" även till annat än just exemplen.

Om inget av ovanstående syften är aktuellt behöver frågan kanske inga exempel. Många frågor fungerar mycket bra utan exempel. I praktiken utgör frågor med exempel undantag snarare än regel.

Det finns inte så mycket forskning om exempel i frågeformuleringar. Den forskning som finns visar bland annat att exempel kan påverka svaren när de ändrar vad respondenten kommer att tänka på och om detta (t.ex. gym) skiljer sig på svarsdimensionen (t.ex. hur många gånger i veckan) mot vad de annars skulle ha utgått ifrån (t.ex. löpning). Bra exempel är därför de som är relevanta för respondenten och som han eller hon annars inte skulle ha tänkt på (Tourangeau, Conrad, Couper & Ye, 2014).

Om den högre kategorin är vag och otydlig för respondenterna (som i exemplet "motion" ovan) kan säkert illustrerande exempel bidra till en bättre förståelse för vad som efterfrågas. Men om den är tydlig för respondenterna behövs inte illustrerande exempel.

Var dock uppmärksam på att en fråga med endast ovanliga exempel kan göra respondenten osäker på vad som egentligen menas. Då kan det vara bättre att blanda vanliga och ovanliga företeelser i exemplen eller att man delar upp det i flera frågor, till exempel att man frågar om den ovanliga

företeelsen (som man befarar att respondenten kan missa) i en enskild fråga.

Sammanfattande punkter:

- Exempel i frågor kan hjälpa respondenten att förstå vad som menas med t.ex. "motion" eller påminna dem om saker som de annars inte skulle ha kommit ihåg.
- Om inte ovanstående funktioner är nödvändiga så behöver frågan antagligen inte heller några exempel.
- Undvik att bara ha "ovanliga exempel" eftersom det kan förvirra.

Korta eller långa beskrivningar?

Att ordval kan påverka svaren är väl dokumenterat. Samtidigt kan företeelser beskrivas på olika sätt, med olika ord och olika många ord. En studie undersökte om allmänhetens åsikter om statens utgifter på olika områden påverkades av om områdena beskrevs med en kort etikett (till exempel "miljön") eller en längre beskrivning ("förbättra och skydda miljön"). Den längre beskrivningen resulterade i en större andel som tyckte att staten spenderade för lite på områdena. Vad kan det bero på? En anledning kan vara att en längre beskrivning ger annorlunda och kanske mer konkreta associationer hos respondenterna. En annan delförklaring kan vara att de längre beskrivningarna i den här studien i flera fall också innefattade en förstärkning. De innehöll positiva värdeord som "förbättra" och "skydda" (Rasinski, 1989).

I ovanstående exempel inkluderade den längre frågan mer meningsbärande information och blev därför, på sätt och vis, en annan fråga än den korta. Men frågor kan även förlängas utan att tillföra meningsbärande information genom att till exempel lägga till en introduktion till frågan. Jämför exemplen nedan (Davidsson & Wärneryd, 1990).

EXEMPEL:

Hur har ni skaffat den bostad ni har nu?

Nästa fråga handlar om anskaffandet av bostäder. Vi är intresserade av att veta något om detta. Hur har ni skaffat den bostad ni har nu?

Den längre varianten är inget att rekommendera för enkäter, generellt sett, men för intervjuer har faktiskt sådana, längre frågor gett bättre resultat. Förmodligen för att de kan öka respondentens motivation och ge denne mer betänktetid (Davidsson & Wärneryd, 1990). Dock ökar även intervjutiden så det är ingen universallösning men den kan användas i specifika fall.

Sammanfattande punkter:

- Längden på frågeformuleringen kan påverka svaren även om betydelsen är densamma. Det beror bland annat på att en längre frågeformulering innehåller fler ord som kan ge andra associationer än dem i en kort frågeformulering men också på att respondenten (i en intervju) får mer betänktetid.

För eller emot? Förbjuda eller tillåta?

Det finns en förbjud-tillåt-paradox i äldre forskning om frågekonstruktion. Paradoxen innefattar att frågor med "förbjud" och "tillåt" inte är varandras motsatser såsom de borde vara logiskt sett. Det klassiska exemplet är nedanstående frågor från 1940-talet:

EXEMPEL:

Tycker du att USA borde förbjuda offentliga tal mot demokrati?

Tycker du att USA borde tillåta offentliga tal mot demokrati?

Andelen som svarar "Ja" på den ena frågan borde vara densamma som svarar "Nej" på den andra frågan, men så är det alltså inte. Senare forskning har studerat det vidare och funnit liknande resultat även med andra frågor. Ibland kan skillnaderna vara så pass stora som 30 procentenheter (Schuman & Presser, 1981). Hippler & Schwarz (1986) föreslår att paradoxen kan bero på att de som är likgiltiga inför frågan svarar "Nej" på båda varianterna och därmed orsakar förskjutningen i svarsproportioner. Varför skulle då de som är likgiltiga eller neutrala svara "Nej" istället för "Ja"? Hippler och Schwarz (1986) menar att både "förbjuda" och "tillåta" är aktiva handlingar medan "inte tillåta" eller "inte förbjuda" mer är ett uttryck för att avstå från att aktivt ta ställning. Författarna tar även upp en annan kognitiv effekt som kan verka i samma riktning – att människor har lättare för att uppmärksamma och resonera kring konsekvenser av aktiva än passiva handlingar ("feature-

positive effect” på engelska). Det skulle då kunna medföra att respondenterna inte uppmärksammar att frågornas implikation är detsamma (”förbjuda” jämfört med ”inte-tillåta”). Studiens resultat stödjer resonemanget att det är respondenterna som är neutrala eller likgiltiga som svarar likadant på båda frågorna. Effekten borde därför bli större ju fler respondenter som är likgiltiga för det som efterfrågas (Hippler & Schwarz, 1986).

Hur ska man förhålla sig till detta som frågekonstruktör? Om ovanstående förklaring är korrekt är det ganska enkelt: man bör inkludera ett svarsalternativ för de respondenter som är neutrala. Frågor med förbjuda-tillåta är inte så vanliga i Sverige men möjligen kan resultaten beskrivna ovan generaliseras till andra, liknande frågor, till exempel frågor om att vara ”för” och ”emot” något.

Sammanfattande punkter:

- Förbjuda och tillåta (och möjligen för och emot) är inte nödvändigtvis varandras motsatser i frågekonstruktionens värld. Det kan bero på att de som är likgiltiga svarar nekande på båda frågorna.
- Problemet kan lösas genom att inkludera ”Vet inte” eller ”Ingen åsikt” som alternativ.

Vissa tycker si, andra tycker så

Vid känsliga eller kontroversiella frågor rekommenderas det ibland att försöka normalisera det känsliga i frågeformuleringen. Istället för att gå rakt på sak kan frågeformuleringen istället inledas med en mjuk introduktion som är tänkt att avvärja det hotfulla i frågan.

EXEMPEL:

Vissa personer tycker att Vietnamkriget var rättfärdigt, andra tycker att det var fel och omoraliskt. Tycker du det var rättfärdigt eller fel och omoraliskt?

Ovanstående exempel inleds med en mjuk introduktion av typen ”Vissa tycker si, andra tycker så”. Leder den typen av introduktioner till en förbättrad mätning, jämfört med en direkt fråga utan introduktion (Yeager & Krosnick, 2012)?

Syftet med introduktion är alltså att göra frågeämnet mindre hotfullt, men det kan även få andra konsekvenser: exempelvis att respondenten börjar tänka på vilka grupper som tycker vad eller att en respondent som tidigare tyckt att något var bortom all tvivel blir varse om att personer faktiskt kan tycka annorlunda. En metaanalys visar att "vissa tycker si, andra tycker så" inte ökade rapporteringen av okonventionella åsikter. Tvärtom var det relaterat till lägre validitet (läs mer i kapitel 8) jämfört med direkta frågor (Yeager & Krosnick, 2012).

En annan studie fann att avdramatiserande introduktioner inte hade någon huvudeffekt i sig. Däremot ledde de till högre rapportering av känsliga ämnen för yngre personer med personlighetstyper för vilka det är viktigt att framstå i god dager (personer med hög risk för social önskvärdhet). Även om denna typ av introduktioner inte tycks ha någon större effekt som helhet kan det alltså möjligen påverka vissa undergrupper (Peter & Valkenburg, 2011).

Det är dock viktigt att den här formen av introduktioner (oavsett om de är allmänna eller av den avdramatiserande typen) är tydligt relaterade till frågan men att frågan i sig är självständig och kan förstås utan introduktionen. Annars kan det stjälpa mer än hjälpa (Lenzner, Kaczmirek & Galesic, 2011).

Sammanfattande punkter:

- Att ha avdramatiserande introduktioner (vissa tycker si, andra tycker så) till känsliga frågor verkar generellt inte förbättra mätningen.

Visionära och generella begrepp

Ibland används enkäter för att utvärdera verksamhetsplaner, visionära målbeskrivningar eller andra generella styrdokument. Inte sällan överförs dokumentets nyckelord direkt till frågorna:

EXEMPEL:

Är du nöjd med utbildningens pedagogiska strategier, koncept och redskap?

Problemet är att styrdokument och liknande, ofta av nödvändighet, ligger på en mycket generell nivå för att kunna appliceras på många olika varianter av verksamheter. En grundregel i frågekonstruktion är dock att

vara specifik så att det inte råder någon tvekan om vad som avses. Ur det perspektivet fungerar sådana generella nyckelord ofta mycket dåligt. Respondenterna kommer säkert att besvara frågorna men det är oklart vad de egentligen har svarat på eftersom begreppen är så breda.

Sammanfattande punkter:

- När generella beskrivningar ska utvärderas, bryt ner de visionära ledorden till ord med mer precis betydelse. Annars blir det svårt att veta vad svaren baseras på.

Vanskliga ord

Det finns en del ord som är vanskliga att använda i frågekonstruktion. Trots det är de ganska populära och återfinns i många enkätfrågor.

Ett sådant ord är *tillräckligt*. Tillräckligt kan vara ett vanskligt ord eftersom det kräver en referenspunkt – tillräckligt jämfört med vad eller i vilket syfte? Om inte frågan är tydlig med vad som utgör referenspunkten kan det komma att variera mellan olika personer. Man kan förstås acceptera upplägget att låta respondenterna själva välja en referenspunkt men den blir då alltså okänd och varierar sannolikt mellan olika personer.

Andra exempel på vanskliga ord är "lagom" och "måttligt". Dessa typer av ord är inte bara kvantifieringar utan dessutom värderande. De beskriver inte bara en mängd utan relationen mellan en mängd och en förväntning, ett behov och så vidare, och går därför ofta lite på tvärs mot övriga alternativ i skattningsskalor.

Ett annat vanskligt ord är *regelbundet* eftersom det ofta används för att beskriva något med hög frekvens trots att ordet faktiskt inte säger något om frekvensen annat än att den är just regelbunden (återkommande med jämna intervall).

Sammanfattande punkter:

- Överväg användningen av vanskliga ord som "tillräckligt" och "regelbundet".

Synonymer

I vardagstext rekommenderas det ibland att använda synonymer för att variera språket. I enkäter är målet ett helt annat. Där finns få skönhetspoäng att plocka, utan det handlar om att vara så tydlig och konsekvent som möjligt.

Om till exempel en enkät växlar mellan begreppen "arbete", "sysselsättning" och "jobb" kan respondenten tolka det som att frågeställaren syftar på olika saker. Varför används nya beteckningar om det inte syftar på något nytt? Detta kan skapa förvirring hos respondenten och kan också leda till feltolkningar. Hur stor den risken är beror på hur mångtydiga begreppen är och hur van respondenten är vid dem.

Ännu värre är om begreppen är tänkta att vara synonymer men faktiskt inte är det. Till exempel kan begreppen "inte arbetar", "inte anställd", "arbetslös", "arbetsökande" syfta på olika saker. Att man "inte arbetar" betyder inte att man inte har en anställning (man kan vara tjänstledig, sjukskriven, föräldraledig etc.). Man kan vara arbetslös utan att söka arbete. Man kan också arbeta utan att vara anställd (till exempel som egen företagare). Och så vidare.

I enkäter finns det alltså inget egenvärde i att variera sig. Tvärtom bör man definiera centrala begrepp och sedan hålla sig till dem.

Sammanfattande punkter:

- I enkäter finns det inget egenvärde i att variera sig (använda synonymer). Tvärtom bör man definiera centrala begrepp och sedan hålla sig till dem.

X/Y

Ibland används två nyckelbegrepp, med snedstreck, i frågeformuleringen.

EXEMPEL:

Har du bytt jobb/anställning under 2016?

Syftet med formuleringen är sannolikt att förtydliga eller täcka in flera sätt att uttrycka sig. Det är inte nödvändigtvis något större fel att göra på det här sättet men det är mer svårläst och bör därför undvikas om det

går. Det kan dock uppstå problem om de två begreppen inte är synonyma (se föregående avsnitt). Jobb och anställning behöver till exempel inte betyda samma sak. Man kan till exempel byta inriktning (genom ett ansökningsförfarande) inom samma anställning och då se det som att man fått ett nytt jobb trots att det är samma anställning.

Alltså, om formen måste användas så se till att de två begreppen inte kan leda till olika svar, annars kan formuleringen stjälpa mer än hjälpa.

Sammanfattande punkter:

- Det kan bli fel för respondenten om två begrepp är tänkta att vara synonyma men inte är det i praktiken.

Att tänka på med frågor som mätinstrument

Hypotetiska frågor

Hypotetiska frågor uppmanar respondenten att ta ställning till en icke-existerande situation. Exempelvis vad man ska göra i framtiden eller om man vann på lotto. Hypotetiska frågor bryter, på sätt och vis, mot principen att man inte ska fråga om sådant som respondenterna inte har någon information om.

EXEMPEL:

Om du blev arbetslös, hur ofta skulle du kontakta (gå till kontor, ringa, besöka webbsida) Arbetsförmedlingen?

Givetvis finns det inget sant svar om framtida eller alternativa situationer utan svaret blir en subjektiv bedömning från respondentens sida.

Forskning i psykologi visar dessutom att individer ofta är rätt dåliga på att förutsäga sitt eget beteende i framtiden (Myers, 2002). Det ligger en del i talesättet "Det blir aldrig som man tänkt sig". Det beror bland annat på svårigheten att uppskatta alla de faktorer som är inblandade i en ny situation. Detta gäller speciellt obekanta situationer. Det är lättare om det är en situation som respondenten har erfarenhet av sedan tidigare och då borde också svaren bli mer korrekta. Det betyder att en person som varit arbetslös tidigare förmodligen kan ge en bättre uppskattning på exempelvis frågan än någon som aldrig varit det.

En del textböcker avråder från hypotetiska frågor (till exempel ESS, 2006) just på grund av ovanstående svagheter. Men hur man ser på hypotetiska

frågor beror också på hur svaren ska användas. Som prognoser för hur det faktiskt kommer att bli är hypotetiska frågor ofta ett trubbigt instrument. Däremot kan de användas för att få en uppfattning om hur respondenterna just nu tror att det kommer bli, till exempel vad som kommer inträffa i framtiden.

Orsakssamband

Det är inte något fel i att fråga respondenterna om orsakerna till deras beteende eller situation. Det är dock viktigt att särskilja om man frågar efter medvetna motiv eller bakomliggande orsaker till att något inträffat. Det förstnämnda går bra att fråga om men det sistnämnda kan vara svårare. Människor har ofta inte något givet svar på frågor om bakomliggande orsaker och kanske inte heller har reflekterat över saken tidigare. Det kan dessutom finnas många orsaker till att något inträffade. En fråga kan sällan inkludera alla relevanta orsaker utan det blir ett urval. Många gånger ingår endast en orsak i frågeformuleringen (se nedan), vilken då kan fungera som ledande. Ofta är det bättre att analysera bakomliggande orsaker på egen hand med hjälp av experiment eller longitudinella datainsamlingar, än att fråga respondenterna själva om det (även om upprepande mätningar också har sina svagheter).

EXEMPEL:

Att du inte reste utomlands förra året, beror det på att du inte hade råd?

Ja

Nej

Hur väl frågor om orsaker når sitt mål beror dock på syftet med frågan. Om syftet är att mäta hur respondenterna själva ser på saken är förstås det bästa att fråga dem. Det beror också på hur komplext ämnesområdet är.

Sammanfattande punkter:

- Hypotetiska frågor ger inga bra prognoser och brukar avrådas. De kan dock fungera för att få insikt i hur respondenten bedömer saker här och nu.
- Att fråga om medvetna motiv går bra men att fråga om bakomliggande orsaker är svårare.

Mörkertal på grund av undvikande

Undersökningar kan innehålla frågor om respondenten oroar sig för eller har råkat ut för något negativt.

EXEMPEL:

Hur ofta blir du diskriminerad?

Hur ofta känner du dig orolig för att drabbas av en olycka?

Dessa är angelägna frågor och en surveyundersökning är en bra metod för att införskaffa den här typen av information. Man bör dock vara medveten om att individen är en aktiv agent i sitt eget liv och kan vara benägen (i den mån det går) att undvika situationer som upplevs som negativa. Exempelvis kan en respondent från en minoritetsgrupp undvika situationer som styrs av majoritetsgruppen för att slippa bli diskriminerad. Frågor av typen "hur ofta" ger därför inte hela bilden utan bör kompletteras med annan information (till exempel om hur ofta de befinner sig i de relevanta situationerna, om de aktivt undviker dem etc.).

Sammanfattande punkter:

- Individen är en aktiv agent i sitt eget liv och kan vara benägen (i den mån det går) att undvika situationer som upplevs som negativa (t.ex. att bli diskriminerad).
- Sådana frågor kan behöva kompletteras med annan information (till exempel om hur ofta de befinner sig i de relevanta situationerna, om de aktivt undviker dem etc.).

För vem svarar respondenten?

Respondenten är alltid en individ men ibland ingår individen även i något annat sammanhang, till exempel som representant för ett hushåll eller som anställd i ett företag. Det måste tydligt uttryckas, antingen i frågeformuleringen eller i en allmän introduktion till undersökningen vem individen svarar för – sig själv, hushållet, organisationen eller företaget?

Detta kan ibland lösas genom att explicit skriva ut till exempel hushållet eller företaget i frågeformuleringen. En annan lösning är att använda

plural i frågeformuleringen, som "ni" eller "era". Tänk på att respondenten kan vara en del av en organisation som har flera nivåer.

EXEMPEL:

Ökade ni er omsättning under 2014?

Ovanstående fråga använder pluralformen men det är likväl otydligt om respondenten ska svara för arbetsplatsen, företaget eller koncernen. Om det blir för krångligt att förtydliga sådant i frågeformuleringen kan man göra det i en introduktion istället (men gärna påminna om det då och då).

Sammanfattande punkter:

- Det måste tydligt framgå för vem respondenten svarar – sig själv, hushållet eller företaget?

Förändring – storlek och riktning

Frågeformulär försöker ofta samla in information om förändringar. Om inte longitudinella data finns kan tvärsnittsstudier ibland kompensera för detta genom att låta respondenten själv rapportera förändringarna sedan någon viss referenspunkt (till exempel "sedan förra året"). Det kan finnas en del problem med sådana frågor. Ett kan vara att respondenternas minne är otillräckligt och att de därför förvränger de faktiska förändringarna (läs mer i kapitel 11).

För att mäta förändringar är det alltså bäst att ha två mätpunkter. Då känner man till både de absoluta värdena och förändringen dem emellan – både storlek och riktning. Om man försöker ersätta detta med endast en fråga kan det uppstå problem, beroende på vad man vill mäta.

EXEMPEL:

Hur mår du idag jämfört med för ett år sedan?

Mycket bättre

Bättre

Sämre

Mycket sämre

Ovanstående fråga ger information om riktningen på förändringen – om personens hälsa blivit sämre eller bättre under det senaste året. Däremot

är utgångspunkten okänd och därmed känner vi inte till hur respondenten faktiskt mår idag. Respondenten kan till exempel må bättre än förra året men ändå inte må särskilt bra. Frågan fångar bara in förändringen men inte de absoluta nivåerna (hur respondenten faktiskt mår för ett år sedan och idag).

EXEMPEL:**Motsvarade kursen dina förväntningar?**

Väldigt mycket

Ganska mycket

Lite grann

Inte alls

Ovanstående fråga handlar, på sätt och vis, också om förändringar i synen på kursen innan och efter att respondenterna har gått kursen. Frågan mäter hur väl de två utgångslägena överensstämmer där avvikelser betyder att synen förändrats. Frågan kan dock egentligen inte ge svar på om respondenten var positiv eller negativ till kursen, vare sig före eller efter att ha gått den. Exempelvis kan respondenten ha haft negativa förväntningar på förhand som sedan bekräftades. Detta är viktigt att tänka på. Frågorna och svarsalternativen måste formuleras utifrån det syfte man har med dem.

Man får inte heller glömma bort att även om frågans syfte är att mäta förändringar måste den samtidigt också tillåta stabilitet. Det är ju inte säkert att det man frågar efter har förändrats. Till exempel kan ju respondenten må likadant idag som för ett år sedan (se exemplet ovan som inte har ett svarsalternativ för stabilitet, t.ex. "oförändrat" eller "likadant").

Sammanfattande punkter:

- Om frågor ska fånga förändringar, tänk på att både riktningen på förändringen och de absoluta nivåerna (före och efter) kan vara intressanta att känna till.

Ändrade uppfattningar och förväntningar

När man ställer frågor och vill mäta hur människor uppfattar olika företeelser måste man vara medveten om att uppfattningar kan förändras

på flera sätt. Detta kan bli tydligt vid upprepade mätningar. Ett exempel är när personer ska svara om något i sitt förflutna. Svaren kan ändras över tid trots att det förflutna naturligtvis inte påverkas av att tiden går. Detta kan bero på minnesfel men också på att människors *uppfattning* av det förflutna förändras. Våra minnen om till exempel händelser och de bakomliggande motiven till vårt handlande kan rekonstrueras, bland annat på grund av att vi själva förändras över tid och ser på vårt förflutna med nya ögon (Trost, 1994).

Ett annat exempel är om man vill göra förbättringar av arbetsmiljön på en arbetsplats eller i en skola. En vanlig procedur är att inleda med att göra mätningar av hur personerna upplever miljön. Sedan genomförs ett antal förbättringar av miljön och efter dessa mäts upplevelsen av miljön igen. Objektivt sett borde förstas den andra mätningen ge ett mer positivt resultat eftersom förbättringar utförts mellan mätningarna. Men om man mäter subjektiva upplevelser är det inte säkert att så blir fallet – personers förväntningar och krav på miljön kan nämligen också förändras över tid (de förbättringarna av miljön som gjordes kan i sig ha startat en sådan process). Man bör vara medveten om detta och fundera på om de subjektiva frågorna bör kompletteras med mer objektiva frågor eller med subjektiva frågor där respondenterna får göra jämförelser före och efter de utförda förbättringarna av miljön.

Sammanfattande punkter:

- Människor förändras och kan därför lämna olika svar vid olika tidpunkter på ett sätt som kan tyckas motsägelsefullt i jämförelse med objektiva mått.

Känner du till ...? Visste du att ...?**EXEMPEL:**

Känner du till att SCB producerar jämställdhetsstatistik?

Ja

Nej

Ovanstående typ av frågor är relativt vanligt förekommande men de är, på sätt och vis, en språklig paradox eftersom själva frågan upplyser om det den frågar efter. Till skillnad från typiska kunskapsfrågor (läs mer i

kapitel 10) ska respondenten i den här typen av frågor själv värdera sin kunskap, inte redovisa den. Hur giltig den värderingen blir borde bero på om ämnet är prestigefyllt eller känsligt. Vid områden som är känsliga eller där det finns normer bör nog upplägget undvikas eftersom det är så enkelt för respondenten att bara svara jakande för att framstå i bättre dager. Vid neutrala ämnen kan dock upplägget ibland vara det mest praktiska att använda.

Sammanfattande punkter:

- Frågor som "Visste du att ..." kan vara den mest praktiska lösningen men bör undvikas för områden som är känsliga eller där det finns normer.

Frågehierarkier, till exempel generell-specifik

Det är ganska vanligt i enkäter att ett område innehåller flera frågor och att de dessutom ligger på olika nivåer. Exempelvis att det ingår flera frågor om hur nöjd individen är med den fysiska arbetsmiljön, arbetstider, chefen, arbetskamrater med mera och därefter en generell fråga om hur nöjd individen är totalt sett. Ett mindre uppenbart exempel är att fråga om hur mycket tid respondenten lägger på skolarbete respektive läxor (läxor kan ses som en del av skolarbetet, en enhet i helheten).

Hur ska man organisera sådana här frågor? Från generell till specifik, eller tvärtom? Det beror lite på syftet. Om man vill ha en spontan, generell bedömning är det bäst att ställa den generella frågan först. Om man tvärtom vill att respondenterna ska ha några specifika aspekter i åtanke när de besvarar den generella frågan, är det bättre att inleda med de specifika frågorna.

Om det bara handlar om några få frågor, som i exemplet skolarbete-läxor, behöver man vara uppmärksam för då kan ordningsföljden ha extra stor betydelse för svaren på den generella frågan. Om frågan om läxor ställs först kan den efterkommande generella frågan uppfattas som att syfta på "skolarbete utöver läxor", eftersom respondenten redan svarat för läxor på frågan innan (Sudman, Bradburn & Schwarz, 1996). Uppskattningarna på den generella frågan blir då lägre jämfört med om den hade ställts först eftersom läxor är exkluderade. Denna tendens avtar dock om det är flera specifika frågor som föregår den generella frågan. Den kan även

motverkas i frågekonstruktionen genom att explicit lägga till exempelvis "inkludera även läxor" (läs mer i kapitel 13).

Sammanfattande punkter:

- Vill man ha en spontan, generell bedömning är det bäst att ställa den generella frågan först.
- Om man tvärtom vill att respondenterna ska ha några specifika aspekter i åtanke när de besvarar den generella frågan, är det bättre att inleda med de specifika frågorna.
- Om det bara är några få frågor kan ordningsföljden ha extra stor betydelse, speciellt för den generella frågan.

Frågekonstruktion och intervjuer

Allmänt

Det mesta av det som står i kapitlet är lika relevant för intervju som enkät men vid intervju finns det dessutom några saker till att tänka på. När insamling sker via intervju behöver frågekonstruktören ta hänsyn till intervjuarens roll i förhållande till frågorna:

- Frågan ska vara talspråksvänligt skriven så att intervjuaren kan läsa upp den och respondenten kan uppfatta frågans syfte via hörseln.
- Intervjuaren ska vara informerad om och förstå frågans syfte.
- Intervjuaren ska ha information om hur eventuella motfrågor från respondenten ska hanteras.
- Intervjuaren ska känna till vilken del av frågan som är obligatorisk (dvs. text som alltid ska läsas upp) och vilka delar som läses upp ibland (till exempel svarsalternativ i vissa frågor).

I telefonintervjuer måste alltid varje svarsalternativ i en så kallad checklista ("mark all that apply" på engelska) ställas som en separat delfråga som respondenten svarar "Ja" eller "Nej" på. Denna frågetyp är ganska vanlig och blir alltså annorlunda för enkäter och intervjuer. Detta är viktigt att ha i åtanke vid kombinerad insamling eller vid byte av insamlingsmetod. I en besöksintervju är det möjligt att hantera frågan som en checklista genom att använda svarskort.

I pappers- och webbenkäter kan tabeller användas om flera frågor har samma struktur. Tabellformen innebär allt som oftast att frågorna tar mindre plats i frågeformuläret. Den ger också respondenten intrycket av

att frågorna i tabellen hör ihop och ska besvaras i samma form. Detta behöver dock inte vara en fördel då det kan främja att respondenten besvarar alla frågorna enligt samma mönster. I intervjuer kan frågor formuleras i block som liknar tabeller för att kunna underlätta för både respondenter och intervjuare. Ett sådant block byggs upp genom att inleda med en introducerande text, till exempel "Nu kommer ett antal frågor om X". Därefter presenteras det som är gemensamt för alla frågor, som avgränsningar i tid och rum: "Jag vill att du tänker på de senaste tolv månaderna när du svarar." Därefter följer första frågan, som ställs i sin helhet inklusive avgränsningar i tid och rum samt eventuell uppläsning av svarsalternativ. I den andra eller tredje frågan i blocket kan vissa delar av texten läggas inom parentes för att visa för intervjuaren att den texten ska användas vid behov om intervju-situationen kräver det.

Det finns en risk att respondenten uppfattar frågor i påståendeform som att de uttrycker intervjuarens egna åsikter. För att undvika detta och annan förvirring bör påståenden i telefonintervjuer skrivas i andra person ("Du är nöjd med ditt arbete") istället för första person ("Jag är nöjd med mitt arbete"), som är praxis i självadministrerade datainsamlingsätt.

Avslutningsvis kan man i enkäter använda parenteser eller fet stil för att försvaga eller förstärka vissa delar av frågeformuleringen. Detta fungerar inte på samma sätt i intervjuer. Det måste vara tydligt för intervjuaren vad denne ska läsa upp. Det går att visa om någon del av frågan ska betonas men det måste också gå att läsa frågan på ett naturligt sätt.

Intervjuer och respondentens svar

I intervjuer finns det en del att tänka på i frågeformuleringen vad gäller hur respondenten kommer fram till ett svar på frågan (läs mer i kapitel 3). Två saker att tänka på är:

- Kommer respondentens spontana svar att passa i de underliggande svarsalternativen?
- Börjar respondenten svara på "rätt" ställe?

Dessa frågeställningar diskuteras under varsin underrubrik nedan.

Kommer respondentens svar på frågan att passa i de underliggande svarsalternativen?

Om svarsalternativen inte läses, kommer då respondentens svar att passa i de underliggande svarsalternativen? Det måste man ta reda på som frågekonstruktör. För det första, kommer frågan att förstås på rätt sätt om inte svarsalternativen läses upp? En respondent kan till exempel svara "När mina föräldrar skilde sig" som svar på frågan "När flyttade du hemifrån?" trots att frågeställaren hade tänkt sig ett årtal som svar. Frågan måste alltid förstås på rätt sätt om svarsalternativen inte ska läsas upp och kan därför behöva vara extra tydlig (genom att till exempel ersätta "när" med "vilket år").

För det andra, leder frågan respondenten till rätt svarsdimension? Detta är framförallt relevant vid skattningsskalor. Frågeställaren kanske hade tänkt sig en svarsskala från "Mycket nöjd" till "Mycket missnöjd" som svarsalternativ till frågan "Hur nöjd är du med ditt arbete?", men respondenterna kan komma att svara "Sådär", "Jodå" eller "Relativt". För att undvika detta måste svarsdimensionen framgå tydligare i frågan. Ibland går dock inte sådana här problem att lösa till fullo utan antingen får man acceptera dem eller läsa upp svarsalternativen (korta skattningsskalor är de svarsalternativ som passar bäst att läsa upp).

För det tredje, vid fristående svarsalternativ (inte skattningsskalor) kan man inte lägga ansvaret på respondenten att kunna förstå vilka de bakomliggande svarsalternativen är (till exempel på frågan "Varför började du studera på universitetet?"). Här måste frågekonstruktören på förhand lista ut vilka alternativ som är relevanta och försöka täcka in så många som möjligt. Dessutom måste intervjuaren vara insatt i frågan och kunna klassa respondentens spontana svar i något av de underliggande alternativen. Ju bättre utformade svarsalternativen är, desto enklare blir detta arbete för intervjuaren (desto mindre är då också risken för fel). Ibland kan det hända att det inte går att dela upp verkligheten på ett tillfredsställande sätt i ett antal svarsalternativ. Respondenten kanske ger ett utförligt svar som inte riktigt passar på något alternativ eller, ännu vanligare, att det passar in på flera alternativ. Intervjuarna måste då veta hur sådana svar ska hanteras. Får de markera flera alternativ? Ska de utreda saken vidare, få respondenten att bestämma sig för endast ett alternativ och så vidare.

Vid intervjufrågor där svarsalternativen inte ska läsas upp är ett praktiskt tips därför att granska dem som om svarsalternativen inte fanns (till

exempel genom att hålla för dem eller ta bort dem ur dokumentet) och sedan jämföra vad man kommer fram till med rådande svarsalternativ. Vid skillnader bör man fundera över hur frågan och svarsalternativen kan förbättras och vilka instruktioner intervjuarna behöver ha för att hantera frågan i praktiken.

Börjar respondenten svara på "rätt" ställe?

För intervjufrågor måste frågekonstruktören se till att respondenterna ska vara redo att avge ett svar när frågan är färdigläst. De ska däremot inte vara redo *innan* frågan är färdigläst (Fowler, 1995). Med "frågan" menas här både frågeformuleringen och svarsalternativen. Respondenten ska helt enkelt inte börja svara i förtid, innan intervjuaren gett all information.

En faktor som kan utlösa ett för tidigt svar hos respondenten är om själva frågan ligger tidigt i frågeformuleringen.

EXEMPEL:

Har du undersökts av en läkare under 2012? Med läkare avses ...

Frågor är nämligen något som aktiverar mottagaren. När man får en fråga svarar man per automatik. Om frågan kommer tidigt i frågeformuleringen kan det hända att respondenten börjar svara innan hela frågeformuleringen är färdigläst. I exemplet ovan kan alltså respondenten börja svara direkt efter frågan ("Har du undersökts av en läkare under 2012?") och därmed missa definitionen ("Med läkare avses ..."). Detta kan leda både till felaktiga svar och till oönskade avbrott i intervjun (Beatty, Fowler & Cosenza, 2007). Därför kan frågan behöva komma sist i frågeformuleringen. Om frågeformuleringen ska innehålla definitioner eller om svarsalternativen ska läsas bör detta alltså helst föregå själva frågan.

Ovanstående exempel var en ja-nej-fråga för vilka svarsalternativen är givna (ja eller nej). Ett annat relaterat problem är annars att respondenten börjar svara när frågan är färdigläst men *innan* intervjuaren hinner läsa svarsalternativen. Hur stor denna risk är beror på

- hur enkel och självklar frågan är, det vill säga hur mycket betänktetid respondenten behöver
- hur intervjuaren läser, det vill säga hur lång paus som görs mellan frågeformulering och svarsalternativ.

Det finns några tekniker för att minska risken. Dels kan man i frågeformuleringen tala om att det kommer att komma svarsalternativ ("Vilka av följande ..."). Ett annat sätt, främst för att minska punkt 2 ovan, är att bygga in en brygga (i kursiv stil i exemplet nedan) i slutet av frågeformuleringen som på ett naturligt sätt, i talspråk, leder vidare till svarsalternativen. Denna lösning tydliggör för respondenten att det kommer att komma mer information och hjälper också intervjuaren med läsningen.

EXEMPEL:

Hur ofta åt du på restaurang förra månaden? Var det ...

Varje dag

Några gånger i veckan

Någon gång i månaden

Aldrig

Sammanfattande punkter:

- En intervjufråga måste vara skriven så att den fungerar att läsas och intervjuaren måste känna till hur frågan ska hanteras i praktiken.
- Om inte intervjuaren läser upp svarsalternativen måste frågan vara konstruerad så att den ändå leder respondenten till de underliggande svarsalternativen.
- Frågan måste vara konstruerad så att respondenten inte börjar svara för tidigt, t.ex. innan intervjuaren hunnit läsa färdigt.

Svarsalternativen

Svarsalternativen är en viktig del i en fråga och har ett eget kapitel i denna bok. Nedan redovisas endast några grundregler. För mer information, läs om svarsalternativ i kapitel 3.

Svarsalternativen ska stämma överens med frågan

I idealfallet är fråga och svarsalternativ en helhet, det vill säga svarsalternativen och frågan innehåller samma dimension. I vissa fall uppfylls inte det, kanske på grund av att svarsalternativen innehåller en annan enhet eller dimension än det som efterfrågas. Exempelvis om frågan "Går du på restaurang?" som är en ja-nej-fråga istället har

svarsalternativ som är veckodagar – ”måndag”, ”tisdag”, ”onsdag” och så vidare. I det fallet innehåller svarsalternativen en dimension som inte efterfrågas i frågan. I detta fall förstår kanske respondenten intentionen med frågan ändå men den blir onödigt krånglig att besvara.

Det finns även exempel när svarsalternativen tillför en ytterligare dimension som inte finns i frågeformuleringen. Exempel: ”Går du på restaurang?” med svarsalternativen ”Ja, ofta”; ”Ja, sällan”; ”Nej, Aldrig”. Här tillför svarsalternativen en frekvensdimension (hur ofta) som frågan inte innehåller. Vid enklare frågor och välkända dimensioner utgör detta inget större problem. Huvudregeln är dock att svarsalternativ och fråga ska stämma överens vad gäller svarsdimension.

Ett tredje exempel, som är allvarligare och mer förrädiskt, är när svarsalternativen blandar dimensioner utan att frågekonstruktören kanske har tänkt sig det.

EXEMPEL:

Alltid

Ganska ofta

Ganska sällan

Mycket sällan

Ovanstående är tänkt att vara en subjektiv frekvensskala. Den ska alltså återge hur ofta respondenten gör någonting. ”Alltid” säger i sig inget om frekvenser utan bygger på proportioner. Alltid innebär 100 procent av gångerna, men säger inget om antalet gånger i sig. Det kan, rent språklogiskt, bygga på endast ett tillfälle. Här bör man alltså fundera på vad svarsskalan ska mäta och hur respondenterna kan tänkas använda den.

Svarsalternativen ska vara heltäckande och varandra uteslutande

Målet är att svarsalternativen ska vara heltäckande och varandra uteslutande. Så är dock inte alltid fallet. Om svarsalternativen är överlappande kan det vara svårt för respondenterna att avgöra vilket alternativ de ska välja, eller om de till och med ska markera flera alternativ. Det kan också ge missvisande resultat. Den enklaste typen av detta fel är numeriska intervall som överlappar varandra.

Ett annat problem är när giltiga svarsalternativ saknas. Huvudmålet är förstås att alla respondenter ska hitta ett svarsalternativ som överensstämmer med hur de känner eller tycker. Annars kan det resultera i både mätfel (att de svarar något som inte riktigt stämmer) och partiellt bortfall (att de hoppar över frågan). Detta innebär att alla giltiga svar helst ska täckas av svarsalternativen. Vid faktafrågor kan det finnas ett givet antal utfall. Vid attityd- eller upplevelsefrågor krävs det en bedömning av hur många svarsalternativ som bör ingå. För vissa typer av frågor kan det finnas väldigt många giltiga svar, fler än vad som är rimligt att ha med som svarsalternativ. Det finns ju ett värde i att begränsa antalet svarsalternativ. I dessa fall kan man överväga att lägga till ett öppet alternativ – "Annat" – i slutet där respondenterna själva får fylla i sina svar.

Svarsalternativen ska stå i en logisk ordning

I svarsskalor har svarsalternativen en given inbördes ordning (till exempel från "högt" till "lågt"). Däremot är det inte alltid givet vilken riktning svarsskalan ska ha (om man till exempel ska börja med "högt" eller "lågt"). För flera av de vanligare svarsskalorna har en viss riktning blivit vedertagen praxis. Det är till exempel mycket vanligare att inleda med "Instämmer helt" eller "Stämmer mycket bra" än de andra ändarna av respektive svarsskalor. Det kan finnas ett värde i att följa den konventionen men man bör vara medveten om risken för tendensen att respondenten till exempel "håller med" (läs mer i kapitel 3).

Även för svarsalternativ utan inbördes ordning (svars kategorier/nominalskalor) finns det ibland praxis. Men för många frågor finns det ingen given ordningsföljd på förhand (till exempel för olika idrotter eller politiska partier). Svarsalternativen bör emellertid ha någon form av logisk ordning. I brist på annat kan alltid bokstavsordning användas.

Sammanfattande punkter:

- Svarsalternativen ska vara i samma enhet/svarsdimension som frågan och inte tillföra någon ny enhet eller dimension, utöver frågeformuleringen.
- Svarsalternativen ska vara varandra uteslutande och heltäckande.

- Svartalternativen ska stå i logisk ordning. I skattningsskalor är ordningen given. I andra typer av svartalternativ kan, i brist på annat, bokstavsordning användas.

Det ska vara tydligt hur frågan ska besvaras

Instruktioner är tilläggsinformation till frågeformuleringen. Instruktioner kan innehålla definitioner av termer och begrepp i frågeformuleringen, men de kan även innehålla "regler" för hur frågan ska fyllas i, till exempel "Markera endast ett alternativ". Vid många frågor är det underförstått att respondenten endast ska markera ett svar per fråga. I vissa fall kan det dock behöva betonas med en instruktion. Det vanligaste är annars att lägga till en instruktion i de fall som bryter mot konventionen eller tidigare svarsformat i enkäten (till exempel vid de frågor där flera markeringar får göras). I dessa fall ska det finnas en instruktion.

Sammanfattande punkter:

- "Reglerna" för frågan måste vara tydliga för respondenten. Om reglerna ändras så kan det behövas en instruktion (t.ex. att respondenten får svara mer än ett alternativ).

Känsliga frågor

Undersökningar innehåller ibland känsliga frågor. De brukar definieras som frågor som respondenterna upplever som genanta, provocerande, hotfulla eller obehagliga att besvara. Det är givetvis ibland svårbedömt vad som i praktiken är en känslig fråga. Frågor som frågekonstruktören antar är känsliga kanske inte alls upplevs så av respondenterna. Det motsatta kan också inträffa – att frågor som inte antas vara känsliga ändå uppfattas så av respondenterna. Det kan förstås också variera stort mellan respondenter vad som uppfattas som känsligt och inte. Om man är osäker kan det vara bra att testa frågorna först eller utreda erfarenheter från andra undersökningar inom samma ämnesområde. Observera att känsliga frågor här inte är synonymt med känsliga personuppgifter enligt Personuppgiftslagen (1998:2014).

Social önskvärdhet – att vilja framstå i bättre dager

Konsekvenserna av känsliga frågor beskrivs utförligt nedan. Vad gäller själva mätningen finns det två tendenser som kan ha betydelse för svaren:

- Självförnekelse ("self deception" på engelska) – respondenten erkänner inte känsliga saker för sig själv och då inte heller för andra.
- Social önskvärdhet – svaren styrs av vilken bild respondenten vill ge av sig själv ("impression management" på engelska).

Av dessa är social önskvärdhet mest utforskat i litteraturen. I de flesta samhällen finns förutom lagar även normer (oskrivna regler) och ideal för vad som är positivt och negativt, önskvärt och icke önskvärt. Dessa kan skilja sig mellan olika kulturer och över tid. Det finns dock många gemensamma normer. En kategori av känsliga frågor är de frågor som ber respondenterna att ange hur de förhåller sig till sådana normer. Respondenten kan bli påverkad av dessa normer och förändra sitt svar för att framstå i bättre dager. Detta kallas för social önskvärdhet. Det kan till exempel ge utslag på frågor om sådant som väcker oro för att det är olagligt eller kontroversiellt i något avseende. Men även det motsatta kan leda till justeringar av svaren, till exempel om det frågas efter sådant som upplevs som önskvärt av respondenten (till exempel att rösta eller skänka pengar). Dessa två faktorer skiljer sig åt i en central aspekt. I det ena fallet ombeds respondenten att rapportera normativa beteenden, i det andra fallet icke-normativa. Det är rimligt att anta att det icke-normativa kan leda till underrapportering och det normativa till överrapportering (Andersson, 2011; Bradburn, Sudman, Blair & Stocking, 1978).

Justeringen kan alltså ske i båda riktningarna. Tendenserna är starkare om det sker inför en annan person, jämfört med om det sker privat eller anonymt. Därför är social önskvärdhet starkare i intervjuer än i självadministrerade datainsamlingsätt. En studie fann att i en självadministrerad enkät var det 6 procent som rapporterade att de hade använt kokain men bara 4 procent rapporterade det till en intervjuare (Biemer & Lyberg, 2003). Frågan "Hur ofta kör du bil efter att ha druckit alkohol?" gav signifikanta skillnader mellan pappersenkät ("Aldrig" = 52 %) och telefonintervju ("Aldrig" = 63 %) (Dillman & Tarnai, 1991).

Fenomenet är främst utforskat vad gäller respondenters egna beteenden och inte lika väl dokumenterat vad gäller respondenters utsatthet som till exempel offer.

Ibland är det frågan i sig som är känslig, inte svaret

Förutom att ämnesområden kan ha normer är en del områden privata av naturen. Ofta är det inte nödvändigtvis svaret i sig som är känsligt inom sådana områden utan att få själva frågan – en reaktion kan vara att ingen annan än individen själv "har med det att göra". Exempel på sådana ämnesområden är sexualliv och frågor om den egna kroppen. Som med allting annat varierar det förstås mellan olika individer i vilken grad de uppfattar ett visst ämnesområde som känsligt eller inte. Det kan även förändras över tid.

Risken för röjande

Ytterligare en dimension av en frågas känslighet är respondenternas upplevelse av risken för röjande av svaren till tredje part. Frågan upplevs alltså som extra känslig om respondenten upplever en risk att svaren kan spridas (medvetet eller av misstag) till en tredje part (till exempel närstående, myndigheter, forskare, journalister eller obehöriga medarbetare).

Risker med känsliga frågor

Riskerna med känsliga frågor är:

- Respondenten hoppar över frågan (resulterar i partiellt bortfall).
- Respondenten avbryter sitt deltagande (objektbortfall).
- Respondenterna svarar oärligt (mätfel).
- Respondenterna far illa (oetiskt samt försämrat undersökningsklimat och förtroende).

Alla dessa risker är väldigt viktiga att ta hänsyn till.

Frågekonstruktörens förhållningssätt

Hur ska man som frågekonstruktör förhålla sig till känsliga frågor? Ett sätt är ju att låta bli att ställa dem. Ibland kan det vara den lämpligaste lösningen. Samtidigt kan vissa känsliga frågor vara väl värda att ställa för att de ökar vår kunskap inom viktiga områden. Hur ska man då göra för att minska de känsliga komponenterna i dem?

Att försäkra respondenten om säkerheten

För att minska riskerna kan den mer privata enkätformen användas istället för intervju. Det är även viktigt att på ett förtroendeingivande sätt informera respondenten om att svaren är konfidentiella och att deras identiteter inte kommer att röjas (ESS, 2006). Metastudier visar att det kan ha positiva effekter på objektsbortfall vid känsliga ämnen (Singer, Von Thurn & Miller, 1995). Om ämnet inte är känsligt kan dock sådan extra tydlig information om säkerheten ha motsatt effekt (Singer, Hippler & Schwarz, 1992). Förmodligen för att det väcker tankar hos respondenterna som annars inte skulle ha väckts – att det kan finnas skäl att oroa sig. Det kan även vara bra att inledningsvis förklara att människor är olika och att det inte finns några rätt eller fel svar (ESS, 2006).

Det är också bra att hänvisa till stödresurser om frågeformuläret skulle riva upp känslor för respondenten.

Att justera frågan eller frågeformuläret

Frågeformuleringarna är ytterligare en sak att tänka på. Det är viktigt att de är anpassade efter målgruppen och ämnets känslighet, samt att de inte är klumpigt eller provocerande ställda. Detta bör man testa innan frågorna tas i bruk, speciellt om frågorna inte har använts tidigare.

Angående socialt önskvärda svar förekommer det flera tekniker i litteraturen. En är att använda öppna frågor (se nedan). Dessa är betydligt krångligare att koda men kan öka rapporteringen av socialt oönskade beteenden. Orsaken till detta anses vara att med fasta svarsalternativ tvingas till exempel en person med hög alkoholkonsumtion att välja det mest extrema svarsalternativet. Detta är något individer tenderar att undvika generellt sett och den tendensen borde vara ännu starkare om det gäller oönskade beteenden. Den problematiken försvinner dock om frågan är öppen (Bradburn, Sudman & Wansink, 2004).

Andra tekniker är följande (Davidsson & Wärneryd, 1990; Bradburn, Sudman & Wansink, 2004):

- Göra frågan avsiktligt ledande genom att anspela på auktoriteter ("Många läkare anser numera ...").
- Avdramatisera beteendet som något alla gör ("Till och med de tålmodigaste föräldrar blir arga på sina barn ibland ...").

- Använda slang eller familjära uttryck istället för de formella som ofta är mer värdeladdade eller tabu.
- Ta beteendet för givet ("Hur många cigaretter röker du?") men ha med "Inga" i svarsalternativen.
- Ha med skäl i frågan för det socialt icke önskvärda beteendet ("Många förare rapporterar obehag av att bära bilbälte ...").

Uppmärksamma att teknikerna ovan bryter mot många av de tidigare reglerna i det här kapitlet (läs mer i avsnittet Ledande frågor). Idén är att det vid känsliga frågor kan minska underrapportering. Dock finns det inte så mycket forskning om hur dessa tekniker fungerar i praktiken. Det finns en risk att respondenterna reagerar på de ledande frågeformuleringarna så teknikerna bör inte användas utan att testas ordentligt först.

För att reducera tonvikten på det socialt oönskade beteendet kan frågan i frågeformuläret blandas med frågor om mindre hotfulla beteenden. Långa frågor har visat sig bättre än korta vad avser känsliga områden.

Det finns också indirekta frågetekniker som används just för känsliga frågor. Ett exempel är "item-count-tekniken" i vilken respondenterna ska uppge från en lista hur många händelser de varit inblandade i. De svarar alltså gemensamt för hela listan, inte för enskilda händelser. Urvalet delas upp så att en del får den ordinarie listan, medan den andra delen får den ordinarie listan plus den känsliga händelsen som man är särskilt intresserad av. Genom att jämföra de två urvalen kan då andelen som varit inblandade i den känsliga händelsen räknas ut. Eftersom respondenterna svarar gemensamt för hela listan blir besvarandet nog mindre känsligt och de svarar förmodligen ärligare (Holbrook & Krosnick, 2010; Tsuchiya, Hirai & Ono, 2007). Andra exempel på frågetekniker som använts vid känsliga frågor är den korsvisa eller triangulära modellen (läs mer i till exempel Jann, Jerke & Krumpal, 2012).

Läs mer om känsliga frågor och datainsamlingsmetod i kapitel 6.

Sammanfattande punkter:

- Känsliga frågor är sådana som respondenterna tycker är genanta, hotfulla, provocerande eller obehagliga att besvara.
- Vad som är känsligt kan variera från person till person, men vissa ämnen är, generellt sett, känsligare än andra.

- Social önskvärdhet är när respondenten svarar oärligt för att följa en norm och framstå i bra dager. Det kan vara både underrapportering och överrapportering.
- Ibland kan frågan i sig vara känslig att få, oavsett svaret.
- Känsligheten påverkas också av den upplevda risken för röjande. Konsekvenserna av känsliga frågor kan bli mätfel, partiellt bortfall, objektsbortfall och att respondenten far illa. Alla är väldigt viktiga att försöka undvika.
- Känsligheten kan mildras genom att använda enkät, försäkra respondenten om undersökningens säkerhetsgrad, ha respektfulla frågeformuleringar, blanda känsliga frågorna med mindre känsliga frågor och att ställa de känsliga frågorna med ledande eller särskilda frågetekniker.

Sammanfattning

Frågeteknik – sammanfattning

Övergripande mål för frågekonstruktion:

- Frågan ska kunna förstås av alla respondenter.
- Frågan ska förstås på samma sätt av alla respondenter.
- Frågan ska mäta det som avses.
- Frågan ska vara så kort som möjligt (men inte förkortas på förståelsens bekostnad).
- Respondenten ska inte kunna reagera negativt på frågan.
- Frågan ska ha ett syfte (annars bör den inte ställas och ta upp respondenternas tid).
- Det ska finnas ett svarsalternativ för alla respondenter.
- Svarsalternativens format ska vara förenligt med de analyser man planerar att göra.

Två frågor i samma fråga:

- Ställ inte mer än en fråga åt gången.
- Problemet med flera frågor i samma fråga kan enklast lösas genom att dela upp det på flera frågor.

- Om det är kombinationen av flera saker som åsyftas i frågan måste det tydligt framgå i frågeformuleringen, till exempel genom att skriva "både x och y".

Inbyggda förutsättningar:

- Om frågan tar vissa förutsättningar för givna ställer det till problem för de respondenter som inte uppfyller dem (kan leda till både bortfall och mätfel).
- De inbyggda förutsättningarna kan vara dolda i den betydelsen att de inte uttrycks direkt i frågeformuleringen.
- De inbyggda förutsättningarna kan också vara explicita genom att frågan först konstaterar något (som kanske inte stämmer för respondenten) och sedan frågar något utifrån det.
- Problemet löses genom att antingen erbjuda ett svarsalternativ för dem som förutsättningarna inte gäller för eller inleda med en filterfråga.

Definiera frågan i tid och rum:

- De tidsmässiga utgångspunkterna i frågan ska framgå för respondenten.
- Om det är ett "normalläge" som åsyftas i frågan bör det tydliggöras genom tillägg som till exempel "vanligtvis", "oftast" eller "generellt sett".
- Om frågan syftar på en specifik period (en så kallad referensperiod) ska den vara tydlig och rimlig, givet det som ska mätas.
- Framträdande händelser kan ha en längre referensperiod (sträcka sig flera år bakåt i tiden).
- Triviala händelser bör ha en kort referensperiod (vecka/veckor).

Ledande frågor:

- En ledande fråga är när frågeformuleringen i sig (snarare än innehållet i frågan) leder respondenterna till ett visst snarare än ett annat svar.
- En fråga kan vara ledande genom att använda värdeladdade och icke-neutrala ord.
- Att anspela på auktoriteter (till exempel forskare eller läkare) är också ledande.

- Att plantera en norm för vad andra tycker eller gör är också ledande (till exempel "de allra flesta tycker si, vad tycker du?").
- Att bara redovisa ena sidan av en sakfråga kan vara ledande.
- En s.k. "minimal balansering", att ha "eller inte" i slutet av frågan, kan vara ett fungerande sätt att göra en ledande fråga neutral.
- Att ställa frågor om komplicerade sakfrågor på ett sådant förenklat sätt att få kan tycka annorlunda är också ledande, även om frågeformuleringen i sig kan vara neutral.
- Svartalternativen kan vara ledande på samma sätt som frågeformuleringen.
- Svartalternativen kan också vara ledande genom att den ena sidan representeras av fler svartalternativ än den andra sidan.

Filter- och följdfrågor:

- Ett frågeformulär innehåller ibland frågor som inte är relevanta för alla respondenter utan bara för dem som svarat på ett visst sätt.
- En lösning är filterfrågor – ett slags förgreningsfråga – som separerar respondentgrupperna och leder dem vidare till respektive aktuell plats i frågeformuläret.
- Ett annat sätt är följdfrågor som har inbyggda villkor – att de bara ska besvaras om respondenten svarade "Ja" på frågan innan.
- Följdfrågor fungerar bra om det endast är en fråga, annars är filterfrågor bäst. I pappersenkäter, där respondenten sköter hoppen själv, finns det en del risker med filterfrågor. De bör därför inte vara för många i en enkät.

Frågeformuleringens innehåll:

- Undvik negationer i frågor.
- En fråga består av olika typer av innehåll, t.ex. olika typer av avgränsningar såsom referensperiod och specificeringar.

- Det är ofta inte givet i vilken ordning de olika innehållen bäst ska komma men det är värt att reflektera över.
- Exempel i frågor kan hjälpa respondenten att förstå vad som menas med t.ex. "motion" eller påminna dem om saker som de annars inte skulle ha kommit ihåg.
- Om inte ovanstående funktioner är nödvändiga så behöver frågan antagligen inte heller några exempel.
- Undvik att bara ha "ovanliga exempel" eftersom det kan förvirra.
- Längden på frågeformuleringen kan påverka svaren även om betydelsen är densamma. Det beror bland annat på att en längre frågeformulering innehåller fler ord som kan ge andra associationer än dem i en kort frågeformulering men också på att respondenten (i en intervju) får mer betänketid.
- Förbjuda och tillåta (och möjligen för och emot) är inte nödvändigtvis varandras motsatser i frågekonstruktionens värld. Det kan bero på att de som är likgiltiga svarar nekande på båda frågorna.
- Problemet kan lösas genom att inkludera "Vet inte" eller "Ingen åsikt" som alternativ.
- Att ha avdramatiserande introduktioner (vissa tycker si, andra tycker så) till känsliga frågor verkar generellt inte förbättra mätningen.
- När generella beskrivningar ska utvärderas, bryt ner de visionära ledorden till ord med mer precis betydelse. Annars blir det svårt att veta vad svaren baseras på.
- Överväg användningen av vanskliga ord som "tillräckligt" och "regelbundet".
- I enkäter finns det inget egenvärde i att variera sig (använda synonymer). Tvärtom bör man definiera centrala begrepp och sedan hålla sig till dem.
- Det kan bli fel för respondenten om två begrepp är tänkta att vara synonymer men inte är det i praktiken.
- Det måste tydligt framgå för vem respondenten svarar – sig själv, hushållet eller företaget?

- Om frågor ska fånga förändringar, tänk på att både riktningen på förändringen och de absoluta nivåerna (före och efter) kan vara intressanta att känna till.
- Människor förändras och kan därför lämna olika svar vid olika tidpunkter på ett sätt som kan tyckas motsägelsefullt i jämförelse med objektiva mått.
- Frågor som "Visste du att ..." kan vara den mest praktiska lösningen men bör undvikas för områden som är känsliga eller där det finns normer.

Att fråga om det okända:

- Hypotetiska frågor ger inga bra prognoser och brukar avrådas. De kan dock fungera för att få insikt i hur respondenten bedömer saker här och nu.
- Att fråga om medvetna motiv går bra men att fråga om bakomliggande orsaker är svårare.

Undvikande mörkertal:

- Individerna är en aktiv agent i sitt eget liv och kan vara benägen (i den mån det går) att undvika situationer som upplevs som negativa (t.ex. att bli diskriminerad).
- Sådana frågor kan behöva kompletteras med annan information (till exempel om hur ofta de befinner sig i de relevanta situationerna, om de aktivt undviker dem etc.).

Frågehierarkier:

- Vill man ha en spontan, generell bedömning är det bäst att ställa den generella frågan först.
- Om man tvärtom vill att respondenterna ska ha några specifika aspekter i åtanke när de besvarar den generella frågan, är det bättre att inleda med de specifika frågorna.
- Om det bara är några få frågor kan ordningsföljden ha extra stor betydelse, speciellt för den generella frågan.

Frågekonstruktion och intervjuer:

- En intervjufråga måste vara skriven så att den fungerar att läsas och intervjuaren måste känna till hur frågan ska hanteras i praktiken.

- Om inte intervjuaren läser upp svarsalternativen måste frågan vara konstruerad så att den ändå leder respondenten till de underliggande svarsalternativen.
- Frågan måste vara konstruerad så att respondenten inte börjar svara för tidigt, t.ex. innan intervjuaren hunnit läsa färdigt.

Svarsalternativen:

- Svarsalternativen ska vara i samma enhet/svarsdimension som frågan och inte tillföra någon ny enhet eller dimension, utöver frågeformuleringen.
- Svarsalternativen ska vara varandra uteslutande och heltäckande.
- Svarsalternativen ska stå i logisk ordning. I skattningsskalor är ordningen given. I andra typer av svarsalternativ kan, i brist på annat, bokstavsordning användas.
- "Reglerna" för frågan måste vara tydliga för respondenten. Om reglerna ändras så kan det behövas en instruktion (t.ex. att respondenten får svara mer än ett alternativ).

Känsliga frågor:

- Känsliga frågor är sådana som respondenterna tycker är genanta, hotfulla, provocerande eller obehagliga att besvara.
- Vad som är känsligt kan variera från person till person, men vissa ämnen är, generellt sett, känsligare än andra.
- Social önskvärdhet är när respondenten svarar oärligt för att följa en norm och framstå i bra dager. Det kan vara både underrapportering och överrapportering.
- Ibland kan frågan i sig vara känslig att få, oavsett svaret.
- Känsligheten påverkas också av den upplevda risken för röjande. Konsekvenserna av känsliga frågor kan bli mätfel, partiellt bortfall, objektsbortfall och att respondenten far illa. Alla är väldigt viktiga att försöka undvika.
- Känsligheten kan mildras genom att använda enkät, försäkra respondenten om undersökningens säkerhetsgrad, ha respektfulla frågeformuleringar, blanda känsliga frågorna med mindre känsliga frågor och att ställa de känsliga frågorna med ledande eller särskilda frågetekniker.

3 Svarsalternativ

Författare: Andreas Persson (SCB)

Svarsalternativ

Svarsalternativen – en viktig del av frågan

Svarsalternativen fyller en viktig funktion. De fastställer vilka svar som är giltiga och det är dessa som sedan analyseras när all data har samlats in. Det är lätt att underskatta svarsalternativens roll, att de ses som naturligt givna av frågeformuleringen och därför mindre viktiga. Det är en felaktig syn. I självadministrerade frågeformulär är svarsalternativen en del av det informationspaket som respondenten ska ta ställning till och kan till och med påverka hur själva frågan förstås.

I kapitlet "Frågeteknik" finns en del grundregler för svarsalternativ. Svarsalternativen ska stämma överens med frågan, vara heltäckande och varandra uteslutande och stå i en logisk ordning. Det ska också vara tydligt hur svarsalternativen ska användas (om till exempel ett eller flera alternativ får markeras). Dessa grundregler gäller *alltid*.

Olika skaltyper

Svarsalternativen kan vara av olika skaltyper. Med det menas att mätnivån är olika vad gäller om svarsalternativen kan särskiljas, rangordnas och om måttenheten är konstant mellan svarsalternativen. De olika skaltyperna är (Dahmström, 2005; Fråga rätt, 2001):

Kvotskala: Svarsalternativen kan särskiljas, rangordnas och måttenheten är konstant. Det finns en absolut nollpunkt (t.ex. vikt eller längd).

Intervallskala: Svarsalternativen kan särskiljas, rangordnas och måttenheten är konstant (t.ex. celsiuskalan).

Ordinalskala: Svarsalternativen kan särskiljas och rangordnas (t.ex. utbildningsnivå).

Nominalskala: Svarsalternativen kan särskiljas (t.ex. yrke).

Alla skaltyper förekommer i statistiska undersökningar men de två sistnämnda är nog vanligast. Vilken skaltyp en fråga kan vara av styrs framförallt av det som ska mätas. Skaltypen styr i sin tur vilka databearbetningar och analyser som kan göras.

Olika svarsalternativ – olika beslut

Svarsalternativens utformning ska styras av vilka analyser man planerar att göra med datamaterialet. Till vilken grad vill man till exempel kunna särskilja mellan olika respondenter på en variabel? Det påverkar hur många svarsalternativ man bör ha. Dessa grundläggande förutsättningar bör finnas på plats innan man designar svarsalternativens utformning och de diskuteras inte vidare i den här boken.

Det finns många olika valmöjligheter som frågekonstruktören ska ta ställning till när svarsalternativen utformas. Detta kapitel fokuserar på vad man bör tänka på i sådana situationer. Ska frågorna vara öppna eller slutna? Ska man använda frekvensskala eller ordinalskala med vaga kvantifieringar? Hur många skalsteg ska det vara? Ska man ha ett mittenalternativ i skalan? Ska man ha verbala etiketter på skalpunkterna? Ska man erbjuda ett alternativ för icke-svar (till exempel "Vet inte" eller "Inte aktuellt")? Vilka intervall ska man ha i en frekvensskala och vad ska vara tak och golv? Vilken ordning ska man ha i en nominalskala med självständiga svarsalternativ? Dessa frågeställningar måste frågekonstruktören ta ställning till.

Svarsalternativ kan alltså variera på många olika sätt (Tourangeau, Rips & Rasinski, 2000). Svarsalternativen kan

- vara öppna eller slutna (kallas även för "stängda" eller "fasta" svarsalternativ)
- utgöras av siffror, verbala beskrivningar eller både och
- variera i antal
- ha en given eller mer slumpmässig ordningsföljd.

Detta kapitel inleder med att redogöra för öppna svarsalternativ för att sedan gå igenom slutna svarsalternativ. Först kommer de med given inbördes ordning och sedan de utan given inbördes ordning.

Sammanfattande punkter:

- Svarsalternativen fyller en viktig funktion – de styr svaren.
- Svarsalternativen kan, om respondenten får ta del av dem, hjälpa till med förståelsen av frågan.
- Svarsalternativen kan vara av olika skaltyper beroende på om de kan särskiljas, rangordnas och om måttenheten är konstant.

Öppna svarsalternativ kan vara väldigt värdefulla tidigt i ett frågeformulärs levnadscykel. Informationen som fås kan bidra till förbättringar, till exempel en mer komplett frågeformulering eller mer heltäckande fasta svarsalternativ. De kan också vara värdefulla i analyskedet, speciellt om ämnesområdet är relativt nytt eller okänt (ett alternativ till detta är att göra kvalitativa förstudier, läs mer i kapitel 8).

En kompromiss mellan öppna och slutna alternativ är att göra frågan delvis öppen med slutna svarsalternativ plus ett öppet avslutande alternativ (Fowler, 1995; Bradburn, Sudman & Wansink, 2004).

Svagheter

Öppna svarsalternativ anses vara mer krävande, både för respondenten som måste skriva svaret men också för dem som ska koda eller analysera svaren (Bradburn, Sudman & Wansink, 2004; ESS, 2006).

Respondenten måste formulera ett eget svar och uttrycka det i ord vilket är mer krävande än att välja något ur en lista med redan givna svarsalternativ. I självadministrerade frågeformulär finns det därför en risk för partiellt bortfall eller knapphändigt ifyllande på öppna frågor som gör anspråk på ett berättande svar.

Ny forskning antyder dock att öppna svarsalternativ ibland faktiskt kan öka motivationen hos respondenterna eftersom de ges möjlighet att berätta sin egen historia (Wenemark, 2010). Denna forskning är baserad på enkätundersökningar inom hälsoområdet – ett område som ligger nära individen och som många är intresserade av, vilket kan passa öppna svarsalternativ extra bra.

Öppna svarsalternativ kräver också mycket efterarbete. Svaren kan inte skannas till siffror i en datafil och det kan bli problematiskt att läsa vissa svar på grund av svårläst handstil. Det kan också uppstå svåra bedömningar eftersom svaren sannolikt kommer att variera mycket i innehåll men måste klassificeras på ett standardiserat vis för att kunna analyseras kvantitativt (Wärneryd, 1990; Bradburn, Sudman & Wansink 2004; ESS, 2006; Fowler, 1995). Ska man använda resultatet kvantitativt kan det alltså krävas att man utvecklar en gedigen kodningsmall. Dessa svårigheter måste frågekonstruktören ha i åtanke.

Lämpliga situationer för öppna svarsalternativ

Det finns situationer när öppna svarsalternativ är extra lämpliga. Ett exempel är när de möjliga svarsalternativen är fler än vad som är rimligt att innefatta i ett frågeformulär:

EXEMPEL:

Vilket tv-program tittar du mest på?

En möjlighet i sådana situationer är att ha frågan delvis öppen, där de vanligaste svarsalternativen finns som slutna men med ett öppet svarsfält för övriga svar.

Ett annat exempel på när öppna svarsalternativ kan vara lämpliga är när frågor är för komplexa för att kunna besvaras med ett, oftast kortfattat, svarsalternativ.

EXEMPEL:

Hur var feodalsystemet i det medeltida Europa uppbyggt?

Den här typen av frågor är väldigt ovanliga i frågeformulär utan hör nog mer hemma i klassrummet. Öppna frågor kan vara bra för kunskapsfrågor, eftersom det inte finns några svarsalternativ som kan vägleda respondenten till rätt svar. Utan fasta svarsalternativ kan inte heller respondenten chansa rätt (vilket är 50 % chans vid ja-nej-frågor).

Öppna svarsalternativ kan också användas när man befärar att svarsalternativen i sig kan påverka svaren i någon riktning. Slutna svarsalternativ är inte bara boxar som respondenten kryssar i utan de förmedlar också information, till exempel vad som är normalt eller extremt. Genom svarsalternativens intervall och position kan personer bli varse om att de till exempel har en väldigt hög alkoholkonsumtion, tittar mycket på tv eller motionerar väldigt lite. Svarsalternativen ger alltså feedback till respondenten och kan påverka deras svar (läs mer om social önskvärdhet i kapitel 2). När man befärar att informationen i svarsalternativen kan leda till skevhet kan öppna svarsalternativ därför vara ett alternativ.

Öppna svarsalternativ kan också vara en fördel när man vill mäta respondenternas opåverkade uppfattning eller erfarenhet av något. Det kan till exempel vara vad de spontant kommer att tänka på (Wärneryd, 1990).

Öppna svarsalternativ är också ofta den bästa lösningen om frågeområdet är okänt eller om frågekonstruktören saknar tillräcklig kunskap för att fastställa fasta svarsalternativ.

Sammanfattande punkter:

- I öppna svarsalternativ för text får respondenten formulera svaret själv, med sina egna ord.
- Styrkor: svaret är inte avgränsat till frågekonstruktörens syn på saken, det kan bidra med ett kvalitativt inslag, det kan visa vilka svarsalternativ som är vanligast samt visa på vad som bör ingå som fasta alternativ.
- Svagheter: kräver mer ansträngning av respondenten att besvara samt mer kodning och efterarbete.
- Öppna svarsalternativ för text är lämpligt när antalet möjliga svar är många, när svarsalternativen i sig kan påverka svaren eller när man vill få en djupare förståelse.

Öppna svarsalternativ där respondenten ska fylla i siffror

Öppna svarsalternativ kan också användas för siffror (istället för text, som i förra avsnittet). I självadministrerade enkäter används ofta ett speciellt svarsfält:

EXEMPEL 1:

Vilket år är du född?

År

EXEMPEL 2:

Hur stor är din bostadsyta?

Kvadratmeter

Fälten ovan har ett begränsat utrymme. Det är rekommenderat att fältet innehåller så många positioner som svaret förväntas innehålla, varken

fler eller färre. Det är också rekommenderat att man skriver ut svarsenheten i anslutning till svarsfältet så att inte respondenterna svarar i fel enhet, till exempel i kronor när det ska vara i tusentals kronor.

Styrkor

Öppna svarsalternativ där respondenten ska fylla i siffror har en del fördelar jämfört med slutna svarsalternativ. De kan ge mer exakt och detaljerad information än om intervall används i en skala med slutna svarsalternativ. I en skala med intervall kan man aldrig särskilja mellan personer inom samma intervall, vilket kan vara en brist speciellt om intervallet är stort. Med ett öppet svarsalternativ slipper man det bekymret. I det fall ämnet är känsligt, eller vid mindre undersökningar, kan det kännas olustigt för respondenterna att ge för detaljerad information. Då kan svarsintervall vara att föredra.

Med öppna svarsalternativ slipper man också systematiska avvikelser som har med informationen i svarsalternativen att göra. Det kan till exempel vara så att vissa respondenter justerar sina svar för att undvika att hamna i det lägsta intervallet på en fråga om inkomst.

En annan fördel med öppet svarsalternativ för siffror är att man slipper så kallade golv- eller takeffekter som kan uppstå för att man lagt intervallen i svarsalternativen fel, till exempel för lågt så att väldigt många hamnar i det högsta intervallet (en takeffekt, en golveffekt är motsatsen). Öppna svarsfält för siffror kan dessutom, till skillnad från öppna svarsfält för text, skannas in till siffror i en datafil.

Svagheter

Precis som för öppna svarsfält för text finns dock risker för att svårsläslig handstil eller slarvigt ifyllande ställer till problem när svaren ska skannas eller tydas. Det är också viktigt att tydliggöra så långt det är möjligt vilka enheter som respondenten ska svara i för att undvika ifyllningsfel. Öppna svarsalternativ ökar också i allmänhet respondentbördan jämfört med slutna svarsalternativ.

Avrundning av siffror i svar

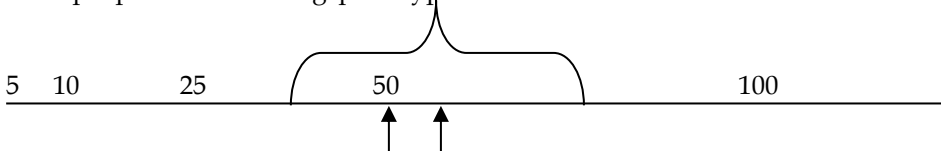
Tourangeau, Rips och Rasinski (2000) diskuterar ett annat potentiellt problem med öppna svarsalternativ där respondenten ska fylla i siffror, nämligen avrundning. I undersökningar om så vitt skilda ämnen som respondenters ålder till antalet partners under ens livstid finns tydliga

indikationer på att respondenterna avrundar och att svaren ofta koncentreras till jämna femtal (5, 10, 15 etc.). I en annan studie (Tourangeau, Rips & Rasinski, 2000; Huttenlocher, Hedges & Bradburn, 1990) där respondenterna skulle ange antal dagar sedan förra intervjutillfället verkade avrundning ske baserat på talen 7 eller 30 (vilket benämns som "kalenderprototypen").

Avrundningen kan bero på att respondenterna av olika skäl inte kan ange ett exakt svar. Kanske minns de inte det efterfrågade på den nivån. Forskning tyder dock på att individer avrundar även när uppgiften inte bygger på minnet utan på bedömningar i realtid, till exempel att avgöra hur långa olika linjer är (Schaeffer & Bradburn, 1989; Tourangeau, Rips & Rasinski, 2000). Kanske är inte siffersystemet så neutralt som vi oftast tänker oss? Kanske är vissa tal mer framträdande än andra?

Avrundning verkar alltså förekomma. Den viktiga frågan är då vilka konsekvenser det kan ha. Kan det leda till systematiska fel? Tourangeau, Rips och Rasinski (2000) beskriver att det finns skäl att tro att det i vissa fall kan göra det. En anledning är att i de avrundningsprototyper som används ökar nämligen avståndet mellan avrundningspunkterna successivt och blir större och större.

Exempel på en avrundningsprototyp:



Eftersom avståndet mellan avrundningspunkterna ökar ligger inte avrundningspunkten mitt i intervallet utan i den lägre halvan. Se exemplet ovan med avrundningspunkterna 5, 10, 25, 50 och 100. En person som avrundar till 50 ska ligga mellan 38 och 75. Men 50 är inte mittpunkten i det intervallet, utan den korrekta mittpunkten är 56,5 (pilen till höger). Avrundningspunkten (pilen till vänster) utgör alltså en underskattning. Detta beror på att respondenterna inte avrundar korrekt utan följer vissa tumregler. Om flera respondenter använder samma tumregler kan det resultera i systematiska underskattningar (Tourangeau, Rips & Rasinski, 2000).

En studie (Holbrook, Anand, Johnson, Cho, Shavitt, Chavez & Weiner, 2014) hittade höga nivåer av avrundning för procentfrågor, frågor om

individuella egenskaper och attitydfrågor med termometerskala. Lägre nivåer av avrundning hittades för beteendefrågor och frågor om respondentens ålder vid olika händelser i det förflutna. Författarna fann inte så mycket stöd för att avrundning berodde på bristande motivation. Tvärtom kan det vara ett resultat av att till exempel respondenterna inser att de inte kan ge ett exakt svar.

Vad ska man då göra åt detta? Det går förstås att understryka vikten av exakta siffror på frågor som man misstänker ligger i riskzonen. Om respondenterna inte har exakta svar att ge är det dock en lönlös strategi. Strategier för att öka respondenternas förmåga att minnas borde också minska risken för avrundning. Ett annat alternativ är att dela upp frågan eller förkorta referensperioden.

Sammanfattande punkter:

- I öppna svarsalternativ för siffror får respondenten skriva siffrorna i svaret själv.
- Styrkor: mer exakta svar än i intervall, undviker tak- eller golveffekter och inga svarsalternativ som i sig kan påverka svaren.
- Svagheter: kan leda till ifyllningsfel och svar som är svåra att tolka.
- En annan svaghet är avrundning: att respondenter inte svarar exakt utan avrundar till vissa punkter, till exempel jämna femtal.
- Det finns en risk att avrundningar kan leda till underskattningar eftersom avståndet mellan avrundningspunkterna tenderar att öka för varje punkt (till exempel 10, 25, 50, 100).

Öppna frågor kontra slutna svarsalternativ

Förutom styrkorna och svagheter redovisade i avsnitten ovan kan debatten om öppna och slutna svarsalternativ föras både på ett vetenskapsteoretiskt och på ett praktiskt och produktionsmässigt plan. Vissa förespråkare för öppna svarsalternativ har hävdats att detta och en konverserande intervjustil är det bästa sättet att nå samsyn och förståelse mellan frågeställare och respondent (Schober & Conrad, 1997; Conrad & Schober, 2000). I stora surveyundersökningar med specifika syften kan det dock vara en omöjlighet i praktiken och ställa orimliga krav på intervjuarna eller på kodning i efterhand. Därför är ofta slutna

svarsalternativ normen och öppna frågor förekommer mest vid specifika frågor.

Läs mer om skillnader mellan öppna och slutna frågor i slutet av detta kapitel.

Slutna svarsalternativ

Slutna svarsalternativ (kallas också fasta svarsalternativ) innebär att respondenten endast kan välja något eller några av de svarsalternativ som frågekonstruktören erbjuder. Slutna svarsalternativ är alltså bestämda på förhand.

EXEMPEL 1:

Vilken typ av bostad bor du i?

Egen villa/radhus
Bostadsrätt/andelsrätt
Hyresrätt
Inneboende, studentlägenhet/rum
Annat

EXEMPEL 2:

Hur säker känner du dig i ditt bostadsområde?

Mycket säker
Ganska säker
Ganska osäker
Mycket osäker

Det finns olika former av slutna svarsalternativ. En vanlig form är att svarsalternativen är helt slutna och att respondenten endast ska välja ett alternativ. Frågor kan dock vara delvis öppna genom att svarsalternativet "annat" förekommer tillsammans med ett öppet svarsfält. Detta kan vara en möjlighet om det av praktiska skäl är omöjligt att ange alla svarsalternativ i en lista eller om frågekonstruktören saknar kunskap om alla svarsalternativ (läs mer i avsnittet Öppna svarsalternativ).

Det är viktigt att svarsalternativen är heltäckande. I faktafrågor kan antal möjliga svar vara givet. I andra typer av frågor måste det bedömmas hur många som är lämpliga att ha. Att ha väldigt många svarsalternativ kan dock också leda till problem. Respondenten kanske inte kan skilja mellan

alla på ett meningsfullt sätt, orkar inte läsa igenom hela listan eller klarar inte av att behålla alla svarsalternativ i minnet (Bradburn, Sudman & Wansink, 2004). Antalet svarsalternativ utgör alltså en avvägning.

Slutna svarsalternativ är enklare för respondenterna att besvara eftersom de då inte behöver formulera något eget svar. Svarsalternativen kan även bidra till respondentens förståelse av frågan och de kräver därmed oftast mindre ansträngning än öppna svarsalternativ. Slutna svarsalternativ kräver även mindre efterarbete än öppna svarsalternativ där svaren måste klassificeras för att kunna analyseras statistiskt (läs mer i avsnittet Öppna svarsalternativ). Frågor med slutna svarsalternativ kräver emellertid mer förarbete för att fastställa vilka svarsalternativ som bör ingå (Bradburn, Sudman & Wansink 2004; ESS, 2006).

Sammanfattande punkter:

- I slutna svarsalternativ väljer respondenten (oftast) ett av några på förhand givna svarsalternativ.
- Styrkor: förhållandevis enkla att sammanställa och analysera samt kräver mindre ansträngning av respondenten.
- Svagheter: kräver mycket förarbete och många beslut (till exempel om vilka och hur många svarsalternativ som ska ingå) samt begränsar respondenten till frågekonstruktörens referensram.

Slutna svarsalternativ med given inbördes ordning

Detta avsnitt fokuserar på slutna svarsalternativ som har en given ordningsföljd, en rangordning. Svarsalternativen kan då inte byta plats med varandra utan att innebörden förändras eller förloras (se exempel 2 i föregående avsnitt). Dessa svarsalternativ kallas ofta för skalor och kan, som redovisats i inledningen av kapitlet, antingen vara av kvot-, intervall- eller ordinalskaletyp:

Kvotskala: Den kan särskiljas, rangordnas, mätas med konstant måttenhet och med absolut nollpunkt (t.ex. vikt eller längd).

Intervallskala: Den kan särskiljas, rangordnas och mätas med konstant måttenhet (t.ex. celsiuskalan).

Ordinalskala: Den kan särskiljas och rangordnas (t.ex. utbildningsnivå).

Slutna svarsalternativ med given inbördes ordning kan vara den svåraste typen att konstruera, eftersom det är många beslut som måste tas och valen kan få konsekvenser. I nedanstående avsnitt diskuteras de val som frågekonstruktören bör ta ställning till.

Vaga kvantifieringar

I undersökningar händer det ofta att svarsalternativen är på ordinal nivå och utgörs av beskrivningar av en mängd. Detta görs genom att svarsdimensionen (till exempel hur ofta en respondent har utfört en aktivitet) tilldelas olika preciseringar såsom "ganska", "viss", "mycket" (se exempel 2 i föregående avsnitt). Ibland används inte preciseringar utan enskilda ord som beskriver en kvantitet, till exempel "ibland" eller "sällan". Dessa beskrivningar kallas för vaga kvantifieringar eftersom det är otydligt vilken faktisk kvantitet som inryms i dem (Bradburn & Miles, 1979). Det får respondenten själv bedöma. Exempel på vaga kvantifieringar är "Mycket bra", "Ganska sällan", "I stort sett alltid", "I viss utsträckning" och "Instämmer lite grann".

EXEMPEL 1:

Är det vanligt att grannarna i ditt bostadsområde utbyter några ord med varandra?

Ja, mycket vanligt

Ja, ganska vanligt

Nej, ganska ovanligt

Nej, mycket ovanligt

I mitten på 1900-talet gjordes en del forskning om vaga kvantifieringar och hur de förhöll sig till faktiska frekvenser (för en genomgång, se Bradburn & Miles, 1979). Forskning visade att personer tillskrev ord som "ibland" väldigt olika kvantiteter. Annan forskning undersökte effekten av ord som ökar intensiteten och kom fram till att "mycket ofta" var ungefär 1,317 gånger så ofta som "ofta" (Cliff, 1959).

EXEMPEL 2

Kan du anpassa dina arbetsuppgifter ifall du skulle bli sjuk?

Sällan eller aldrig

Ibland

Ofta

Problemet med den vaga kvantifieringen ligger just i kopplingen mellan det objektiva och faktiska. Forskning har till exempel visat att ords kvantitativa innebörd varierar bland annat beroende på kontext. "Ibland" innebär olika faktiska frekvenser beroende på om det är förekomsten av jordbävningar eller huvudvärk som ska skattas. Skalpunkterna kan alltså betyda olika saker beroende på vad som utgör referenspunkten – vad det är som efterfrågas (Wärneryd, 1990). Annan forskning visar att betydelsen inte bara varierar beroende på kontext utan även mellan personer, till exempel med olika inställning, upplevelser eller vanor (Bradburn & Miles, 1979). "Väldigt ofta" kan tolkas som "mer än vad som är typiskt för sådana som jag" (Schaeffer, 1991). Vad som är "typiskt" kan dock variera mellan olika grupper av individer. Vaga kvantifieringar kan leda till skillnader i svar mellan grupper trots att det inte finns några faktiska skillnader (Wärneryd, 1990).

EXEMPEL:**Har du besökt följande institutioner ofta, ibland, sällan eller aldrig?**

Vårdcentral
Arbetsförmedling
Försäkringskassan
Grundskola
Mataffär
Bank
Kollektivtrafik

I ovanstående exempel kan säkert till exempel "ofta" eller "sällan" betyda olika saker för de olika alternativen eftersom förväntningarna skiljer sig åt mellan dem (man besöker till exempel mataffären oftare än banken eller vårdcentralen).

Det kan alltså vara lite problematiskt att jämföra svaren mellan olika frågor och mellan olika grupper av respondenter eftersom de bakomliggande faktiska frekvenserna kan variera.

Varför används då vaga kvantifieringar istället för faktiska frekvenser som svarsalternativ? Det är lätt hänt att se på faktiska frekvenser som en objektiv mätning och på vaga kvantifieringar som en subjektiv mätning och att den objektiva då skulle vara mer korrekt. Men det är förmodligen en förenklad bild. På sätt och vis mäter båda en konstruerad bedömning

hos respondenten och båda har sannolikt sina respektive felkällor (Schaeffer, 1991).

Vilken av dem är bäst, det vill säga förknippad med minst fel? Det beror sannolikt på vad man frågar om. Vissa frågeområden frågar efter företeelser som har vaga gränser (till exempel känslor eller tillstånd) där faktiska frekvenser kan passa illa som svarsalternativ. Andra frågor frågar efter sådant som respondenterna kan förväntas minnas på en detaljerad nivå. Då innebär vaga kvantifieringar en informationsförlust och faktiska frekvenser är bättre (Schaeffer, 1991; Bradburn & Miles, 1979).

På grund av ovanstående rekommenderas att man använder faktiska frekvenser i de fall då respondenten förväntas kunna svara på den nivån. I andra fall kan det vara lämpligare att använda vaga kvantifieringar. Detta beror också på vad som är syftet med frågan. Om målet är att mäta den subjektiva upplevelsen av något så är förstås den subjektiva, vaga kvantifieringen lämpligast.

Om vaga kvantifieringar ska användas är det bra att ha ganska stora intervall, eftersom respondenterna sannolikt inte är så säkra på sina svar (Wärneryd, 1990).

Sammanfattande punkter:

- "Vaga kvantifieringar" är när ord används för att beskriva en frekvens eller styrka. Det kan innebära en precisering (till exempel "ganska" eller "väldigt") eller ett enskilt ord som beskriver en frekvens (till exempel "ibland").
- Nackdelen med vaga kvantifieringar är att de inte har någon given betydelse utan utgår från en referenspunkt - "ganska ofta" jämfört med något, men vad?
- Vad som är referenspunkten kan variera mellan olika personer, beroende på till exempel personlighet och erfarenheter.
- Vad som är referenspunkten kan också variera beroende på sammanhang. "Ganska ofta" kan syfta på helt olika bakomliggande faktiska frekvenser beroende på om det är t.ex. bio- eller teaterbesök som efterfrågas.
- På grund av dessa variationer (se ovan) kan det vara problematiskt att jämföra frågor och även grupper av respondenter med varandra.

- Vaga kvantifieringar kan uppfattas som en subjektiv och osäker mätning men ibland är det den bästa lösningen. Vissa saker kan inte besvaras med faktiska frekvenser.

Antal skalsteg

Valet av antal svarsalternativ utgör ett dilemma för frågekonstruktören. Att ha många svarsalternativ möjliggör förfinade analyser, att kunna skilja mellan olika grupper och olika nivåer av mätobjektet. Men man måste också ta hänsyn till respondentens förmåga att minnas, uppmärksamma och särskilja mellan olika alternativ. Om skalstegen är för många förlorar de sin mening för respondenten, och om det händer påverkar det förstås mätningen negativt (Krosnick & Berent, 1993). Som frågekonstruktör är det viktigt att vara medveten om att även för få skalsteg kan ge en missvisande bild. Det kan innebära att man tvingar respondenten att svara något som inte riktigt stämmer. Att bara ha "Ja" och "Nej" som svarsalternativ till en attitydfråga tvingar dem som har en mindre extrem inställning att ta ställning. Frågekonstruktören måste därför hitta en god balans avseende antalet svarsalternativ (ESS, 2006).

Hur många skalsteg en skala kan ha hänger också ihop med om skalstegen ska ha etiketter (att varje skalsteg har en beskrivning i ord) eller inte. Vid beskrivande etiketter sätter ofta språket en begränsning. Språket är tyvärr inte tillräckligt nyanserat för att kunna särskilja för många olika grader av en svarsdimension på ett entydigt sätt. Detta är även relaterat till beslutet om skalan är unipolär eller bipolär (se nedan). I bipolära skalor som är centrerade kring en mittpunkt går det åt färre etiketter eftersom samma etikett kan användas för två skalsteg i svars-skalan, ett steg på varje sida av mittpunkten ("*Ganska positiv*" och "*Ganska negativ*", se exemplet nedan).

EXEMPEL:

Mycket positiv
Ganska positiv
Varken positiv eller negativ
Ganska negativ
Mycket negativ
Ingen åsikt

Hur många skalsteg är då lämpligt? Det finns mycket forskning inom surveyområdet om det optimala antalet skalsteg. Även inom psykologin har det länge forskats mycket om skalsteg eftersom den typen av svarsalternativ är väldigt vanlig inom det området (för en genomgång, se Preston & Colman, 2000). I en studie skulle 149 respondenter skatta servicen på restauranger och affärer på skalor som gick från två till elva skalsteg. Olika analyser om reliabilitet och validitet visade att skalorna med två, tre och fyra skalsteg presterade förhållandevis dåligt men att det blev kontinuerligt bättre upp till sju skalsteg. Respondenterna själva föredrog främst tiogradiga skalor, därefter de sju- och niogradiga (Preston & Colman, 2000).

Litteraturen inom surveyområdet är inte entydig utan rekommendationerna varierar en del. De flesta studier och böcker hävdar dock att det inte är meningsfullt ha fler än fem–nio svarsalternativ vid postenkäter, eftersom respondenterna inte kan skilja mellan fler än så (Fowler, 1995; ESS, 2006; Wärneryd, 1990; Bradburn, Sudman & Wansink, 2004). Forskning visar att reliabiliteten ökar vid fler skalsteg (Alwin, 1992, Alwin & Krosnick, 1991; Saris & Gallhofer, 2007), men att sju skalsteg kan vara en bra kompromiss (Alwin, 2007). I telefonintervjuer bör svarsskalan vara kortare än så, speciellt om så kallad branchingteknik inte används (läs mer i kapitel 6).

Om svarsskalan är numerisk kan flera svarspunkter innefattas (Bradburn, Sudman & Wansink, 2004). Läs dock om unipolära och bipolära skalor nedan.

Sammanfattande punkter:

- Många svarsalternativ möjliggör mer detaljerade analyser.
- Om svarsalternativen är så många att de tappar sin mening påverkas mätningen negativt.

- För få skalsteg kan också påverka mätningen negativt – att respondenten tvingas svara något som inte riktigt stämmer.
- Om skalan ska ha etiketter vid skalstegen begränsar språket hur många skalsteg skalan kan ha.
- Antalet skalsteg beror också på om skalan är uni- eller bipolär. Vid bipolära skalor kan samma etikett användas för två skalsteg, ett på varje sida av mittpunkten.
- Rekommendationerna varierar men en tumregel är fem–nio svarsalternativ i postenkäter och färre i telefonintervjuer.

Riktningen på svarsskalan

Riktningen på svarsskalan, om den till exempel ska inledas med den höga eller låga änden, är en aspekt av frågekonstruktion som inte tillägnats så mycket forskning som den förtjänar. Rekommendationerna från litteraturen är också lite motsägelsefulla.

Det finns två anledningar till att det kan ha betydelse vilken ända man har först (Yan & Keusch, 2015). Den första anledningen är kognitiv. Respondenterna kan komma att utgå från något av de förstkommande svarsalternativen som ett ankare och sedan försöka anpassa sitt svar efter detta (strategin kallas anchor and adjustment på engelska).

Anpassningen blir dock ofta bristfällig och hamnar för nära ankaret, vilket kan leda till en snedvridning i riktning mot den ända av svarsskalan som kommer först. Den andra anledningen har med motivation att göra. De respondenter som försöker besvara enkäten med minsta möjliga ansträngning kan vara mer benägna att välja ett alternativ som kommer tidigt snarare än sent i en svarsskala. I det mest extrema fallet väljer de det allra första alternativet. En studie testade dessa två förklaringar och fann att riktningen på svarsskalan hade effekt i tre fall av tio. Eftersom effekten inte drabbade alla frågor drar författarna slutsatsen att förklaringen inte borde ha med bristande motivation att göra, för i så fall borde alla frågor ha påverkats (Yan & Keusch, 2015).

Det är svårt att designa bort dessa företeelser. Det handlar snarare om att vara medveten om att de kan påverka – den förstkommande änden kan få mer stöd än vad som är riktigt korrekt eller vad den skulle fått om riktningen varit den omvända. Om man är särskilt intresserad av en ända av svarsskalan och vill kunna lita på de svar som hamnar där, kan det vara bäst att inleda med den andra änden.

Sammanfattande punkter:

- Riktningen på svarsskalan kan ha betydelse. Det kan finnas en tendens (p.g.a. både kognitiva och motivationsmässiga skäl) att välja ett tidigt snarare än sent alternativ.
- Det kan vara svårt att designa bort denna tendens.

Instämmer-skalor

Instämmer-skalor används ofta i kombination med frågor i påståendeform (kallas ibland Likertskala, läs mer i kapitel 12). Respondenten ska ta ställning till olika påståenden och markera i vilken grad han eller hon instämmer i det som påståendet hävdar. Instämmer-skalor är ifrågasatta av en del forskare (Schaeffer & Presser, 2003) och vissa avråder från att använda dem (Converse & Presser, 1986). En av de främsta kritikerna, Jon. A. Krosnick, höll till exempel 2009 ett keynote-föredrag med titeln "The end of agree-disagree scales" på en stor surveykonferens.

EXEMPEL:

	Stämmer mycket bra	Stämmer ganska bra	Stämmer ganska dåligt	Stämmer mycket dåligt	Vet inte
På min arbetsplats respekterar vi varandra	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Anledningen till denna skepsis är bland annat att instämmer-skalor har förknippats med så kallad "acquiescence bias", en tendens att hålla med. Det är lättare att hålla med om någonting än att ta avstånd ifrån något. Påståendet att "Regeringen gör ett dåligt jobb" leder alltså i princip till ett lite mer negativt resultat om regeringens prestation än om det omvända, "Regeringen gör ett bra jobb" använts, trots att påståendena borde mäta samma sak (fast spegelvänt). Tendensen kan bero på att respondenterna vill vara artiga och tillmötesgående eller att de uppfattar frågeställaren som en expert med kunskapsmandat. Det kan också vara ett utslag av att respondenten tar kognitiva genvägar för att besvara frågorna med så liten ansträngning som möjligt. Allt detta kan leda till systematiska fel i data (Revilla, Saris & Krosnick, 2014).

Det kan även finnas andra typer av problem förknippade med instämmer-skalan. Om syftet till exempel är att mäta hur bra eller dåligt respondenten upplevde att något var, är det generellt sett enklare att svarsalternativen utgår ifrån bra och dåligt istället för att respondenten ska instämma i eller ta avstånd från ett påstående.

EXEMPEL:

Vården på sjukhuset var bra

För att besvara ovanstående påstående måste respondenten komma fram till hur bra eller dålig vården på sjukhuset var. I instämmer-skalar är svarsprocessen däremot inte avslutad när respondenten har kommit fram till hur bra eller dålig vården var. Respondenten måste dessutom anpassa svaret till påståendet och instämmer-skalan. Förutom att det kräver extra möda av respondenten kan det också uppstå problem. Vad händer till exempel om respondenten tyckte att vården var bättre än bra? Ska respondenten då instämma eller inte? (Revilla, Saris & Krosnick, 2014)? Rent logiskt är det inte givet. Förmodligen förstår de flesta respondenter ändå intention med formatet, till exempel, att de besvarar instämmer-skalan i exemplet ovan som om den i praktiken vore en bra-dåligt-skala. Alltså, att respondenten på sätt och vis på egen hand kompenserar för formatets begränsningar.

På grund av detta kan det vara värt att, där det är möjligt, överväga frågespecifika svarsalternativ istället för en instämmer-skala som svarsalternativ (Saris, Krosnick & Schaeffer, 2005). Man bör dock ha i åtanke att frågeformuläret då oftast blir längre och innehåller fler svarsskalor, vilket beroende på omfattningen också kan öka bördan och störa mätningen.

Påståendeformen kan möjligen fylla en speciell funktion vid mjukare mätningar om till exempel attityder, känslor och upplevelser. Det gäller speciellt om området eller de enskilda påståendena innehåller intima, känsliga eller kontroversiella ämnen. Vanliga frågesatser är ganska konfrontativa i sin framtoning. Det finns som en oskriven regel mellan frågeställare och respondent att en fråga ska vara giltig att ställa, annars borde den inte ställas. Detta härrör från de regler som styr konversation (läs mer i kapitel 7). Då kan påståenden vara att föredra eftersom de inte är lika konfrontativa eller personliga utan just påståenden, som till sin natur kan vara antingen sanna eller falska.

Avslutningsvis kan det nämnas att Krosnick genomförde en annorlunda studie som inför testpersonerna förklarades vara parapsykologiskt. Testpersonerna skulle använda sina paranormala förmågor för att lista ut vilket påstående forskaren tänkte på och sedan instämma eller ta avstånd från det. Mellan 10–20 procent fler testpersoner instämde än tog avstånd ifrån påståendet som de upplevde att forskaren tänkte på. Det kanske kan ses som en hålla-med-tendens, som i detta fall inte var beroende av frågeformuleringen eftersom det inte fanns någon sådan (Ping, 2005).

Sammanfattande punkter:

- Instämmer-skalar (baserade på instämmer eller stämmer) är vanliga svarsskalor till frågor i påståendeform.
- Skalan är förknippad med hålla-med-tendenser ("acquiescence bias").
- Svarsprocessen är mer komplicerad än frågesatser – respondenten måste inte bara ta ställning till innehållet i sig utan även anpassa sitt svar till påståendet och svarsskalan.

Unipolära och bipolära skalor

Svarsskalor kan vara antingen unipolära eller bipolära. Unipolär betyder att skalan endast har en pol, där skalan sträcker sig från mycket av någonting till lite eller inget alls av samma sak.

EXEMPEL 1:

Inte alls

Lite grann

Ganska mycket

Väldigt mycket

Bipolära skalor sträcker sig däremot från en pol till dess motsats.

EXEMPEL 2:

Mycket positiv
Ganska positiv
Varken positiv eller negativ
Ganska negativ
Mycket negativ

Unipolära skalor används oftast för att mäta frekvenser av något (till exempel Våldigt ofta-Aldrig eller som i exempel 1 ovan). Bipolära skalor används oftast för att mäta attityder (till exempel Mycket positivt-Mycket negativt).

Unipolära skalor har bara en pol och har därför generellt sett färre skalsteg än bipolära, som har två poler (läs mer ovan). Krosnick och Tahk (2008) fann att för unipolära skalor ger fem skalsteg det mest tillförlitliga resultatet medan sju skalsteg presterade bäst för bipolära skalor.

Bipolära skalor, som går från en pol till dess motsats, har en naturlig mittpunkt mellan den ena och den andra polen (punkten "Varken positiv eller negativ" i Exempel 2 ovan). Därför är det framförallt för bipolära skalor som diskussionen under nästa rubrik – mittenalternativ eller inte – är relevant.

Det händer att unipolära skalor innehåller en mittpunkt men det är svårt att tolka vad den egentligen betyder. En unipolär skala återger en kontinuerlig ökning. Det är utifrån det perspektivet svårt att förstå varför till exempel skalsteg tre på en femgradig skala skulle ha någon neutral betydelse (som mittenalternativet utgör i bipolära skalor). I praktiken kan det dock vara lite lurigt att avgöra vad som är unipolära och bipolära skalor och det är lätt att förvirras av beskrivningarna av skalstegen.

Sammanfattande punkter:

- Unipolär betyder att skalan endast har en pol – den utgår från lite av något till mycket av samma sak.
- Bipolära skalor innehåller två poler och går från en pol till dess motsatta pol.

- Unipolära skalor används mest för att mäta frekvenser, bipolära för att mäta åsikter.
- Det är främst vid bipolära skalor som en mittpunkt är aktuell, placerad mittemellan de två polerna.
- Bipolära skalor har traditionellt sett fler skalsteg än unipolära.

Mittenalternativ eller inte

Om mittenalternativ ska ingå i en svarsskala eller inte är ett vanligt förekommande beslut för frågekonstruktören. Mittenalternativ är främst relevant vid bipolära skalor (läs i förra avsnittet och se sista exemplet). Att erbjuda ett mittenalternativ eller inte påverkar resultatet. Beroende på vilken typ av fråga det handlar om kan en betydande del av respondenterna markera mittenalternativet som ofta utgör en neutral punkt. Alla som markerar mittenalternativet gör det dock troligtvis inte av önskvärda skäl. Det kan istället bero på bristande motivation (så kallad "satisficing", Krosnick, 1991), vilket resulterar i mätfel. Dock finns det tecken på att flykten till mittenalternativet drabbar båda polerna proportionerligt (Presser & Schuman, 1980). Det är viktigt eftersom det antyder att mittenalternativet inte leder till skevhet även om det påverkar fördelningen. Om detta behövs det dock mer forskning och det kan också variera mellan olika typer frågor.

Om skalan *inte* erbjuder ett mittenalternativ riskerar man å andra sidan partiellt bortfall hos dem som inte vill ta ställning. Respondenter som är neutrala tvingas också att välja någon av polerna vilket ger mätfel. Hur stort felet blir borde bero frågans ämne och om respondenter har en tydlig åsikt om det (lutar åt någon av polerna) eller inte. Vare sig man har ett mittenalternativ eller inte i sin svarsskala riskerar man alltså fel.

Vad ska man då göra? Forskningen på området ger en del vägledning i hur man kan resonera. Fler väljer till exempel mittalternativet om det erbjuds explicit (10–20 %) jämfört med om det erbjuds först om respondenten tvekar eller uttrycker sig ligga mittemellan. Detta är alltså en möjlig kompromiss men främst aktuell i telefonintervjuer där svarsalternativen inte direkt delges respondenten utan först "vid behov".

En studie fann att antalet som valde mittenalternativet minskade med en femgradig skala jämfört med en tregradig (41,4 % jämfört med 53,7 %). Det skulle kunna tolkas som att en del av dem som valde mittenalternativet i den tregradiga skalan faktiskt lutade åt ett håll, men

att det saknades svarpunkter för dem (Presser & Shuman, 1980). Ett sätt att minska betoningen på mittpunkten kan alltså vara att differentiera skalan.

Det finns även studier som visar att fler väljer mittpunkten om frågan är svår att förstå och läsa, vilket kan tyda på att motivation kan vara en viktig faktor (Velez & Ashworth, 2007). Detta borde då kunna motverkas genom att göra frågorna lättare att förstå.

Ett annat sätt att undvika att respondenter som egentligen är tveksamma eller omotiverade markerar mittpunkten är att erbjuda ett ytterligare svarsalternativ för dem, till exempel "Vet inte" eller "Tveksam" (Wärneryd, 1990).

Hur ska man då benämna mittenalternativet? Det finns en rik flora av variationer. I vissa fall har mittenalternativen etiketter som beskriver dess centrala mittposition. I andra fall har de en annan typ av benämning, till exempel "Osäker" eller "Vet inte riktigt". Forskning visar att en benämning som uttrycker skalpunktens mittposition och inte någon form av osäkerhet är bäst för att undvika svar som inte hör hemma där (Hernández, Espejo & González-Romá, 2006).

Bradburn, Sudman & Wansink (2004) rekommenderar att man ska erbjuda mittalternativ generellt. Forskningen på området antyder att det också ökar reliabiliteten (Saris & Gallhofer, 2007; Krosnick & Tahk, 2008). En generell rekommendation är därför att använda mittenalternativ men vara medveten om att det kan påverka resultatet och att tänka på resonemangen ovan när man utformar skalan.

Sammanfattande punkter:

- Mittenalternativ är främst relevant vid bipolära skalor.
- Om man inte har ett mittenalternativ tvingar man dem som är neutrala att ta ställning (mätfel) och det kan också riskera partiellt bortfall.
- Om man har mittenalternativ finns det risk att flera respondenter kommer att markera det p.g.a. bristande motivation, inte för att det är sant (mätfel).
- "Flykten till mittpunkten" kan begränsas genom att ha fler skalpunkter och erbjuda svarsalternativ för icke-svar (till exempel "Vet inte").

- Mittpunkten bör ha en etikett som beskriver dess position mittemellan punkterna, så att den inte tolkas som något svarsalternativ för allmän osäkerhet (till exempel "Osäker" eller "Vet inte riktigt").
- Forskningsresultaten varierar en del när det gäller mittpunkten. Rekommendationen är ändå att generellt använda en mittpunkt men att vara medveten om hur det kan påverka resultatet.

Alternativ för icke-svar ("vet inte", "ingen åsikt" och "inte aktuellt")

Svarsalternativ som "Vet inte", "Ingen åsikt" eller "Inte aktuellt" brukar kallas för icke-svar. Det beror på att de är tänkta att fånga upp de respondenter som inte kan placera sig på något av de ordinarie svarsalternativen. Det kan vara värt att inledningsvis påpeka att de svarsalternativ som oftast är icke-svar ibland kan vara meningsbärande och en del av ordinarie svarsalternativ. Det beror på frågans syfte. Om frågan syftar till att mäta huruvida respondenten vet någonting eller inte så är "Vet inte" ett högst meningsbärande svarsalternativ och inte något icke-svar. I detta avsnitt behandlas dock de situationer där dessa svarsalternativ kompletterar ordinarie svarsalternativ. Den mesta forskningen är gjord på attitydundersökningar. Därför är "Ingen åsikt" det mest utforskade alternativet för icke-svar.

Om man ska ha ett "Ingen åsikt"-alternativ i attitydfrågor har länge varit en het fråga i surveylitteraturen. Debatten inleddes med att Converse (1970) hävdade att respondenter i undersökningar utan "Ingen åsikt" som svarsalternativ ibland rapporterar att de har en attityd trots att de rimligtvis inte kan ha det (Krosnick, Holbrook, Berent, Carson, Hanemann et. al, 2002). I en klassisk studie från 1940-talet var det 70 procent som angav en åsikt om påhittade sakfrågor – det vill säga sakfrågor som inte existerade och som ingen därför borde kunna ha någon åsikt om. Denna studie och även andra indikationer väckte en diskussion om åsiktsfrågor verkligen mäter endast åsikter. Studien har senare kritiserats och avfärdats. En mer nutida studie syftade till att göra om attitydmätningen med påhittade ämnen men under kontrollerade former. Resultatet visade att 30 procent angav en åsikt trots att attitydobjektet ("Agricultural trade act" i den studien) inte existerade. Det förekom dock att respondenter öppet erkände att de inte visste men i slutändan ändå valde ett svarsalternativ. De 30 procenten sjönk dock till

10 procent om ett "Ingen åsikt" fanns som svarsalternativ (Schuman & Presser, 1980).

Att ha ett "Ingen åsikt" som alternativ skulle alltså kunna rensa bort dem som faktiskt inte har någon åsikt och på så sätt förbättra mätningen. Detta förutsätter emellertid att de som väljer "Ingen åsikt" faktiskt inte har någon åsikt och att de skulle ha fabricerat en åsikt om inte det svarsalternativet hade erbjudits. Det finns dock motsatta idéer – att vissa respondenter väljer "Ingen åsikt" som en enkel väg vidare trots att de faktiskt har en åsikt (så kallad "satisficing"). Det finns alltså två tendenser: 1) utan "Ingen åsikt" hamnar respondenter utan en åsikt på svarsskalan eller hoppar över frågan, och 2) med "Ingen åsikt" väljer en del som faktiskt har en åsikt det alternativet på grund av låg motivation. Dessa olika tendenser går åt olika håll men vilken har allvarligast konsekvenser?

En del av problemet bottnar i att det inte är alldeles självklart vad en åsikt eller en attityd är. De 30 procent som angav åsikter om fiktiva ämnen i undersökningen ovan ("Agricultural trade act") kan ha besvarat frågorna utifrån sina åsikter på en mer generell nivå än frågan (vars attitydobjekt alltså var påhittat), till exempel utifrån deras syn på till exempel handelsregleringar i allmänhet. Betyder det att deras svar var felaktigt och fabricerat? Förfarandet, att formulera en åsikt om något specifikt utifrån en mer generell princip, är förmodligen relativt vanligt i attitydundersökningar. Det beror på att människor inte alltid har färdiga åsikter om allt som de tillfrågas om i undersökningar (läs mer i kapitel 12). Det är alltså inte helt självklart vad som är en åsikt eller inte.

Frågan om icke-alternativ ska erbjudas eller inte är inte nödvändigtvis antingen-eller utan det finns kompromisser. I intervjuer kan alternativet finnas dolt, det vill säga att det erbjuds först om respondenten tvekar eller att intervjuaren markerar det på egen hand om han eller hon bedömer att respondenten inte har någon åsikt. Det erbjuds alltså endast "vid behov". Ett liknande tillvägagångssätt kan i viss mån användas i webbenkäter. Exempelvis kan svarsalternativet erbjudas först om respondenten går vidare i enkäten utan att ha besvarat en viss fråga. Forskning har visat att det kan ge mindre partiellt bortfall (Derouvray & Couper, 2002).

Hur ska man annars göra? Det finns trots allt ganska mycket forskning som jämfört skillnader i datakvalitet mellan att erbjuda "Ingen åsikt" och att inte göra det. Forskningen är dock inte entydig. Studier har visat att

”Ingen åsikt” kan förbättra, försämma eller inte ha någon betydelse för datakvaliteten (Andrews 1984; Alwin & Krosnick 1991; Alwin 2007; McClendon & Alwin 1993; Scherpenzeel & Saris 1997; Wikman & Wärneryd, 1990; Krosnick, Holbrook, Berent, Carson, Hanemann et. al, 2002). Helt klart behöver man som frågekonstruktör ta ställning till både respondenternas motivation, hur pass troligt det är de inte har en åsikt om frågeområdet och om icke-alternativet är ett rimligt alternativ eller inte, givet frågeområdet och respondenterna. Vissa saker kan respondenterna förväntas veta (se nedan) och då framstår ett ”Vet inte” -alternativ som överflödigt.

EXEMPEL:**Hur bra eller dåligt känner du till reglerna för föräldrapenning?**

Mycket bra
Ganska bra
Ganska dåligt
Mycket dåligt

Ett skäl till att ha med ett icke-svar är att man då kan vara ganska säker på att de som svarat på ordinarie svarsalternativ faktiskt hör hemma där och inte tvingats dit på grund av bristande svarsalternativ. Detta kan vara en trygghet i analyskedet. Respondenterna kan också uppfatta det som positivt att frågekonstruktören erbjuder ett icke-alternativ. En generell rekommendation är därför att inkludera ”Ingen åsikt” men att göra en bedömning utifrån det som stått tidigare i avsnittet.

Den mesta forskningen om icke-svar bygger på attitydfrågor och svarsalternativet ”Ingen åsikt”. Det finns mindre forskning om ”Vet inte” vid faktafrågor men en studie (Poe et. al, 1988; Krosnick, Holbrook, Berent, Carson, Hanemann et. al, 2002) visar dock, i likhet med en hel del forskning om ”Ingen åsikt”, att reliabiliteten var densamma oavsett om ”Vet inte” fanns som svarsalternativ eller inte.

Sammanfattande punkter:

- Om icke-svar saknas som svarsalternativ tvingas respondenter som frågan inte är relevant för att ändå ta ställning (till exempel för att de inte har en åsikt).
- Om icke-svar finns som svarsalternativ kan, å andra sidan, respondenter välja det p.g.a. låg motivation.

- Forskningen varierar huruvida datakvaliteten blir bättre eller sämre med alternativ för icke-svar.
- I intervjuer (och i viss mån i webbenkäter) kan icke-svar erbjudas "vid behov".
- Det finns också ett respondentperspektiv, att respondenterna inte ska känna sig tvingade till svar som inte stämmer.
- Vid beslutet om icke-svar ska ingå bör man också utgå från frågans innehåll. Hur relevant är frågan för respondenterna? Hur motiverade är de?
- Om icke-svar finns kan man i alla fall vara (mer) säker på att de som ligger på svarsskalan faktiskt hör hemma där.

Beskrivande etiketter till skalpunkterna

Ett annat beslut för frågekonstruktören är om skalstegen ska ha beskrivande etiketter eller inte. Med det menas att skalstegen har beskrivningar som med ord definierar skalsteget, till exempel "mycket bra", "ibland" eller "instämmer helt".

Som beskrivits ovan hänger detta beslut ihop med antalet skalsteg i skalan. Vid fler än sju (bipolära) eller fem (unipolära) skalsteg börjar språket tryta för att kunna beskriva skillnaderna mellan skalstegen. Det är vanligt att åtminstone ändpunkterna på skalan har etiketter. Detta är oftast nödvändigt för att respondenterna ska förstå svarsdimensionen. I bipolära skalor är det också vanligt att mittpunkten har en etikett. I övrigt kan det variera om skalstegen har en etikett, en siffra, både och eller är helt blanka.

När man skriver etiketter bör man tänka på att de ska ha relativt precisa betydelser för respondenterna. De måste också uttrycka lika stora intervall längs dimensionen (Krosnick & Fabrigar, 1997).

Det kan vara viktigt att känna till att etiketten bara är en av flera faktorer som kan påverka hur skalpunkterna tolkas. Siffror vid skalpunkterna kan också påverka (läs mer i avsnittet *Tolkning av svarsalternativens numrering* nedan). Skalpunktens position kan också ha betydelse – samma etikett kan ges olika betydelser beroende på skalpunktens position i skalan. Exempelvis betyder "Bra" olika mycket beroende på om det är det näst högsta alternativet (7,3 på en tiogradig numerisk

skala) eller det tredje högsta alternativet (5,4) (Wilson, Altman, Whitaker & Callegaro, 2004).

Ett förekommande argument för att använda siffror istället för etiketter är att skalan då blir mer av en intervallskala med jämnt avstånd mellan skalstegen. Det är möjligt att siffrorna signalerar en konstant måttenhet (så kallad ekvidistans) till respondenterna men det är inte givet hur detta går till eftersom den underliggande dimensionen (det siffrorna betyder) är densamma oavsett om vi namnger skalstegen med ord eller inte. Respondenten borde ändå behöva omvandla siffrorna för att kunna anpassa sitt svar till svarsdimensionen (Krosnick & Fabrigar, 1997; Krosnick, Judd & Wittenbrink, 2005).

En studie undersökte skillnader i mätning mellan att ha etiketter vid alla skalsteg jämfört med att bara ha det vid ändpunkterna. Varianten med etiketter vid alla skalsteg hade högre tillförlitlighet. De undersökte även hur länge olika etiketter var i blickfånget och fann att respondenterna fokuserar längre på ändpunkterna om de är de enda skalpunkterna som har etikett, jämfört med om alla skalpunkter har det. Trots detta fungerade alltså skalan med etiketter vid varje skalpunkt bättre, förmodligen för att de mellersta skalpunkternas etiketter bidrar till att respondenten får en bättre uppfattning av skalan som helhet (Menold, Kaczmirek, Lenzner & Neusar, 2014).

Annan forskning visar samma sak - att etiketter ökar tillförlitligheten eftersom de hjälper respondenterna att förstå skalpunkternas innebörd (Krosnick & Berent, 1993). Det är också lättare att redovisa resultatet på ett förståeligt sätt till tredje part om etiketter används. Därför rekommenderas generellt etiketter till skalstegen.

Sammanfattande punkter:

- Etiketter betyder att skalstegen definieras med ord.
- Om skalan har många skalsteg kan det bli svårt att formulera etiketter som tydligt skiljer sig åt.
- En idé är att skalan blir mer av en intervallskala om skalan har siffror än om den har etiketter. Men det är inte givet hur detta fungerar eftersom respondenterna ändå borde behöva omvandla siffrorna till den underliggande svarsdimensionen för att kunna svara.
- Forskningen antyder att etiketter förbättrar mätningen eftersom det hjälper respondenten att förstå skalstegens innebörd.

Siffror och färg

Respondenterna använder ibland ledtrådar eller tumregler när de tolkar visuell information. En sådan tumregel är att objekt som är visuellt lika också är konceptuellt lika (lagen om likhet). En studie undersökte detta med hjälp av färger. Hypotesen var att om färgnyanserna vid ändpunkterna är lika varandra så uppfattas skalan ha kortare konceptuell vidd än om färgnyanserna är mer olika. Denna hypotes bekräftades, men gällde bara när ändpunkterna hade etiketter och inte de mellanliggande skalpunkterna. Numrering av skalstegen hade starkare effekt än färgnyanser. Dessa resultat tyder på någon form av hierarki av visuell information, där siffror upplevs som viktigare än färg av respondenterna. (Tourangeau, Couper & Conrad, 2007). Det viktiga är att tänka på att även den här typen av information kan påverka respondentens tolkning av svarsskalan.

VAS-skalor

En variant av skattningsskalor är så kallade VAS-skalor ("visual analog scales"). Till skillnad från andra skalor markerar respondenten i VAS-skalor svaren på en linje, inte i en svarsruta. VAS-skalor har funnits länge men är inte så vanliga i surveyundersökningar av flera skäl. De kräver mycket administration efter datainsamlingen och eftersom de är visuella fungerar de inte i telefonintervjuer. En distinktion görs mellan VAS och GRS ("graphic rating scale"). VAS är en blank linje med endast ändpunkterna angivna. GRS har även skalsteg angivna längs linjen. VAS-

skalan i dess rena form är alltså kontinuerlig, det vill säga respondenten är inte begränsad av vissa angivna skalsteg som i en diskret skala.

EXEMPEL:

Väldigt nöjd _____ Väldigt missnöjd

När man skapar en VAS- eller GRS-skala måste man bland annat ta hänsyn till hur lång linjen ska vara, vilka ändpunkterna ska vara, om det ska finnas markeringar för skalstegen och om det ska finnas en mittpunkt. Beroende på vägvalen kan skalan i praktiken bli ganska lik en diskret skala.

VAS-skalan har en del potentiella fördelar men hur fungerar den i jämförelse med andra skalor med fasta svarsalternativ? Det finns en del ny forskning om detta i webbundersökningar som jämför VAS-skalor (som har så kallade "sliders", en markör som kan flyttas längs en linje) med skalor med fasta svarsalternativ ("radioknappar"). En studie hittade ingen skillnad vad gäller validiteten (Bayer & Thomas, 2004) men kom fram till att VAS-skalan tog längre tid att använda. En annan studie prövade många olika varianter men kom fram till samma sak samt och att VAS-skalorna dessutom hade högre andel partiellt bortfall. En rekommendation kan därför vara att undvika VAS-skalor framförallt i webbundersökningar, om man inte har ett tydligt syfte med just den skaltypen (Couper, 2008).

Slutna svarsalternativ utan given inbördes ordning

I slutna svarsalternativ utan given inbördes ordning är svarsalternativen självständiga. De kan särskiljas men inte rangordnas (till exempel olika idrotter eller politiska partier). Denna typ av svarsalternativ kallas nominalskala.

EXEMPEL:

Vilken är din nuvarande sysselsättning? Du kan ange flera alternativ

Yrkesarbetar
Tjänstledig eller föräldraledig
Studerar, praktiserar
Arbetsmarknadsåtgärd
Arbets sökande

Ålderspensionär
Avtalspensionär
Förtidspensionerad, sjukpensionerad
Sjukskriven
Långtidssjukskriven (mer än 3 månader)
Sköter eget hushåll

För vissa typer av frågor finns det alltså ingen given ordningsföljd för svarsalternativen. De bör ändå ha någon form av struktur, speciellt om de är många till antalet. Annars kan det bli väldigt tidsödande och rörigt för respondenterna att leta reda på rätt svarsalternativ.

Om frågan innehåller många svarsalternativ som inte går att gradera kan det vara lämpligt att överväga att dela upp dem i flera frågor eller åtminstone gruppera svarsalternativen (ESS, 2006). I brist på annat kan man alltid använda bokstavsordning. Ett annat alternativ, särskilt om listan med svarsalternativ är lång, är att ha de vanligaste alternativen först för att minska respondentbördan.

När listan med svarsalternativ är lång ökar dock risken för positionseffekter – att svarsalternativens position påverkar svarsfördelningen. Denna risk är större vid nominalskalor än vid andra typer av skalor (läs mer nedan men framförallt i kapitel 13).

Sammanfattande punkter:

- I nominalskalor – svarsalternativ som kan särskiljas men inte rangordnas – bör det ändå finnas något slags ordning, speciellt om alternativen är många, så att respondenten kan orientera sig.

Ja-nej

Den enklaste typen av nominalskala är de skalor som bara innehåller två alternativ, till exempel "Ja" eller "Nej". Ja-nej-skalor är ganska vanliga i frågeformulär. Det är lätt att förstå varför. De innebär ett tydligt och lättolkat ställningstagande. Det är dock viktigt att vara medveten om att ja-nej i vissa fall kan innebära en kraftig informationsförlust. Det kan finnas många olika nyanser eller grader mellan "Ja" och "Nej" som missas. Ja-nej-skalor kan alltså bidra till en förenklad bild (Wärneryd, 1990).

EXEMPEL:

Har du varit arbetslös någon gång under de senaste tolv månaderna?

Ja

Nej

För mätningens skull, och även med tanke på att respondenterna ska kunna känna att de gett en rättvis bild, kan det ibland finnas goda skäl att ha en mer nyanserad skala än ja-nej.

Ja-nej-skalar bör sålunda användas med viss försiktighet när det gäller att mäta saker som kan variera i mer än två punkter. Med andra ord kan det finnas vissa risker med att använda ja-nej-skalar som ordinalskala (med endast två skalpunkter). När det gäller mätområden som är tydligt nominala (egenskaper som kan särskiljas men inte rangordnas, till exempel om respondenten äger en bil) är tvärtom ja-nej-skalan ofta den mest lämpliga.

Sammanfattande punkter:

- Ja-nej ger ett tydligt resultat men kan, beroende på frågans innehåll, innebära en informationsförlust och framstå som inskränkande av respondenterna.
- Om frågan ska mäta sådant som kan variera i mer än två punkter finns det, av flera skäl, ett värde att ha en mer differentierad skala.
- I övrigt är ja-nej väldigt tydliga svarsalternativ.

Rankingskalor

I rankingskalor ska respondenten utifrån frågans syfte skapa en ordningsföljd (till exempel från viktigast till minst viktigt) av svarsalternativ som inledningsvis är osorterade. Skalan besvaras oftast genom att respondenten sätter en siffra bredvid varje svarsalternativ enligt ett givet system, till exempel att 10 = viktigast. Det finns även andra lösningar.

Rankingskalan har en del egenskaper som kan innebära svårigheter för respondenterna. Det är en ganska avancerad kognitiv uppgift att ranka objekt. Varje objekt behöver inte bara bedömas för sig, vilket är standarden i nominalskalor, utan även jämfört med alla andra objekt i

skalan. Dessutom måste en ordningsföljd upprättas bland objekten. För att skalan ska fungera krävs det att respondenten strängt följer formatet. Det går generellt inte att ge flera objekt samma position. Samtidigt ger rankingformatet mer frihet än traditionella format, vilket kan leda till att respondenten hittar på egna lösningar.

Formatet tar betydligt längre tid än andra format (tre gånger så lång tid enligt en studie; Munson & McIntyre, 1979). Rankingskalor kräver att mycket information ska bearbetas. Den kräver därför i praktiken visuellt stöd och fungerar mindre bra i telefonintervjuer (Alwin & Krosnick, 1985). Data från rankingfrågor kräver också särskild hänsyn eftersom svaren är beroende av varandra (Blasius, 2012). Ofta kan dessutom en lösning där respondenterna får gradera varje objekt för sig, till exempel på en skala 1-5, fungera som likvärdigt alternativ till en rankingskala. En rekommendation är därför att undvika rankingskalor om det inte finns särskilda skäl till att använda dem.

En studie (Blasius, 2012) använde webbens tekniska möjligheter för att testa olika format av rankingskalor:

Format 1 innebär ranking med hjälp av pilar som förflyttar objekt uppåt eller neråt tills önskad ranking uppnåtts.

Format 2 innebär ranking genom att flytta objekt från en box till en annan med datormusen.

Format 3 innebär ranking genom uppdelning av frågan, där respondenten endast markerar de viktigaste och minst viktiga alternativen. Sedan ställs frågan igen med de tidigare valda alternativen borttagna. Proceduren upprepas tills alla alternativ har täckts.

Format 4 innebär ranking genom att sätta en siffra som motsvarar objektets position bredvid varje objekt. Denna variant utnyttjar inte webbens möjligheter utan är samma format som brukar användas i pappersenkäter.

Blasius (2012) rekommenderar formatet med att dra objekt (format 2 ovan) till en ny box om det inte är väldigt många objekt som ska rankas. Formatet är en relativt simpel manöver och rankingen blir visuellt tydlig i den nya boxen, jämfört med att sätta siffror vid varje objekt där respondenten själv måste visualisera rankingen. Både formatet med pilar och att dela upp rankingen i flera frågor ökar respondentbördan och kan leda till både mätfel och partiellt bortfall.

Sammanfattande punkter:

- Rankingskalor kräver mycket av respondenten – att objekt ska bedömas i förhållande till varandra och att en ordning ska upprättas.
- Det finns även risker för att respondenterna kommer att hitta på egna lösningar.
- I telefonintervjuer fungerar rankingskalor dåligt. I webbenkäter finns det fler lösningar än i pappersenkäter.
- En rekommendation är att undvika rankingskalor, speciellt i pappersenkäter och telefonintervjuer. De bör endast användas om det finns särskilda skäl och i så fall förenklas så långt som möjligt.

Checklista

En vanlig typ av fråga och svarsalternativ är så kallade checklistor där svarsalternativen består av en lista av svarsalternativ med en ruta per svarsalternativ. Respondenterna ska "bocka av" de alternativ som stämmer för dem men lämna övriga svarsrutor blanka. Ofta finns det en instruktion i stil med "Markera alla alternativ som stämmer för dig" ("mark all that apply" på engelska).

Frågetypen har flera fördelar. Genom möjligheten att täcka in många olika områden på en gång är både utrymmet och respondentbördan minimal. Frågetypen är dock även ifrågasatt. Det är svårt att bedöma vad avsaknaden av markering egentligen betyder. Respondenten kan till exempel ha bedömt att alternativet inte är relevant, missat alternativet eller inte kunnat avgöra om det är relevant eller ej.

Ett annat skäl är att forskning visar att det görs färre markeringar i checklistor än om varje alternativ hade ställts som en ja-nej-fråga (Rasinski, Mingay & Bradburn, 1994; Smyth, Dillman, Christian & Stern, 2005). Detta är viktigt att känna till, speciellt för undersökningar med kombinerad insamling med både telefonintervju och självadministrerat. I telefonintervjun är frågor i formen av checklistor svåra att genomföra och blir av nödvändighet ja-nej-frågor. Skillnader i frågeformatet kan då leda till olika mätningar mellan datainsamlingsmetoderna.

En studie (Smyth, Dillman, Christian & Stern, 2005) undersökte tre potentiella förklaringar till skillnaden i svar mellan checklista och ja-nej-frågor:

1. Satisficing (att respondenten svarar med minsta möjliga ansträngning, läs mer i kapitel 9) undersöktes genom att studera primacy-effekter (se avsnittet Positionseffekter nedan).
2. Hur djupt informationen i frågorna begrundas, vilket mättes genom svarstiden.
3. Hålla med-tendenser undersöktes genom att lägga till ett tredje svarsalternativ – "Neutral" eller "Vet inte" – på några frågor.

Resultatet gav mest stöd för förklaringen om hur länge informationen i frågan begrundas. Det verkar som om respondenterna funderar längre på frågor uppställda i ja-nej-format jämfört med checklistor. Ja-nej-frågor tycks alltså leda till en djupare informationsbehandling. Det är dock viktigt att uppmärksamma att denna studie bara undersöker orsaker till skillnader men att vi inte med säkerhet kan veta vilket av de två frågeformaten som kommer närmast sanningen. För det ändamålet krävs separata valideringsstudier (Smyth, Christian & Dillman, 2008; Smyth, Dillman, Christian & Stern, 2005).

På SCB:s enkätenhet finns erfarenheten att respondenter ibland besvarar listor av ja-nej-frågor i tabellform som om de vore checklistor. Det innebär att de bara fyller i frågorna som de svarar "Ja" på och lämnar frågorna för "Nej" blanka, vilket då blir partiellt bortfall.

Sammanfattande punkter:

- Checklistan har många fördelar men även nackdelar – vad betyder avsaknaden av markering?
- Checklistan ger färre markeringar än antalet "Ja" i ja-nej-frågor med samma alternativ.
- Vår erfarenhet på SCB är att respondenterna ibland besvarar listor av ja-nej-frågor som checklistor, det vill säga enbart markerar "Ja".

Scenariobeskrivningar

Det händer att svarsalternativen har formen av längre scenariobeskrivningar. Med det menas att de beskriver ett scenario eller

en situation mer utförligt än vanligt. När man skriver sådana scenarier måste man vara noggrann med att scenarierna inte överlappar för mycket så att det blir svårt för respondenterna att välja. En annan problematik är att ju mer detaljerade scenarierna är, desto större är risken att någon del inte överensstämmer med respondenten, som då kanske avfärdar alternativet. Denna risk kan mildras genom att till exempel lägga till "Välj den beskrivning som passar bäst in på dig".

Den här typen av svarsalternativ är primärt på nominal nivå men kan även vara på ordinalnivå. Då krävs det att man verkligen har det i åtanke när scenarierna skrivs.

Positionseffekter

Nominalskalor, svarsalternativ utan inbördes ordning, är den skaltyp som är mest drabbad av så kallade positionseffekter. Det betyder att svarsalternativens position i en lista påverkar svaren jämfört med en annan position. Detta beror sannolikt på att nominalskalor är den skaltyp som oftast har flest antal svarsalternativ. Ibland är det långa listor där alternativen dessutom är fristående och saknar inbördes ordning.

Positionseffekter kan vara av främst två typer. Ett förhöjt stöd för svarsalternativ som kommer tidigt i listan (kallas för "primacy effect" på engelska) eller för svarsalternativ som kommer sent i listan (kallas för "recency effect" på engelska). Primacy-effekter är vanligast i självadministrerade enkäter och recency-effekter i telefonintervjuer. Dessa effekter behandlas inte vidare i detta kapitel men finns utförligt beskrivna i kapitel 13, Ordningsföljder och kontext.

Svarsalternativ i intervjufrågor

Mycket av innehållet i detta kapitel gäller även vid intervjufrågor, men intervjun har några egna förutsättningar som kan påverka vilken typ av svarsalternativ man väljer som frågekonstruktör. Den stora skillnaden jämfört med enkät är att vid intervjuer kan inte respondenten själv läsa svarsalternativen. Det finns ett antal olika format man kan använda som frågekonstruktör:

1. Öppen fråga utan svarsalternativ – respondentens spontana svar registreras i klartext av intervjuaren

2. Öppen fråga med underliggande svarsalternativ som inte delges respondenten

EXEMPEL:**Varför valde du att avbryta dina studier?**

Jag fick arbete

Innehållet motsvarade inte mina förväntningar

Det var för höga krav

...

3. Angiven svarsdimension i frågeformuleringen

EXEMPEL:**Hur ofta träffar barnet sin andra biologiska förälder?**

Flera gånger i veckan

Ungefär en gång i veckan

Ungefär varannan vecka

...

4. Svarsalternativen ingår i frågeformuleringen

EXEMPEL:**Bor du i eget rum eller delar du rum med någon?**

Eget

Delar

5. Ja-och-nej-fråga

EXEMPEL:**Använder du glasögon?**

Ja (alltid eller ibland)

Nej

6. Svarsalternativen läses upp efter avslutad frågeformulering

Vilket format (1-6) man använder i frågan får förstås stor påverkan på respondentens svarprocess men även på intervjuarens roll. Om formatet fungerar som tänkt innebär 1-3 att intervjuaren tolkar respondentens svar och antingen återger det i klartext eller väljer rätt underliggande

svarsalternativ. Det blir alltså intervjuaren, på sätt och vis, som väljer svarsalternativ åt respondenten. Respondenten själv känner inte till svarsalternativen eftersom de inte har lästs upp. Detta kallas för fältkodning, intervjuaren kodar respondentens fria svar, vilket kan vara förknippat med problem. Intervjuaren måste ta ett snabbt beslut och kan missförstå eller missbedöma respondentens fria svar. Risken är störst i formattyperna 2 och 3 ovan. Hur stor risken för problem är beror bland annat på hur avgränsat frågeämnet är. I exempel 2 ovan kan till exempel respondenten komma att svara något som inte riktigt passar in på något alternativ eller som passar in på flera.

Format 3 innebär i grunden mindre tolkningsutrymme än format 1 och 2 eftersom frågan då åtminstone är avgränsad i svarsdimensionen ("Hur ofta"). Det finns förstås en risk för fel om inte svarsdimensionen i frågan och svarsalternativen överensstämmer som i format 3 (svarsalternativen är mer specifika än frågeformuleringen som kan ge svar som "Titt som tätt" eller "Ganska ofta").

Format 4 och 5 innebär inte någon risk för fältkodning. Respondenten styrs genom frågeformuleringen till giltiga svarsalternativ som finns representerade i intervjuarens datorstöd. Detta gäller om respondenten följer formatet, vilket inte är helt säkert.

Om format 1-3 är förknippade med risker för tolkningsfel, varför överhuvudtaget överväga dem? Det säkraste är att läsa upp svarsalternativen (format 6 ovan) så varför inte göra det varje gång? Det har med tre saker att göra:

- frågans ämnesområde
- hur finfördelad informationen behöver vara för dataanalysen
- intervjuens begränsning – att frågor och svarsalternativ behöver vara korta.

Om ämnesområdet är lättförståeligt och om det är fullt tillräckligt att få en indelning på grov nivå så kan en ja-nej-fråga vara den bästa lösningen. Men för vissa ämnen och informationskrav blir sådana smala format inskränkande. Att tvinga in en fråga i en "för trång kostym" kan dessutom få allvarigare konsekvenser än fältkodning eftersom respondenterna då kan komma att bryta formatet, till exempel för att de inte tycker att det beskriver deras verklighet korrekt. På liknande sätt kan det i vissa typer av frågor vara en omöjlighet att läsa upp alla svarsalternativ (Format 6). Om svarsalternativen är många ökar de

respondentbördan, administrationstiden för frågan och risken för fel eftersom respondenten inte kan hålla kvar alla alternativ i huvudet. I sådana situationer kan ett mer fritt format vara att föredra (1–3).

Alltså, ju lägre siffra i listan ovan (från 6 till 1) desto mer tolkningsutrymme – att intervjuaren behöver tolka eller följa upp respondentens svar. Man måste också ta hänsyn till ämnesområdet och vilken informationsnivå som krävs. Alla format kan vara förknippade med fel och man får avgöra vilket som är bäst i en given situation. Det gäller att hitta en bra balans mellan fördelar och nackdelar. Svarsalternativens format är alltså ett viktigt beslut vid frågekonstruktion i intervjuer. Utfallet är oftast inte givet utan det går i många fall att strukturera ett innehåll på flera sätt, till exempel genom att dela upp en generell fråga med format 1 till flera specifika delfrågor med format 6.

Sammanfattande punkter:

- En stor skillnad mot enkät är att vid intervjuer kan inte respondenten själv läsa svarsalternativen.
- Valet av svarsformat påverkar både respondentens svarsprocess och intervjuarens roll.
- Det finns 6 format som varierar från att respondenten får svara fritt (och intervjuaren återger svaret i klartext) till att alla svarsalternativen läses upp av intervjuaren.
- Om inte respondenten känner till svarsalternativen blir det, på sätt och vis, intervjuaren som väljer åt respondenten. Detta kan leda till fel.
- Mer reglerade format har dock sina egna svagheter (de kan vara inskränkande eller alternativen kan vara för många för att läsas upp).
- Vilket format som passar en fråga bäst är därför en bedömningsfråga.

Svarsalternativ och tolkningar

Detta avsnitt redovisar utifrån ett antal klassiska forskningsresultat hur respondenterna kan tolka, tänka kring och använda svarsalternativen. Syftet är att ge ökad förståelse för svarsalternativens funktion när respondenten besvarar frågan.

Tolkning av svarsalternativens numrering

Att ha etiketter vid alla skalsteg ger mer tillförlitliga svar än att bara ha etiketter vid ändpunkterna (se ovan). Sådana resultat visar att respondenterna faktiskt uppmärksammar beskrivningarna vid skalstegen. Det finns en del forskning som undersökt hur andra delar av svarsformatet kan påverka. En studie av Schwarz, Knäuper, Hippler, Noelle-Neumann & Clark (1991) studerade hur siffror i kombination med etiketter kan påverka svaren. Frågan gällde hur framgångsrika respondenterna upplevde att de hade varit i livet. De olika svarsskalorna var:

0	10
Inte alls framgångsrik	Extremt framgångsrik
-5	+5
Inte alls framgångsrik	Extremt framgångsrik

Det är lätt att se formatet, i detta fall kombinationen av siffror och etiketter, som en neutral del av mätinstrumentet. Men ovanstående studie tyder på att det inte alls är så. I skalan som gick från 0 till 10 var det 34 procent som placerade sig på den lägre halvan. Motsvarande del på skalan -5 till +5 var endast 13 procent – en skillnad på mer än 20 procentenheter (Schwarz, Knäuper, Hippler, Noelle-Neumann & Clark, 1991).

Denna typ av resultat förbryllar. Om man begriper den bakomliggande mekanismen blir resultatet däremot förstäligt och något som kan användas i frågekonstruktionen. Ovanstående svarsskalor är endast definierade i ändpunkterna. Detta gör att etiketterna vid ändpunkterna får stor betydelse i hur svarsdimensionen som helhet uppfattas. Vad betyder då "Inte alls framgångsrik"? Det är vagt och kan innebära och innefatta många olika saker. Det som verkar hända (Schwarz m.fl. bekräftade även detta i ett senare experiment) är att när respondenterna försöker förstå vad ändpunkten betyder så får den tillhörande siffran betydelse. När "Inte alls framgångsrik" finns tillsammans med siffran noll tolkas det som "avsaknad av framgång". Det är inte svårt att föreställa sig denna tolkning eftersom noll ofta betyder avsaknad av någonting. När "Inte alls framgångsrik" finns tillsammans med -5 uppfattas det däremot som "misslyckad", vilket är mycket starkare än

”avsaknad av framgång” (Schwarz, Knäuper, Hippler, Noelle-Neumann & Clark, 1991).

Utifrån denna förklaring är resultatet i studien lättare att förstå.

”Misslyckad” är i själva verket motsatsen till ”framgångsrik”. Det verkar alltså som om det undre svarsformatet (-5 till +5) gör att skalan uppfattas som bipolär (går från en pol till dess motsats) snarare än unipolär (en kontinuerlig ökning). Inte heller detta är förvånande eftersom -5 till +5 beskriver ett bipolärt förhållande (med 0 som neutral mittpunkt) medan 0 till 10 är unipolärt (Schwarz, Knäuper, Hippler, Noelle-Neumann & Clark, 1991).

Vad ska man då dra för slutsats av detta? Jo, att man bör vara tydlig med svarsdimensionen så att respondenten inte behöver förtydliga den på egen hand. Det är också viktigt att se till att svarsskalans siffror matchar skalans format. Om skalan är unipolär ska inte siffror som antyder att den är bipolär användas (Schwarz, Knäuper, Hippler, Noelle-Neumann & Clark, 1991).

Sammanfattande punkter:

- Perifer information, som inte är avsedd att vara meningsbärande, kan ändå påverka respondenterna.
- Ett exempel är skalans numrering – om den går från minus till plus snarare än från noll och uppåt.
- Förmodligen beror det på att skalan uppfattas som bipolär (där ändpunkterna är motsatser) om siffrorna går från minus till plus.
- Svarsskalan bör alltså vara tydlig och perifer information bör vara i linje med svarsskalan, inte signalera något motstridigt.

Den psykologiska mittpunkten

Hur många pinnar behövs för att få ett tio meter långt staket som har en meter mellan varje pinne? Ett sätt att försöka besvara frågan kan vara att använda division (dela 10 med 1) vilket ger svaret tio. Detta är dock felaktigt eftersom det missar att ta hänsyn till den första pinnen. Det rätta svaret fås genom samma division, men att den första pinnen läggs till (elva pinnar behövs) (Zigerell, 2011).

Motsvarande tankefel kan även förekomma i enkäter. Exempelvis att 5 skulle vara mittpunkten i en skala som går från 1 till 10. Det är det inte.

En tiogradig skala har ingen mittpunkt utan mitten på skalan ligger mittemellan skalpunkt 5 och 6, vilket inte representeras med någon egen punkt. Att respondenter som vill vara neutrala väljer skalpunkt 5 innebär alltså en underskattning i den typen av skala (Zigerell, 2011).

Problemet med att den psykologiska och faktiska mittpunkten inte överensstämmer kan motverkas genom att ha beskrivningar vid skalpunkterna som tydliggör deras innebörd eller att ha en skala med elva skalsteg (om man vill erbjuda en mittpunkt).

Sammanfattande punkter:

- Den psykologiska och faktiska mittpunkten är inte alltid förenliga och det kan behöva förtydligas i skalan (5 är inte mittpunkten på en skala som går från 1-10, även om det kan framstå som så).

Tolkning av frekvensskalor – position eller frekvens?

Svarsskalor är inte endast ett neutralt mätformat utan även bärare av information till respondenterna (se ovan). Sällan blir detta tydligare än i ett experiment av Schwarz, Hippler, Deutsch & Strack (1985) om frågor om tv-tittande. Författarna resonerade att svarsskalan vid beteendefrågor kan av respondenterna uppfattas uttrycka frågeställarens förväntan och kunskap om hur verkligheten ser ut. Alltså, att intervallen i en frekvensskala motsvarar en förväntad, "normal", fördelning. Respondenterna kan då använda svarsskalan utifrån denna förväntning. För att testa detta genomfördes ett experiment där frågan "Hur många timmar tittar du på tv en vanlig dag?" jämfördes med två olika svarsskalor: 1) en lågfrekvensvariant som gick från "upp till en halvtimme" till "mer än 2,5 timmar" och 2) en högfrekvensvariant som gick från "upp till 2,5 timmar" till "mer än 4,5 timmar" (Schwarz, Hippler, Deutsch & Strack, 1985).

Resultatet visade att med lågfrekvensvarianten var det 16 procent som tittade på tv 2,5 timmar eller mer. Med högfrekvensvarianten var motsvarande siffra 38 procent, mer än dubbelt så många (Schwarz, Hippler, Deutsch & Strack, 1985). Vad beror detta på? Respondenterna tittar väl lika mycket eller lite på tv oavsett hur vi konstruerar intervallen i svarsskalan? Hur ska ett sådant här resultat förstås?

Frågan gällde ett område som det kan vara svårt att svara exakt på. Vad innebär "titta på tv"? I många hem är nog tv:n påslagen på kvällarna utan att personerna tittar hela tiden. Dessutom varierar det säkert stort från dag till dag, så det kan vara svårt att avgöra vad som är en "vanlig dag". Tv-tittande är dessutom en trivial aktivitet som människor inte så noga lägger på minnet. Frågan är alltså svår att svara exakt på. Det verkar därmed som en del respondenter utgår från svarsalternativen som en referensram – en slags förväntad normalfördelning över hur tv-tittande ser ut i verkligheten där det mittersta alternativet är det vanligaste och ytteralternativen extremer. De tycks sedan svara utifrån hur de själva förhåller sig till detta. Istället för att utgå från faktiska timmar av tv-tittande (som uttrycks i svarsalternativens intervall) verkar vissa alltså basera sitt svar på utgångspunkten "hur mycket jag tittar i jämförelse med andra" (som de uppfattar uttrycks i svarsalternativens position) (Schwarz, Hippler, Deutsch & Strack, 1985).

Hur ska man som frågeställare förhålla sig till detta? För det första bör frågan förtydligas. Otydliga frågor borde i högre grad leda till problem i stil med studien ovan. Ett annat alternativ kan vara använda ett öppet svarsalternativ (Schwarz, Hippler, Deutsch & Strack, 1985).

Sammanfattande punkter:

- Respondenter kan ibland utgå från svarsalternativens position snarare än innebörd. Respondenterna kan alltså svara relativt snarare än faktiskt.
- Svarsalternativ kan förväntas återge en slags normalfördelning där det "normala" representeras av alternativet i mitten.
- Problemet kan åtgärdas genom ett förtydligande av frågan eller genom att använda ett öppet svarsalternativ.

Upplevelsen av skalan – ska icke-svar höra till svarsskalan eller stå separat?

Det finns forskning som föreslår att respondenter inte alltid utgår ifrån svarsalternativens innebörd utan ifrån något annat, till exempel deras position, när de besvarar frågorna (se ovan). På liknande sätt visar en studie att en del respondenter utgår från skalans visuella snarare än konceptuella mittpunkt. Svartsfördelningen blir annorlunda om icke-svaren (till exempel "Vet inte" eller "Ingen åsikt") ligger i anslutning till

övriga svarsalternativ, som en del av skalan, jämfört med om de är separerade med till exempel en linje. Anledningen är förmodligen att de uppfattas tillhöra skalan om de inte åtskiljs och att den visuella mittpunkten då förskjuts. Rekommendationen är därför att separera icke-svar från övriga svarsalternativ visuellt (Tourangeau, Couper & Conrad, 2004).

Den visuella mittpunkten kan också förskjutas genom att svarspunkterna i en vertikal skala ligger på ojämnt avstånd från varandra. Det påverkar också svarsfördelningen enligt ovanstående princip (Tourangeau, Couper & Conrad, 2004) och bör därför undvikas.

Sammanfattande punkter:

- Om icke-svar ligger i direkt anslutning till ordinarie svarsalternativ kan de framstå som en del av svarsskalan och påverka vad som uppfattas som mittpunkten på skalan.

Skillnader mellan öppna och stängda frågor

Skillnaden mellan öppna och stängda frågor behöver inte nödvändigtvis vara så stor men kan i vissa fall vara betydande. En öppen fråga kan leda till en annan svarsprocess än en stängd fråga. För att kunna bedöma detta behöver man på djupet känna till möjliga skillnader mellan öppna och stängda frågor.

Svarsalternativen är en ibland underskattad del av en surveyfråga. De är inte något som respondenterna vänder sig till först när de redan har ett färdigt svar. Om de delges respondenten är de en del av svarsprocessen från start. Om svarsalternativen delges respondenten får alltså denne mer information än om de inte skulle ha delgetts. Detta kan få en del konsekvenser.

Förståelse av frågan

Informationen i svarsalternativen kan bidra till respondentens förståelse av frågan. Om till exempel frågeformuleringen är vag så kan svarsalternativen hjälpa respondenten att avgränsa innebörden. Exempelvis visar forskning att man får väldigt olika resultat på frågan "Hur ofta är du arg?" beroende på vilken tidsperiod som innefattas i svarsalternativen. Om intervallen i svarsalternativen är baserade på "senaste veckan" uppfattas "arg" som något betydligt mildare än om

intervallen är baserade på "senaste året". Det är en ganska naturlig tolkning av respondenten. Om det var mild ilska som avsågs vore en referensperiod på ett år orimligt lång med tanke på minnets begränsningar. På motsvarande sätt, om det var extrem ilska som avsågs vore "senaste veckan" en onaturligt kort referensperiod eftersom de flesta personer inte blir vansinniga på veckobasis (Schwarz, Strack, Muller & Chassein 1988).

Svarsalternativen innehåller alltså information som kan påverka respondentens förståelse av frågeformuleringen. Exempelvis kan svarsalternativen korrigera för en vag frågeformulering.

Svarsalternativen ger förslag på möjliga svar

I öppna frågor får respondenten ingen input utöver frågeformuleringen om vad som kan ingå i svaret. De måste själva generera svarsalternativen. Denna process påverkas bland annat av hur de tolkat frågans nyckelbegrepp, vad som är framträdande i minnet, hur grundligt de genomsöker minnet, intervjusituationen och så vidare. I slutna frågor påverkar svarsalternativen respondenten och ger explicita förslag på vad som kan ingå i svaret. Detta är två väldigt olika utgångslägen.

Hur pass öppna de öppna frågorna faktiskt är kan variera. I nominalskalor är inte svarsalternativen lika givna som i ordinal- och intervallskalor som strikt följer en rangordnande svarsdimension (till exempel från "mycket bra" till "mycket dåligt"). Öppna frågor på nominalskalenivå är därför i regel "mer öppna" än de på ordinal- och intervallskalenivå.

Förutom skaltypen kan också ämnet påverka hur pass öppen en fråga är. Frågan "Vilket politiskt parti tycker du bäst om?" är visserligen öppen i formatet men har en naturlig avgränsning i att det finns en begränsad mängd partier att välja mellan. Frågan "Vilken fråga är viktigast för Sverige i framtiden?" är däremot mycket mer öppen eftersom den har många fler tänkbara svar. Så även om frågan har ett öppet format kan skaltypen och ämnet påverka hur pass öppen den är i praktiken.

Skillnaden mellan öppna och slutna frågor på nominalskalenivå borde bli större ju mer "öppen" frågan är ämnesmässigt. Forskningen visar också detta. På frågan "Vad är det viktigaste för barn för att förbereda dem för livet?" var det 4,6 procent som valde "att tänka kritiskt" om frågan var öppen men 61,5 procent valde det alternativet om frågan var stängd

(Schuman & Presser, 1981). En häpnadsväckande skillnad men den är inte är så svår att förstå: frågan är väldigt vag och generell. Det finns så många olika svar som skulle kunna vara relevanta. Om respondenten måste generera möjliga svar på egen hand är det inte säkert att de kommer att komma fram till just samma svar som om frågan var slutna.

Det finns förstås nackdelar med slutna frågor också. De ger ingen information om vad respondenten spontant, på egen hand, kommer att tänka på. Men de avgränsar och inriktar också respondentens tänkande. Det är mindre troligt att respondenten kommer att börja generera egna svar om de blir erbjudna en lista med färdiga alternativ. I slutna frågor blir alltså frågeställarens, på förhand fastställda, svarsalternativ styrande. Hur stora konsekvenser detta får beror på hur bred frågan är och hur heltäckande svarsalternativen är. Om frågan är väldigt bred (som "Vad är det viktigaste för barn för att förbereda dem för livet?") framstår det som omöjligt att erbjuda alla alternativ som skulle kunna vara möjliga för respondenten. Frågeställarens avgränsning blir därmed styrande. I frågor som är mer naturligt avgränsade och listan av svarsalternativ därmed kan bli mer heltäckande, borde konsekvenserna av detta bli mindre (Schwarz, 1999).

I nominalskalefrågor kan skillnaderna mellan öppna och slutna frågor därför bli ganska stora, beroende på hur bred frågan är ämnesmässigt och vilka svarsalternativ som erbjuds.

Det finns alltså olika fördelar med att ha en öppen fråga jämfört med att ge respondenten svar att välja på. Sammanfattningsvis kan svarsalternativen hjälpa respondenten att förstå frågan och minnas saker som de annars inte skulle ha kommit ihåg. Dock kan de verka begränsande (Schwarz, 1999). Hur stora skillnader det kan leda till mellan öppna och slutna frågor beror på hur "självkla" svaren är för respondenten. Detta beror, i sin tur, på skaltypen och ämnesområdet - hur bred frågan är.

Sammanfattande punkter:

- Svarsalternativen kan bidra till frågans förståelse eftersom de kan förtydliga vilken intensitet som avses i frågan.
- "Arg" uppfattas syfta på olika starka känslor beroende på om svarsskalan utgår från dagar eller månader.

- I frågor som är breda ämnesmässigt (finns många tänkbara svar) kan det bli stora skillnader mellan öppna och stängda frågor. Det beror på att i öppna frågor svarar respondenten fritt. I stängda frågor avgränsar svarsalternativen respondenten.

Sammanfattning

Svarsalternativ – sammanfattning

Olika svarsalternativ – olika beslut:

- Svarsalternativen fyller en viktig funktion – de styr svaren.
- Svarsalternativen kan, om respondenten får ta del av dem, hjälpa till med förståelsen av frågan.
- Svarsalternativen kan vara av olika skaltyper beroende på om de kan särskiljas, rangordnas och om måttenheten är konstant.
- Vilka svarsalternativ en fråga ska ha kräver beslut om till exempel antal, hur intervall ska ligga, om det ska finnas mittpunkt eller icke-svar och så vidare.
- Svarsalternativen kan variera på vissa grundläggande sätt (se punktlistan i avsnittet Olika svarsalternativ – olika beslut).

Öppna svarsalternativ där respondenten ska fylla i text:

- I öppna svarsalternativ för text får respondenten formulera svaret själv, med sina egna ord.
- Styrkor: svaret är inte avgränsat till frågekonstruktörens syn på saken, det kan bidra med ett kvalitativt inslag, det kan visa vilka svarsalternativ som är vanligast samt visa på vad som bör ingå som fasta alternativ.
- Svagheter: kräver mer ansträngning av respondenten att besvara samt mer kodning och efterarbete.
- Öppna svarsalternativ för text är lämpligt när antalet möjliga svar är många, när svarsalternativen i sig kan påverka svaren eller när man vill få en djupare förståelse.

Öppna svarsalternativ där respondenten ska fylla i siffror:

- I öppna svarsalternativ för siffror får respondenten skriva siffrorna i svaret själv.

- Styrkor: mer exakta svar än i intervall, undviker tak- eller golveffekter och inga svarsalternativ som i sig kan påverka svaren.
- Svagheter: kan leda till ifyllningsfel och svar som är svåra att tolka.
- En annan svaghet är avrundning: att respondenter inte svarar exakt utan avrundar till vissa punkter, till exempel jämna femtal.
- Det finns en risk att avrundningar kan leda till underskattningar eftersom avståndet mellan avrundningspunkterna tenderar att öka för varje punkt (till exempel 10, 25, 50, 100).

Slutna svarsalternativ – allmänt:

- I slutna svarsalternativ väljer respondenten (oftast) ett av några på förhand givna svarsalternativ.
- Styrkor: förhållandevis enkla att sammanställa och analysera samt kräver mindre ansträngning av respondenten.
- Svagheter: kräver mycket förarbete och många beslut (till exempel om vilka och hur många svarsalternativ som ska ingå) samt begränsar respondenten till frågekonstruktorrens referensram.

Slutna svarsalternativ med given inbördes ordning – vaga kvantifieringar:

- "Vaga kvantifieringar" är när ord används för att beskriva en frekvens eller styrka. Det kan innebära en precisering (till exempel "ganska" eller "väldigt") eller ett enskilt ord som beskriver en frekvens (till exempel "ibland").
- Nackdelen med vaga kvantifieringar är att de inte har någon given betydelse utan utgår från en referenspunkt – "ganska ofta" jämfört med något, men vad?
- Vad som är referenspunkten kan variera mellan olika personer, beroende på till exempel personlighet och erfarenheter.
- Vad som är referenspunkten kan också variera beroende på sammanhang. "Ganska ofta" kan syfta på helt olika bakomliggande faktiska frekvenser beroende på om det är t.ex. bio- eller teaterbesök som efterfrågas.

- På grund av dessa variationer (se ovan) kan det vara problematiskt att jämföra frågor och även grupper av respondenter med varandra.
- Vaga kvantifieringar kan uppfattas som en subjektiv och osäker mätning men ibland är det den bästa lösningen. Vissa saker kan inte besvaras med faktiska frekvenser.

Antal skalsteg:

- Många svarsalternativ möjliggör mer detaljerade analyser.
- Om svarsalternativen är så många att de tappar sin mening påverkas mätningen negativt.
- För få skalsteg kan också påverka mätningen negativt – att respondenten tvingas svara något som inte riktigt stämmer.
- Om skalan ska ha etiketter vid skalstegen begränsar språket hur många skalsteg skalan kan ha.
- Antalet skalsteg beror också på om skalan är uni- eller bipolär. Vid bipolära skalor kan samma etikett användas för två skalsteg, ett på varje sida av mittpunkten.
- Rekommendationerna varierar men en tumregel är fem–nio svarsalternativ i postenkäter och färre i telefonintervjuer.

Riktningen på svarsskalan:

- Riktningen på svarsskalan kan ha betydelse. Det kan finnas en tendens (p.g.a. både kognitiva och motivationsmässiga skäl) att välja ett tidigt snarare än sent alternativ.
- Det kan vara svårt att designa bort denna tendens.

Instämmer-skalor:

- Instämmer-skalor (baserade på instämmer eller stämmer) är vanliga svarsskalor till frågor i påståendeform.
- Skalan är förknippad med hålla-med-tendenser ("acquiescence bias").
- Svarsprocessen är mer komplicerad än frågesatser – respondenten måste inte bara ta ställning till innehållet i sig utan även anpassa sitt svar till påståendet och svarsskalan.

Uni- och bipolära skalor:

- Unipolär betyder att skalan endast har en pol – den utgår från lite av något till mycket av samma sak.

- Bipolära skalor innehåller två poler och går från en pol till dess motsatta pol.
- Unipolära skalor används mest för att mäta frekvenser, bipolära för att mäta åsikter.
- Det är främst vid bipolära skalor som en mittpunkt är aktuell, placerad mittemellan de två polerna.
- Bipolära skalor har traditionellt sett fler skalsteg än unipolära.

Mittenalternativ eller inte:

- Mittenalternativ är främst relevant vid bipolära skalor.
- Om man inte har ett mittenalternativ tvingar man dem som är neutrala att ta ställning (mätfel) och det kan också riskera partiellt bortfall.
- Om man har mittenalternativ finns det risk att flera respondenter kommer att markera det p.g.a. bristande motivation, inte för att det är sant (mätfel).
- "Flykten till mittpunkten" kan begränsas genom att ha fler skalpunkter och erbjuda svarsalternativ för icke-svar (till exempel "Vet inte").
- Mittpunkten bör ha en etikett som beskriver dess position mittemellan punkterna, så att den inte tolkas som något svarsalternativ för allmän osäkerhet (till exempel "Osäker" eller "Vet inte riktigt").
- Forskningsresultaten varierar en del när det gäller mittpunkten. Rekommendationen är ändå att generellt använda en mittpunkt men att vara medveten om hur det kan påverka resultatet.

Alternativ för icke-svar (till exempel "Vet ej"):

- Om icke-svar saknas som svarsalternativ tvingas respondenter som frågan inte är relevant för att ändå ta ställning (till exempel för att de inte har en åsikt).
- Om icke-svar finns som svarsalternativ kan, å andra sidan, respondenter välja det p.g.a. låg motivation.
- Forskningen varierar huruvida datakvaliteten blir bättre eller sämre med alternativ för icke-svar.
- I intervjuer (och i viss mån i webbenkäter) kan icke-svar erbjudas "vid behov".

- Det finns också ett respondentperspektiv, att respondenterna inte ska känna sig tvingade till svar som inte stämmer.
- Vid beslutet om icke-svar ska ingå bör man också utgå från frågans innehåll. Hur relevant är frågan för respondenterna? Hur motiverade är de?
- Om icke-svar finns kan man i alla fall vara (mer) säker på att de som ligger på svarsskalan faktiskt hör hemma där.

Beskrivande etiketter till skalpunkterna:

- Etiketter betyder att skalstegen definieras med ord.
- Om skalan har många skalsteg kan det bli svårt att formulera etiketter som tydligt skiljer sig åt.
- En idé är att skalan blir mer av en intervallskala om skalan har siffror än om den har etiketter. Men det är inte givet hur detta fungerar eftersom respondenterna ändå borde behöva omvandla siffrorna till den underliggande svarsdimensionen för att kunna svara.
- Forskningen antyder att etiketter förbättrar mätningen eftersom det hjälper respondenten att förstå skalstegens innebörd.

Slutna svarsalternativ utan given inbördes ordning:

- I nominalskalor – svarsalternativ som kan särskiljas men inte rangordnas – bör det ändå finnas något slags ordning, speciellt om alternativen är många, så att respondenten kan orientera sig.

Ja-nej-svar:

- Ja-nej ger ett tydligt resultat men kan, beroende på frågans innehåll, innebära en informationsförlust och framstå som inskränkande av respondenterna.
- Om frågan ska mäta sådant som kan variera i mer än två punkter finns det, av flera skäl, ett värde att ha en mer differentierad skala.
- I övrigt är ja-nej väldigt tydliga svarsalternativ.

Rankingskalor:

- Rankingskalor kräver mycket av respondenten – att objekt ska bedömas i förhållande till varandra och att en ordning ska upprättas.

- Det finns även risker för att respondenterna kommer att hitta på egna lösningar.
- I telefonintervjuer fungerar rankingskalor dåligt. I webbenkäter finns det fler lösningar än i pappersenkäter.
- En rekommendation är att undvika rankingskalor, speciellt i pappersenkäter och telefonintervjuer. De bör endast användas om det finns särskilda skäl och i så fall förenklas så långt som möjligt.

Checklista:

- Checklistan har många fördelar men även nackdelar – vad betyder avsaknaden av markering?
- Checklistan ger färre markeringar än antalet "Ja" i ja-nej-frågor med samma alternativ.
- Vår erfarenhet på SCB är att respondenterna ibland besvarar listor av ja-nej-frågor som checklistor, det vill säga enbart markerar "Ja".

Svarsalternativ i intervjufrågor:

- En stor skillnad mot enkät är att vid intervjuer kan inte respondenten själv läsa svarsalternativen.
- Valet av svarsformat påverkar både respondentens svarsprocess och intervjuarens roll.
- Det finns 6 format som varierar från att respondenten får svara fritt (och intervjuaren återger svaret i klartext) till att alla svarsalternativen läses upp av intervjuaren.
- Om inte respondenten känner till svarsalternativen blir det, på sätt och vis, intervjuaren som väljer åt respondenten. Detta kan leda till fel.
- Mer reglerade format har dock sina egna svagheter (de kan vara inskränkande eller alternativen kan vara för många för att läsas upp).
- Vilket format som passar en fråga bäst är därför en bedömningsfråga.

Svarsalternativ och tolkningar:

- Perifer information, som inte är avsedd att vara meningsbärande, kan ändå påverka respondenterna.

- Ett exempel är skalans numrering – om den går från minus till plus snarare än från noll och uppåt.
- Förmodligen beror det på att skalan uppfattas som bipolär (där ändpunkterna är motsatser) om siffrorna går från minus till plus.
- Svarsskalan bör alltså vara tydlig och perifer information bör vara i linje med svarsskalan, inte signalera något motstridigt.
- Den psykologiska och faktiska mittpunkten är inte alltid förenliga och det kan behöva förtydligas i skalan (5 är inte mittpunkten på en skala som går från 1-10, även om det kan framstå som så).
- Respondenter kan ibland utgå från svarsalternativens position snarare än innebörd. Respondenterna kan alltså svara relativt snarare än faktiskt.
- Svarsalternativ kan förväntas återge en slags normalfördelning där det "normala" representeras av alternativet i mitten.
- Problemet kan åtgärdas genom ett förtydligande av frågan eller genom att använda ett öppet svarsalternativ.
- Om icke-svar ligger i direkt anslutning till ordinarie svarsalternativ kan de framstå som en del av svarsskalan och påverka vad som uppfattas som mittpunkten på skalan.
- Svarsalternativen kan bidra till frågans förståelse eftersom de kan förtydliga vilken intensitet som avses i frågan.
- "Arg" uppfattas syfta på olika starka känslor beroende på om svarsskalan utgår från dagar eller månader.
- I frågor som är breda ämnesmässigt (finns många tänkbara svar) kan det bli stora skillnader mellan öppna och stängda frågor. Det beror på att i öppna frågor svarar respondenten fritt. I stängda frågor avgränsar svarsalternativen respondenten.

4 Enkätmetodik

Författare: Lina Fjelkegård (SCB) och Andreas Persson (SCB)

Detta är främst en bok om frågekonstruktion men det finns en del relaterande områden som frågekonstruktören bör ta hänsyn till eller kan bli inblandad i. Formulärkonstruktion – när enskilda frågor ska sättas samman till en helhet som sedan ska förmedlas till respondenterna – är ett tydligt exempel på en sådan uppgift. Ofta blir frågekonstruktören även inblandad i, eller ansvarig för, annat undersökningsmaterial än just enkäten. Det kan till exempel vara skrivandet av förbrev, introduktioner och påminnelser. Kanske beror det på att frågekonstruktören ofta är intresserad av språk och van att formulera sig i skrift. Nedan beskrivs några tumregler och en del forskningsresultat uppdelat i två större avsnitt: 1) formulärkonstruktion och 2) annat utskicksmaterial. Målet är att förbättra mätningen men också att få så hög svarsfrekvens och så nöjda respondenter som möjligt.

Formulärkonstruktion

När de enskilda frågorna är färdigskrivna är nästa uppgift att sammanställa dem i det frågeformulär (intervju eller enkät) som respondenterna ska besvara. Detta är en uppgift med många faktorer att beakta. Hur ska innehållet struktureras?

Avsnittsindelning

Frågekonstruktören må ha formulerat frågorna var och en för sig men så presenteras de inte för respondenterna. För att underlätta ifyllandet av frågeformuläret bör den som sätter samman frågeformuläret sortera de enskilda frågorna i logiska grupper, besluta om frågegruppernas inbördes ordning och dela in dem i avsnitt (ESS, 2006).

Avsnittsindelning i enkäter

Frågorna ska inte komma huller om buller utan bör delas upp i avsnitt. Ett avsnitt bör innehålla minst två frågor. Oftast är en ämnesmässig uppdelning det allra bästa. Det innebär att respondenterna inte behöver tänka om hela tiden utan kan besvara ett helt avsnitt innan förutsättningarna ändras. Men det kan finnas situationer när en annan uppdelning är

mer lämplig, till exempel kronologisk (att respondenterna svarar för olika tidsperioder av sitt liv).

Varje avsnitt bör ha en avsnittsrubrik. Det händer ganska ofta att enkäter saknar det, men det är dumt att missa eftersom det hjälper respondenten att lättare ta till sig innehållet.

Läs exempelvis nedanstående text (Passer & Smith, 2007):

Först måste man sortera och avgöra vilka saker som hör ihop och ska hållas isär. Man lägger dem i olika högar. Om man blandar kan det bli fel i slutändan, till exempel att storleken blir för liten. Sedan är det bara att köra igång. Proceduren tar ungefär en timme beroende på vad man väljer.

Texten var ganska svår att förstå eller hur? Tänk om den hade haft rubriken "Att tvätta" – då skulle det ha varit lättare (Passer & Smith, 2007). Rubriker fyller nämligen funktionen av att introducera ett ramverk som innehållet kan förstås utifrån. En förvarning om vad som komma skall. Rubriker är alltså viktiga för förståelsen. Det förutsätter dock att rubriken är meningsbärande.

EXEMPEL:

Några frågor till dig

Ovanstående rubrik innehåller ingen information om frågornas innehåll och kan därför inte hjälpa respondenterna med att skapa ett ramverk. Den typen av rubriker bör formuleras om.

Det bör även tydliggöras i layouten när ett nytt avsnitt kommer så att respondenten uppmärksammar detta och kan "koppla om" efter de nya förutsättningarna. Detta kan göras genom att ha en rubrik som är synlig och på en ny sida. Rubrikerna ska dock inte ersätta frågetexten. Man ska kunna förstå varje fråga för sig, utan att behöva läsa den övergripande avsnittsrubriken.

Avsnittsindelning i intervjuer

Avsnittsindelning fyller samma viktiga funktion i intervjuer som i enkäter (se ovan). I intervjuer kan dock inte respondenten se när frågeformuläret byter ämne utan det måste intervjuaren tala om. Vanligt är att intervjuaren berättar det genom en kort introducerande text. Det är viktigt att sådana introduktioner är korta och koncisa men meningsbärande.

EXEMPEL:

Nu kommer några frågor om ditt boende ... (Dahmström, 2005)

Precis som för enkäter ska inte sådana introduktioner ersätta frågeinnehållet utan frågorna ska vara självständiga.

Sammanfattande punkter:

- Frågorna ska inte komma huller om buller utan bör delas upp i avsnitt. Ett avsnitt bör innehålla minst två frågor. Oftast är en ämnesmässig uppdelning det allra bästa.
- Varje avsnitt bör ha en avsnittsrubrik som förbereder respondenten på innehållet.
- I intervjuer kan dock inte respondenten se när frågeformuläret byter ämne utan det måste intervjuaren tala om. Vanligt är att intervjuaren berättar det genom en kort introducerande text.

Ordningsföljden mellan och inom avsnitt

Ett frågeformulär består av innehåll på olika nivåer som tillsammans ska utgöra en helhet. Det är viktigt att den helheten är så naturlig och förståelig som möjligt för respondenten – det påverkar både mätningen och respondentens upplevelse på ett positivt sätt. Den mest naturliga ordningsföljden är oftast att följa trattprincipen – att gå från generellt till specifikt. En annan aspekt är att frågeformuläret gärna ska inledas med frågor som respondenten tycker är intressanta eller roliga att besvara.

Det kan emellertid finnas en del andra aspekter att ta hänsyn till. Frågeformulären ska helst inledas med frågor som är lätta att besvara så att respondenterna känner att de kan hantera uppgiften. Detta gäller framförallt i enkäter där risken för tidiga avbrott förmodligen är betydligt större än i intervjuer.

Kan man uppfylla alla dessa rekommendationer är det förstås bra, annars får man bedöma vad som är bäst i varje enskilt fall.

Det är väldigt vanligt att frågeformulär inleds med ett antal demografiska frågor om till exempel kön, inkomst, ålder och utbildning. Dessa är förmodligen lätta att besvara men sannolikt inte särskilt intressanta. Varför är de då så vanliga att inleda med? I intervjuer kan det vara ett sätt att kontrollera att man pratar med rätt person. I enkäter är

syftet inte lika självklart. Varför inte inleda med mer ämnesmässiga frågor? Ett problem är dock var de demografiska frågorna i så fall ska placeras. Det passar inte bra att ha dem mitt i enkäten eftersom det bryter flytet. Demografiska frågor måste nog antingen komma först eller sist i enkäten. En risk med att ha dem sist är att de då inte blir lika korrekt ifyllda eftersom respondenterna ofta är mindre koncentrerade i slutet än i början. Detta kan få tråkiga konsekvenser eftersom de demografiska frågorna ofta utgör viktiga sorteringsvariabler i analyskedet. Det vanliga är därför att inleda med dem.

Känsliga frågor bör inte ligga först eftersom de kan ha en avskräckande effekt på hela frågeformuläret. De kan ge intrycket av att hela frågeformuläret kommer att vara känsligt, trots att det kanske bara rör sig om ett par frågor. Om de känsliga frågorna kommer längre in i frågeformuläret kan också respondenten ha fått ett sammanhang genom tidigare frågor, och en förståelse för varför frågorna är viktiga att ställa. Det finns alltså flera skäl till att ha känsliga frågor längre in i frågeformuläret (läs mer i kapitel 2).

Om frågeformuläret innehåller frågor som kan antas vara beroende av sammanhanget (att svaren på frågorna kan påverkas av placeringen i frågeformuläret) bör man behålla samma ordning som i tidigare frågeformulär eller som i andra undersökningar man vill jämföra med. Då håller man i alla fall eventuella fel konstanta och minimerar risken för att tolka ändringar på grund av skifte av sammanhang som faktiska förändringar.

Sammanfattande punkter:

- Den mest naturliga ordningsföljden är oftast att följa trattprincipen – att gå från generellt till specifikt.
- Frågeformulären ska helst inledas med frågor som är intressanta och lätta att besvara så att respondenterna känner att de kan hantera uppgiften.
- Känsliga frågor bör inte ligga först eftersom de kan ha en avskräckande effekt på hela frågeformuläret.

- Om frågeformuläret innehåller frågor som kan antas vara beroende av sammanhanget (att svaren på frågorna kan påverkas av placeringen i frågeformuläret) bör man behålla samma ordning som i tidigare frågeformulär eller som i andra undersökningar man vill jämföra med.

Layout

Om frågeformuläret är en pappers- eller webbenkät ska det också formges. Detta fyller framförallt två funktioner. Den första är att enkäten ska ha ett tilltalande intryck, den andra är att layouten ska underlätta ifyllandet av enkäten, till exempel genom att göra viktiga aspekter mer framträdande. I layoutarbetet ingår även att anpassa enkäten för eventuell skanning (läs mer i kapitel 5).

Navigering

Ett smidigt flöde genom en enkät är viktigt för att minimera bortfall och mätfel. Det gäller främst enkäter och i särskild grad pappersenkäter. I intervjuformulär och webbenkäter sköter inte respondenten navigeringen på egen hand i samma grad utan mycket av flödet är redan programmerat i webbenkäten eller intervjuarens datorstöd.

I pappersenkäter sköter däremot respondenten navigeringen själv. Då behövs en tydlig guidning, till exempel angående vad som ska besvaras, i vilken ordning och vad som eventuellt inte ska besvaras. Detta görs med hjälp av instruktioner och genom att dra respondentens uppmärksamhet till rätt plats i enkäten (läs mer i kapitel 5).

Filterfrågor

Filterfrågor är ett viktigt verktyg att styra respondenterna rätt med. De gör så att respondenten kan hoppa över frågor som inte är relevanta. Samtidigt måste respondenten ha information nog för att göra ett korrekt val, annars finns risken att frågor felaktigt hoppas över. Filter ska också vara enkelt konstruerade för att inte förvirra respondenten. Generellt rekommenderas att det svarsalternativ som tar respondenten till nästa fråga ska stå först, och att de alternativ som medför hopp ska komma därefter (läs mer i kapitel 2).

Filterfrågor innebär dock ett störningsmoment och bör därför användas med viss försiktighet.

Instruktioner

Det händer att ord, termer och begrepp missförstås av respondenterna. Ett sätt att undvika detta är att erbjuda ytterligare förklaringar och instruktioner. Instruktioner är alltså tilläggsinformation till frågeformuleringen. De kan utgöras av exempelvis definitioner eller regler för hur frågan ska besvaras. En instruktion kan variera från ett kort exempel till en lång redogörelse för vad som innefattas och inte innefattas i ett begrepp.

I standardiserade intervjuer är sådana instruktioner ofta just standardiserade i syfte att alla respondenter ska få samma information. I friare och mer flexibla intervjuer kan intervjuarna förklara med egna ord (Fowler & Mangione, 1990). I självadministrerade enkäter är sådana instruktioner också standardiserade eftersom alla respondenter får samma information. I pappersenkäter krävs det dock en avvägning av hur mycket instruktioner man kan ha utan att röra till layouten eller öka informationskravet för mycket. Det kan också bli irriterande för de respondenter som inte behöver instruktionerna.

Man kan skilja på *nödvändiga* och *förtydligande* instruktioner. Nödvändiga instruktioner är instruktioner som respondenten verkligen bör läsa (eller höra) för att kunna besvara frågan korrekt. Förtydligande instruktioner är mer av information "vid behov", till exempel om vissa respondenter skulle stöta på något speciellt problem. Nödvändiga instruktioner är naturligtvis viktigare än förtydligande. Om instruktioner verkligen har en avgörande betydelse bör de ligga så nära frågan som möjligt. Det bästa är om det går att införliva dem i frågeformuleringen. Detta kan vara lämpligt om informationen i instruktionen har stor påverkan, till exempel hur ett ord eller en term ska förstås. Annars kan man placera dem direkt under frågan som en separat text. Det är dock en svagare placering.

Förtydligande instruktioner kan också placeras efter frågetexten, eller om de gäller flera frågor, i avsnittets inledning.

I intervjuer måste det vara tydligt för intervjuaren vilka instruktioner som är nödvändiga och alltid ska läsas och vilka som endast ska läsas vid behov, till exempel om respondenten tvekar eller missförstår. Detta kan tydliggöras för intervjuarna i datorstödet, intervjuarinstruktionerna och vid utbildningen inför undersökningen.

Instruktioner i webbenkäter

I webbenkäter finns tekniska möjligheter att ge instruktioner vid behov. Exempelvis kan instruktioner finnas som klickbar länk eller bara vara synlig om respondenten markerar den med muspekaren. Ett problem är dock att respondenterna kanske inte anser att de behöver instruktioner. Ett annat problem är att även om respondenterna inser det kanske de väljer att låta bli att aktivera instruktionen av motivationsskäl.

När instruktionen är ständigt synlig i webbenkäter används den mer än om den endast är synlig när den markeras med datormusen. Det verkar alltså som att det extra steg som krävs för att markera instruktionen gör att färre läser den. De som markerar instruktionen läser den däremot mer noggrant, jämfört med om instruktionen är ständigt synlig. Hur ska man förhålla sig till detta resultat? Ett tips kan vara att ha instruktionen ständigt synlig, till exempel placerad mellan fråga och svarsalternativ, men att göra den framtung i den betydelsen att det viktigaste kommer först (Peytchev, Conrad, Couper & Tourangeau, 2010).

En studie (Conrad, Couper, Tourangeau & Peytchev, 2006) undersökte instruktioner i en webbundersökning. Det visade sig att instruktioner användes i liten grad generellt sett, även om det bara krävdes ett musklick för att aktivera dem. Användningen var ännu lägre om det krävdes två klick eller ett klick och sedan scrollning. Det var vanligare att aktivera instruktioner för tekniska än för vardagliga termer. Om instruktionen innehöll någon värdefull information (något som gick bortom eller kanske till och med motsade vad respondenten redan visste på förhand) ökade användningen av instruktionen fortsättningsvis i formuläret, jämfört med instruktioner som inte tillförde något nytt utöver vad respondenten redan kunde förväntas veta.

För att stimulera användning av instruktioner ska de därför vara så enkla att aktivera som möjligt. Allra enklast är om de finns vid frågetexten och inte behöver aktiveras, men det är inte alltid möjligt. Om instruktionen behöver aktiveras för att bli synlig är det viktigt att den tillför någon värdefull information för respondenten (Conrad, Couper, Tourangeau & Peytchev, 2006).

Separata instruktioner

Ibland kan instruktionerna vara så långa att det inte går att placera dem vid den aktuella frågan eller ens i enkäten. Enda möjligheten är då att ha instruktionen separat, till exempel som ett instruktionshäfte eller i ett

elektroniskt dokument. Detta kan ibland vara nödvändigt, som vid komplicerade mätningar med avancerade begrepp. Det bör dock undvikas så långt det är möjligt eftersom det finns en risk att respondenterna inte läser separata instruktioner i önskvärd utsträckning. Vid frivilliga undersökningar bör de överhuvudtaget inte användas (inte ens kortare instruktioner i följebrev). Vid undersökningar med uppgiftslämnarskyldighet kan de ibland vara nödvändiga. Men även där bör man arbeta med att underlätta instruktionerna så mycket som möjligt.

Sammanfattande punkter:

- Instruktioner är tilläggsinformation till frågeformuleringen. De kan utgöras av exempelvis definitioner eller regler för hur frågan ska besvaras.
- I pappersenkäter krävs det en avvägning av hur mycket instruktioner man kan ha utan att röra till layouten eller öka informationskravet för mycket.
- Man kan skilja på *nödvändiga* och *förtydligande* instruktioner. Nödvändiga instruktioner är instruktioner som respondenten verkligen bör läsa (eller höra) för att kunna besvara frågan korrekt. Förtydligande instruktioner är tänkta att användas mer "vid behov", till exempel om vissa respondenter skulle stöta på något speciellt problem.
- Om instruktioner verkligen har en avgörande betydelse bör de ligga så nära frågan som möjligt.
- I intervjuer måste det vara tydligt för intervjuaren vilka instruktioner som är nödvändiga och alltid ska läsas och vilka som endast ska läsas vid behov.
- I webbenkäter finns tekniska möjligheter att ge instruktioner vid behov. Exempelvis kan instruktioner finnas som klickbar länk eller bara vara synlig om respondenten markerar den med muspekaren.
- Det kan dock vara svårt att få respondenter att läsa instruktioner.
- För att stimulera användning av instruktioner ska de vara så enkla att aktivera som möjligt. Om instruktionen behöver aktiveras för att bli synlig är det viktigt att den tillför någon värdefull information för respondenten.

Längden på frågeformulär

Hur lång ska frågeformuläret (enkäten eller intervjun) vara? Det är ett ständigt dilemma. Ett långt frågeformulär innebär att mer information samlas in. Det är ganska naturligt att de flesta som genomför en undersökning vill passa på att samla in så mycket information som möjligt. Ofta kan det dessutom vara flera intressenter inblandade i frågeformuläret som alla vill ha med sina frågor. Här är dock det gamla ordspråket "Den som gapar efter mycket..." värt att ha i åtanke. Frågeformulärets längd kan påverka både mätningen och bortfall. Om det är för långt kan respondenterna tröttna och fylla i det mindre noggrant eller, i värsta fall, avbryta besvarandet helt. Oavsett mätningen, bör man för respondenternas skull inte heller ställa fler frågor än man behöver.

Enkätlängd

Bortfall

Det kan tyckas självklart att längden på enkäten borde påverka bortfallet i riktningen att ju längre enkäten är, desto högre blir bortfallet. Forskningen har dock haft svårt att bekräfta att det faktiskt förhåller sig så. För pappersenkäter brukar längd oftast mätas i antalet sidor, men det är inte alltid ett rättvisande mått. Förmodligen är *upplevelsen av längden* också viktig, vilket inte bara beror på längden i sig utan kan även påverkas av layouten och ämnesområdet. Det är alltså svårt att dra några generella slutsatser om längden och bortfall för det är många faktorer som spelar in. Större undersökningar har ofta längre frågeformulär men också mer resurser att lägga på förhandsinformation, belöningar och uppföljningar. Det kan kompensera en eventuell negativ effekt av enkätlängden (Fritz & Lundström, 2007). Dillman (2007) hävdar att det finns en negativ effekt av längd efter tolv sidor men på grund av ovanstående faktorer kan det i praktiken vara svårt att ange ett exakt sidantal.

Den allmänna erfarenheten är ändå att längden har betydelse för bortfall och även för datakvaliteten. Den betydelsen är extra stor i undersökningar som inte har så många andra attraktionsfaktorer, till exempel de som berör ett relativt sett smalt ämne. Man bör alltså undvika att fylla enkäten med frågor som "kan vara bra att ha" och inte fråga mer än nödvändigt.

I webbenkäter där respondenten klickar sig fram till nästa fråga eller avsnitt är längden okänd för respondenterna. Antingen får man upplysa

dem om det i introduktionen, vilket kan påverka deltagandet (Galesic & Bosnjak, 2009). Ett annat alternativ är att använda en så kallad "progressbar", som visar hur långt respondenten har kommit och hur mycket som återstår av enkäten (en "progressbar" passar dock illa för enkäter som varierar mycket i längd beroende hur respondenterna svarar på filterfrågor etc.). Om man frågar respondenterna så vill de ofta ha en sådan, men om man erbjuder det som ett klickbart alternativ är det få som använder det. Forskningen om progressbars är i allmänhet otydlig och effekten varierar i olika studien (Couper, 2008) (läs mer i kapitel 5). En faktor som kan ha betydelse är respondenternas förväntan på enkätlängd och i vilken grad enkäten är i linje med eller bryter mot dessa förväntningar. En progressbar kan alltså både hjälpa och stjälpa. Om en sådan ska användas måste därför bedömas från fall till fall (Yan, Conrad, Tourangeau & Couper, 2010). Dock kan en *för* exakt visning av antalet återstående frågor vara störande för besvarandet. En rekommendation är därför att ha med en progressbar men som på ett mer övergripande och mindre exakt sätt redovisar var respondenten är i frågeformuläret (Couper, 2008).

Mätning

Vad gäller mätning har en del studier undersökt men inte hittat några negativa svarstendenser vad gäller enkätlängd. Andra har hittat en ökad markering av mittenalternativet och färre svar i ändpunkterna på skalor i enkätens senare delar (för en genomgång av äldre forskning, läs Herzog & Bachman, 1981). En studie fann indikationer på att enkätlängden kan orsaka så kallad "straight-lining" – att respondenten lämnar samma svarsalternativ på många frågor i ett gemensamt tabellformat. Resultatet förklaras med respondenternas sjunkande motivation. Det är svårt att dra generella slutsatser om längdens betydelse för datakvaliteten. Det finns många faktorer som påverkar, till exempel ämnesområdet, enkätens formatering och layout samt frågornas formuleringar och krav (Herzog & Bachman, 1981). Men precis som för bortfallet är nog mångas erfarenhet att längden kan påverka mätningen negativt.

Intervjulängd

Vid en telefonintervju var en tidigare rekommendation att den inte ska pågå längre än cirka 30 minuter (Bergman & Wärneryd, 1982). Idag är rekommendationen att de bör vara ännu kortare. Det kan vara svårt att behålla respondenternas koncentration och intresse längre än så via ett så

pass distanserat medium som telefonen. Idag är det dessutom inte säkert att respondenterna befinner sig i hemmet när intervjuaren ringer utan de kan vara i miljöer som är sämre lämpade för intervju. En studie hittade tecken på uttrötningseffekter och att "Vet inte"-alternativet valdes oftare beroende på hur sent i intervjun frågan ställdes. De kombinerade effekterna var, i kronologisk ordning, 16 procent, 17,9 procent, 28,2 procent och 19,1 procent. Att den sista placeringen är lägre än den näst sista (28,2 jämfört med 19,1 %) är förvånande. Det kan bero på att respondenten fått indikationer på att slutet på intervjun nalkades och därför skärpte sig (Smith, 1992). Vid telefonintervjuer finns det också en risk för störningsmoment i respondentens närmiljö, till exempel tv:n och familjemedlemmar. Denna risk ökar om respondentens intresse för undersökningen avtar.

Vid besöksintervjuer finns bättre möjligheter till längre intervjuer. Ofta utförs den i lugn och ro och innebär ett större åtagande från respondenten. Det finns därför förmodligen andra förväntningar vad gäller tidsåtgång. Rekommendationen för besöksintervjuer är cirka 60 minuter (Bergman & Wärneryd, 1982).

Det går att gå över de rekommenderade tiderna men även om den är okänd och varierar för olika personer finns det likväl en tidpunkt när respondentens intresse och koncentration börjar tryta så pass mycket att det kan leda till bortfall eller mätfel.

Sammanfattande punkter:

- Det kan tyckas självklart att längden på enkäten borde påverka bortfallet i riktningen att ju längre enkäten är, desto högre blir bortfallet. Forskningen har dock haft svårt att bekräfta att det faktiskt förhåller sig så.
- Förmodligen är upplevelsen av längden också viktig, vilket inte alltid beror på längden utan kan även påverkas av layouten och ämnesområdet.
- Den allmänna erfarenheten är ändå att längden har betydelse för bortfall och även för datakvaliteten. Den betydelsen är extra stor i undersökningar som inte har så många andra attraktionsfaktorer.

- Vid en telefonintervju var en tidigare rekommendation att den inte ska pågå längre än cirka 30 minuter (Bergman & Wärneryd, 1982). Idag är rekommendationen att de bör vara ännu kortare.
- Vid besöksintervjuer finns bättre möjligheter till längre intervjuer.

Börda i panelundersökningar

En faktor som kan påverka svarsbenägenheten är alltså den börda som uppgiften innebär. Bördan kan bestå av den tid det tar att besvara frågeformuläret här och nu – själva formulärlängden. I panelundersökningar, där respondenterna deltar vid flera insamlingstillfällen, får de dessutom bördan av framtida insamlingar att ta ställning till. I engångsundersökningar finns en omedelbar börda men i panelundersökningar finns också en longitudinell börda. Den kan ses som att respondenten förbinder sig att delta igen i framtiden. Blir då bortfallet högre vid rekryteringen till en panel jämfört med till en enskild undersökning? Det beror mycket på hur förfrågan formuleras och vilken undersökning det handlar om. Enligt en studie om deltagandet i en intervjuundersökning av paneltyp ledde tillägget (kursiverat) "Det är en entimmesintervju nu och *mer i framtiden*" till fem procents lägre svarsfrekvens (Apodaca, Lea & Edwards, 1998).

Annat utskicksmaterial

Nedanstående avsnitt handlar om annat undersökningsmaterial än själva frågeformuläret, till exempel förhandsinformation, följbrev, påminnelser och belöningar för deltagandet. Sådant material fyller viktiga funktioner för undersökningen. Det kan påverka mätningen, respondenternas upplevelse av undersökningen och motivationen att delta. I avsnittet nedan diskuteras i huvudsak vilken betydelse materialet kan ha på svarsfrekvensen.

Förhandsinformation

Sätt att öka intresset och höja svarsfrekvensen

Ett sätt att öka svarsfrekvensen är att skicka ut information till respondenterna i ett separat utskick innan själva insamlingen börjar. Någon vecka innan enkäten anländer eller en intervjuare ringer får respondenten ett brev som berättar om undersökningen och att

respondenten snart kommer få ett utskick med en enkät alternativt kommer bli uppringd. För enkäter brukar sådan information kallas för förbrev ("pre-notification letter" på engelska). Förhandsinformationen inför en intervjuundersökning kallas för introduktionsbrev ("advance letter" på engelska).

Förhandsinformationens syfte

Syftet med förhandsinformationen är (Luppés, 1995)

- att annonsera intervjun eller enkäten
- att ge information om undersökningen, vad den handlar om och eventuellt något om hur resultaten kommer att användas (för introduktionsbrev till intervju ges vanligtvis även den juridiska information som respondenten har rätt till)
- att motivera respondenterna att delta i undersökningen, väcka nyfikenhet och fungera som en "aptitretare".

Att skicka ut förhandsinformation kan ju vara en trevlig gest, men har det någon betydelse för svarsfrekvensen?

Förbrev - enkät - svarsfrekvens

I en sammanställning av flera studier om olika faktorer som påverkar svarsfrekvenser framkommer att förbrev kan ha en viss positiv effekt (Dillman, 2000). Dillman skriver att generellt sett är fler kontakter bättre för svarsfrekvensen än färre. Det finns också exempel där förbrevet har en unik effekt utöver att endast vara ytterligare en kontakt. I en studie ökade förbrevet svarsfrekvensen med fyra-sex procent. Ökningen kunde inte kompenseras av frankerade svarskuvert, påminnelsekort eller utskick med ytterligare en enkät (Dillman, Clark & Sinclair, 1995). SCB har också gjort några studier för att testa effekten av förbrev med mestadels positiva resultat. Förbrev verkar kunna ha en positiv effekt på deltagandet. Kanske för att det uppfattas som en positiv kontakt (i jämförelse med påminnelser).

Inför undersökningar med webbinsamling bör man dock vara något försiktig med att använda förbrev. Beroende på hur brevet formuleras kan det uppfattas som märkligt av mottagaren att det inte innehåller några inloggningsuppgifter eller någon webbadress. Om brevet innehåller inloggningsuppgifter bör det hanteras som ett missiv (följebrev) och inte som ett förbrev, se nedan.

Introduktionsbrev – telefonintervju

Det finns betydligt mer forskning om effekten av introduktionsbrev inför telefonintervjuer än vad det finns för förbrevens effekt på pappers- och webbenkäter. Att det fokuserats så mycket på telefonundersökningar är inte så konstigt. Internationellt sett har telefonintervjuer länge varit den dominerande insamlingsmetoden vid insamling från individer och hushåll. De senaste åren har telefonintervjuer stött på problem med stora bortfall vilket motiverat forskning med syfte att bryta den negativa trenden.

Forskning har visat att introduktionsbrev har en positiv effekt på svarsfrekvensen (de Leeuw, Callegaro, Hox & Korendijk, 2007). Intervjuarna själva uppskattar dessutom introduktionsbrevet eftersom det medför att de slipper handskas med förvåning och skepsis hos respondenterna i samma grad. Introduktionsbrev kan dock ha en negativ effekt genom att vissa respondenter bestämmer sig för att inte delta på förhand och utan att intervjuaren då kan förklara syftet med undersökningen (Collins, Sykes, Wilson & Blackshaw, 1988; de Leeuw, Callegaro, Hox & Korendijk, 2007).

I en genomgång av ett flertal studier fann de Leeuw, Callegaro, Hox och Korendijk (2007) att introduktionsbrev ökar svarsfrekvensen med åtta procentenheter. Effekten gällde oavsett population, ämne eller avsändare. Effekten verkade inte heller påverkas av introduktionsbrevets längd. Andra studier har dock funnit att postkort också kan ha en positiv effekt jämfört med inget utskick alls (Iredell, Shaw, Howatt, James & Granich 2004; de Leeuw, Callegaro, Hox & Korendijk, 2007). Dillman (2007) menar dock att det krävs ett visst antal sekunder för människor att koda in information i minnet och att ett postkort går alltför fort att kasta eller lägga undan.

Förhandsinformationens innehåll

För att få en bild av vad som vanligtvis ingår i introduktionsbreven gjorde Luppés (1995) en innehållsanalys av introduktionsbrev från sju länder (däribland brev från SCB). De flesta innehöll namnet på undersökningen, namnet på organisationen som genomförde undersökningen, fördelarna med att delta och information om anonymitet och sekretess. Studien kopplade dock inte innehållet till svarsfrekvens.

I metaanalysen av introduktionsbrevens effekter (se ovan) studerades även betydelsen av olika typer av innehåll i introduktionsbrev (de Leeuw, Callegaro, Hox & Korendijk, 2007). Eftersom innehållet var likartat i de flesta brev var det svårt att undersöka skillnader, men det framkom att det hade en positiv effekt att utlova feedback.

Dillman (2007) rekommenderar att förbrevet ska vara personligt, positivt, kortfattat, förklara att undersökningen är viktig, bygga upp en förväntan och skickas ut relativt tätt in på insamlingen.

Skillnader mellan datainsamlingsmetoder

Det finns alltså betydligt mer forskning om introduktionsbrev inför intervjuundersökningar än vad det finns om förbrev till pappers- och webbenkäter. Därför kan det vara intressant att resonera lite kring på vilka sätt enkäter och intervjuer skiljer sig åt när det gäller förhandsinformation (förbrev och introduktionsbrev).

Intervju – en annan typ av kontakt än enkät

En intervjukontakt skiljer sig på många sätt från en enkätkontakt. En aspekt är det personliga intrånget. Det var förr vanligare att få post från främlingar än att de ringde hem men så är det kanske inte längre. Att prata med en främling jämfört med att läsa ett brev från någon man inte känner kan dock upplevas som en starkare situation. Detta kan medföra att potentiella respondenter inledningsvis är mer besvärade och skeptiska i en intervjusituation än de är vid en enkät. Möjligen kan förhandsinformation därför ha en annorlunda effekt på telefon än i enkät.

En annan skillnad är möjligheten att presentera undersökningen. I en telefonintervju är tiden kort och det finns inte så mycket tid att presentera undersökningen. Är urvalspersonen skeptisk minskar det utrymmet ännu mer. I ett introduktionsbrev kan urvalspersonen få ta del av information om undersökningen i sin egen takt och utan oro för att tvingas gå med på något. Intervjuaren kan också hänvisa till introduktionsbrevet för att förtydliga syftet med påringningen. Det är också möjligt att introduktionsbrevet ökar chanserna för att den potentiella respondenten över huvud taget svarar i telefonen när den som genomför undersökningen ringer.

Det finns alltså flera skäl till att förhandsinformation fyller fler och kanske viktigare funktioner för intervjuer än för enkäter. Forskningen bekräftar också detta genom att visa på större effekter för intervjuer.

Inför intervjuundersökningar som genomförs av SCB skickas det alltid ut ett introduktionsbrev. Utöver en kort beskrivning av undersökningen och om vad som ska hända inkluderas obligatorisk information som respondenten har rätt att få

Reaktioner på förebrev – kvalitativa intervjuer med bortfallspersoner

Insynen i hur urvalspersonerna reagerar på förhandsinformation är begränsad. I intervjuundersökningar kan intervjuarna berätta om sina erfarenheter av förhandsinformationens funktion, men hur upplever respondenterna det? Svarefrekvensen är ett väldigt grovt mått på detta. För att få djupare insikter krävs kvalitativa studier.

I en sådan kvalitativ studie med vägrare kunde flera intervjuade bortfallspersoner knappt minnas att de hade fått ett förebrev. Formatet på förbrevet verkar ha haft betydelse eftersom det är viktigt att det inte uppfattas som reklam. Exempelvis menade de att ett postkort som inte låg i kuvert skulle avfärdas direkt utan att det lästes. Många lade inte heller märke till vem som var avsändare till förbrevet. De intervjuade bortfallspersonerna kom inte heller ihåg vad som stod i brevet eller vad undersökningen handlade om (Sztabinski, Dyjas-Pokoska & Zmijewska-Jedrzejczyk, 2006).

Författarna till den kvalitativa studien definierar tre grupper av reaktioner på brevet, där respondenten kunde

1. helt ignorera brevet efter att ha läst det
2. känna oro och misstro, till exempel rädsla för att bli lurade eller att det är märkligt att någon vet deras namn och adress samt irritation över alla som ringer och vill veta eller sälja saker. Den här reaktionen hade främst gamla, låginkomsttagare och lågutbildade
3. känna motvilja. Författarna menar dock att det var svårt att få fram den egentliga orsaken till den starka motviljan.

Man ska dock komma ihåg att alla dessa negativa reaktioner på brevet kom från personer som valt att inte delta. För andra kan det ha haft en positiv betydelse och tvärtom fått dem att delta i undersökningen (Sztabinski, Dyjas-Pokoska & Zmijewska-Jedrzejczyk, 2006).

Sammanfattande punkter:

- Ett sätt att öka svarsfrekvensen är att skicka ut information till respondenterna innan insamlingen inleds – innan enkäten anländer eller intervjuaren ringer.
- Syfte: att annonsera intervjun eller enkäten, att ge information om undersökningen och att motivera respondenterna att delta i undersökningen, fungera som en "aptitretare".
- I en studie ökade förbrevet svarsfrekvensen med fyra–sex procent.
- En sammanställning av flera studier fann att introduktionsbrev ökar svarsfrekvensen med åtta procentenheter.
- De flesta introduktionsbrev innehåller namnet på undersökningen, namnet på organisationen som genomförde undersökningen, fördelarna med att delta och information om anonymitet och sekretess.
- Dillman rekommenderar att förbrevet ska vara personligt, positivt, kortfattat, förklara att undersökningen är viktig, bygga upp en förväntan och skickas ut relativt tätt in på insamlingen.

Missiv (följebrev)

Missiv kallas det brev som medföljer en enkät. Benämningen kan också användas för brev med inloggningsuppgifter till en webbenkät. Detta skiljer sig alltså från förbrevet som skickas ut innan själva datainsamlingen har påbörjats. Missivet är ofta längre och utförligare än förbrevet som är mer av en aptitretare (se ovan). Missiv är endast relevant vid undersökningar som innehåller en självadministrerad del, det vill säga i huvudsak enkätundersökningar. I intervjuer motsvaras missivet av en muntlig introduktion från intervjuaren. Detta är bra att tänka på när man konstruerar sitt introduktionsbrev inför en intervju (se ovan).

Det främsta syftet med missivet är att ge urvalspersonerna all information som behövs för att veta vad som gäller och för att kunna besvara enkäten. Ett annat syfte är att motivera deltagande i undersökningen.

Innehåll

Viss information bör alltid finnas med i brevet:

- information om vad undersökningen syftar till och hur de insamlade svaren ska användas
- information om frivillighet eller eventuell uppgiftslämnarskyldighet
- formell juridisk information som respondenten har rätt till, exempelvis sekretess, hantering av personuppgifter och eventuella utlämnanden
- praktisk och teknisk information om besvarandet, exempelvis inloggningsuppgifter, länk och svarskuvert.

Den här informationen bör alltid vara med och det är viktigt att tänka på hur den formuleras både för att det ska vara lätt att förstå och för att brevet ska bli trevligt och inbjudande. Det ökar chanserna för att mottagaren läser vidare och förhoppningsvis också deltar i undersökningen. Brevet bör inledas på ett sätt så att läsaren snabbt förstår vad det handlar om och varför man har fått brevet.

För att ytterligare höja motivationen kan det vara bra att utveckla beskrivningen av syftet och då på ett sätt så att läsaren förstår att undersökningen och det egna deltagandet är viktigt. Vilken kunskapslucka ska undersökningen fylla och varför behövs den kunskapen? Tänk dock på att inte beskriva undersökningen på ett sätt som kan påverka respondentens svar.

Vår erfarenhet på SCB är också att många respondenter uppskattar om det i missivet framgår om och i så fall var det går att ta del av den statistik som tas fram genom undersökningen.

Vilket språk man använder och vilken layout man vill ha i missivet styrs av vilket intryck man vill ge. Vill man framstå som häftig och modern eller formell och seriös? Missivets layout bör gå i linje med detta. Det viktigaste är att anpassa innehåll och format utifrån brevets mottagare.

Det finns tyvärr inte så mycket forskning som visar att bilder skulle ha någon större betydelse för svarsfrekvensen, även om det kan se trevligt ut. Tänk dock på att även bilder som ses som neutrala kan ha inverkan på respondenten. De ska därför väljas ut med omsorg.

Sammanfattningsvis bör missivet

- vara inbjudande och trevligt
- vara enkelt och lättillgängligt
- vara informativt
- ha den viktigaste informationen i början
- vara mottagaranpassat.

Muntliga introduktioner och övertalning

Att få personer att vilja delta i en intervju

Att få personer att delta i en intervjuundersökning är ibland en muntlig förhandlings- eller övertalningssituation. Storleken på bortfallet varierar mellan olika intervjuare, vilket gör det meningsfullt att undersöka intervjuares olika strategier och tekniker i kontakten med urvalspersoner. Morton-Williams (1993) analyserade interaktionen mellan intervjuare och urvalsperson vid besöksintervjuer och fann att intervjuare som gick utanför manus var mer framgångsrika i att övertala än de som strikt följde manus. Sociala färdigheter är också viktiga, till exempel att kunna beskriva undersökningen på ett intressant sätt för just den person man pratar med. I arbetet med att övertala vägrare bör man naturligtvis ha mätningen och urvalspersoners perspektiv på att bli övertalad i åtanke (Burton, Laurie & Lynn, 2006; Couper, 1997).

Sammanfattande punkter:

- Missiv kallas det brev som medföljer enkäten. Missivet ska ge nödvändig information och motivera till deltagande.
- Respondenterna har rätt att få viss information om undersökningen innan de deltar. Sådan information måste alltid finnas med i missivet.
- För att motivera urvalspersonen att delta i undersökningen bör missivet på ett bra sätt förklara varför undersökningen och deltagandet är viktigt.
- Missivet ska ha en trevlig och lättillgänglig ton.

Påminnelser

Påminnelser är nästintill en nödvändighet i enkätundersökningar. En vanligt förekommande procedur är att ha åtminstone två skriftliga påminnelser. Ofta är dessa förkortade versioner av missivet men med förklaringen att det är en påminnelse det handlar om och ingen ny undersökning. Ibland skickas en ny enkät med i utskicket, ifall urvalspersonen inte skulle ha kvar den förra.

Ibland kombineras tack-för-deltagande och påminnelse. Detta kan vara en bra lösning om undersökningen är anonym så att man inte kan skilja dem som har svarat ifrån dem som inte har svarat. Även i andra fall kan lösningen vara lämplig, till exempel om man inte har hunnit registrera alla inkomna svar innan påminnelsen går ut.

Påminnelser leder generellt till fler svar även om effekten sällan är lika hög som man kanske hade hoppats på. Det gäller speciellt effekten av senare påminnelser.

Påminnelser kan vara irriterande för urvalspersonerna, speciellt för dem som redan beslutat sig för att inte delta. Det finns därför både kostnadseffektiva och etiska skäl till att begränsa antalet påminnelser.

Belöningar

Allmänt

Att ge belöningar för att delta i undersökningar är utbrett i Nordamerika och förekommer även på många andra platser. Oftast handlar det om en summa pengar men det finns också andra typer av belöningar. Forskning på området har till och med utvärderat effekten av att skicka med en chokladbit i första utskicket. Det gav en effekt i inledningen av datainsamlingen, men i slutet var den effekten borta (Brennan & Charbonneau, 2009).

Belöningen kan ha olika funktioner – från en symbolisk gest till en mer regelrätt ersättning för deltagande. Det sistnämnda är ganska ovanligt i urvalsundersökningar men förekommer till exempel i webbpaneler där respondenterna får poäng för varje undersökning de deltar i. Poängen kan de sedan handla för i en webbshop.

Belöningar kan även delas ut vid olika faser av datainsamlingen. De kan till exempel delas ut före eller efter medverkan, endast vid vägran och

dessutom variera i storlek. Vad säger då forskningen om belöningars effekter på svarsfrekvensen?

Det finns många äldre empiriska studier om belöningar. En metaanalys från 1975 (Armstrong, 1975) visade att belöningar som betalades ut på förhand minskade bortfallet rejält (ungefär en tredjedel). Mer moderna sammanställningar visar att (Church, 1993; Singer & Bossarte, 2006):

- pengar är mer effektivt än andra ersättningsformer
- större belöningar ger större effekt på svarsfrekvensen (inom vissa gränser)
- effekten är störst på undersökningar som har en låg svarsfrekvens i grunden
- förhandsbelöningar som bifogas med första utskicket är mer effektiva än utlovad ersättning vid deltagande
- effekten är störst för undersökningar som saknar andra morötter till deltagande. Belöningar kan till exempel kompensera en undersöknings nackdelar (som till exempel lång intervjuetid, Groves, Singer & Corning & Bowers, 1999)
- vissa studier vist att den effekt av belöningar som framkommer försvinner om man gör upprepade uppföljningar på bortfallet
- effekten av belöning inte verkar bero på att intervjuarna agerar annorlunda när de vet att urvalspersonen fått en belöning (Singer, Van Hoewyk & Maher, 2000).

Orsaker till belöningars effekt

Varför fungerar belöningar? Varför fungerar förhandsbelöningar bättre än utlovade belöningar? Ett skäl kan vara att det är mycket starkare och mer konkret att ha något i handen än att bli lovad något. Vid löften måste individen dessutom ha förtroende för att motparten kommer att hålla sitt ord och utbetala belöningen. Det kan säkert finnas en del misstro kring detta.

En förhandsbelöning kan också få respondenterna att uppmärksamma undersökningen och signalera att den är viktig och verkligen värdesätter deras deltagande. En annan möjlig skillnad mellan förhands- och utlovade belöningar är att en förhandsbelöning inleder ett utbyte. Ena parten (i det här fallet undersökningsorganisationen) har startat och gett respondenten något. Respondenten kan då känna sig manad att återgälda detta (denna regel om ömsesidigt byte av tjänster kallas för "reciprocitetsnormen") genom en motprestation – att besvara enkäten.

Vid en utlovad belöning har inget utbyte inletts utan beslutet att delta eller inte tas på andra grunder.

Rättvisa och förväntningar

Ska belöningar ges till alla eller ska de erbjudas tidigast vid första eller andra påminnelsen? Det finns kostnadseffektiva skäl till att erbjuda det tidigast vid påminnelser men det kan förstås bli orättvist. I en fokusgruppsundersökning som diskuterade sådana frågor tyckte inte deltagarna att det var orättvist. Däremot tyckte de inte att samhällsnyttiga undersökningar skulle ha belöningar. Om någon tjänar pengar på undersökningsresultatet är det dock skäligt att betala, ansåg de (Groves, Singer, Corning & Bowers, 1999).

Sammanfattningsvis bör den som funderar på att använda belöningar i sin undersökning väga de olika för- och nackdelarna mot varandra och ta ett beslut utifrån både den enskilda undersökningens förutsättningar och det allmänna undersökningsklimatet.

Sammanfattande punkter:

- Påminnelser leder generellt till fler svar även om effekten sällan är lika hög som man kanske hade hoppats på. Det gäller speciellt effekten av senare påminnelser.
- Det finns därför både kostnadseffektiva och etiska skäl till att begränsa antalet påminnelser.
- Belöningar kan ha en positiv effekt på svarsfrekvensen, speciellt om de ges i förhand.

Andra utskicksfaktorer

Personalisering

I självadministrerade datainsamlingsätt är frågeställaren begränsad till utrymmet i missivet och det eventuella förbrevet. I intervjuer finns större möjligheter att beskriva undersökningsorganisationen, undersökningens bakgrund och syfte. Intervjuaren kan presentera detta och svara på eventuella frågor. Kanske är just detta – det mänskliga mötet – anledningen till att telefonintervjuer traditionellt sett har haft högre svarsfrekvenser än självadministrerade datainsamlingsätt?

Ett sätt att minska det mänskliga avståndet i självadministrerade enkäter är att personalisera undersökningen (Bergman & Wärneryd, 1982). Det kan innebära att man skriver under missivet med en riktig underskrift, tilltalar respondenten direkt samt har ett mjukare och mer informellt språk, som om man skrev till en affärskollega (Dillman, 2007).

Redan på 1960- och 1970-talen gjordes en hel del forskning om att göra följebrevet mer personligt, snarare än att följa en standardiserad mall. Personaliserade följebrev riktade till allmänhet höjde svarsfrekvensen med sju-åtta procent. Personaliserade följebrev till anställda gav däremot mer blandade resultat. Andreasen (1970) utredde saken närmare och fann att personalisering inte nödvändigtvis ökar svarsfrekvensen. Den kan också verka i motsatt riktning och minska känslan av anonymitet. Ju mer personlig en inbjudan är, desto svagare är känslan av anonymitet och i vissa undersökningar kan det ge utslag. Andreasen drar dock slutsatsen att personalisering inte är en särskilt avgörande faktor eftersom det krävs ganska stora förändringar för att ens få en effekt.

Det går alltså att personalisera på olika sätt och det är viktigt att vara medveten om att vissa av dessa sätt kan stå i konflikt med andra mål. Exempel på det är känslan av anonymitet eller önskemålet att framställa undersökningsorganisationen som seriös och förtroendeingivande.

Sista svarsdatum

Ska man ha med ett sista svarsdatum eller inte? Att inkludera det kan signalera att det är bråttom att svara eller att tiden är begränsad, vilket skulle kunna öka svarsfrekvensen. Men det kan också finnas negativa effekter, till exempel att man tappar dem som skulle ha svarat senare, efter sista svarsdatum. I en studie upptäckte man att om man hade med ett svarsdatum i missivet ökade andelen tidiga svar, men den slutliga svarsfrekvensen blev densamma (Henley, 1976).

Erbjuda olika sätt att besvara undersökningen

Att erbjuda flera så kallade datainsamlingsmetoder är något som kan tyckas borde öka svarsfrekvensen, bland annat eftersom olika metoder kan vara olika effektiva för att få kontakt med olika grupper. Vilken datainsamlingsmetod man helst deltar i kan även variera mellan personer. I en studie testade man effekten av att kunna använda olika datainsamlingsmetoder i en medicinsk undersökning. Där hade man möjlighet att erbjuda både postenkät, telefonintervju och besöksintervju.

Telefonintervju visade sig vara bra för att nå dem som var svåra att nå via post, till exempel unga och dem med bra hälsa. 46 procent av dem som inte deltog via pappersenkät eller telefonintervju ställde upp på en besöksintervju (Gallagher, Fowler & Stringfellow, 2005).

Det är dock mer osäkert om man ska erbjuda flera datainsamlingsmetoder samtidigt, till exempel pappers- och webbenkät, i första utskicket. Det finns bland annat forskning av som visar att valmöjligheten i sig kan vara förknippad med lägre snarare än högre svarsfrekvens. En förklaring kan vara att det blir för krångligt när urvalspersoner känner att de måste välja. En annan möjlig förklaring är att ju mer det informeras om valmöjligheter, desto mer varse blir urvalspersonerna om att de faktiskt har ett val och att de då utnyttjar detta genom att inte delta alls (Medway & Fulton, 2012). Om detta behövs det dock mer forskning (läs mer om att blanda datainsamlingsmetoder i kapitel 6).

Svarsfrekvenser och etik

Svarsfrekvenserna sjunker i hela västvärlden. Alla som genomför en undersökning vill förstås att den ska bli så bra som möjligt och då är det lätt hänt att man tänker kortsiktigt. Men man måste ha urvalspersonerna i åtanke och respektera att de kanske inte vill delta.

Av etiska skäl bör respondentens beslut naturligtvis respekteras, men undersökningsorganisationer och företag måste också tänka långsiktigt, både för respondentens och för *sin egen skull*. Det är en sak att försöka övertala urvalspersoner genom att sakligt förklara undersökningens värde. Det är något annat att försöka bemöta respondentens försvar med psykologiska försäljningstekniker. Bara för att ett visst tillvägagångssätt kan öka svarsfrekvensen i en enskild undersökning här och nu bör den kanske inte användas i ett långsiktigt perspektiv. Det är värt att påpeka att väldigt lite av bortfallsforskningen har studerat de långsiktiga konsekvenserna av det de föreslår. Hur vi behandlar respondenterna idag påverkar om de vill delta i undersökningar i framtiden.

Sammanfattande punkter:

- En möjlighet att öka svarsfrekvensen är att personalisera kommunikationen (till exempel genom ett mer personligt tilltal). Effekterna är dock inte så övertygande.

- Om undersökningsorganisationen har urvalspersonernas förtroende kan det öka svarsfrekvensen att lyfta fram det.
- Att erbjuda flera datainsamlingsmetoder ökar svarsfrekvensen men om alla erbjuds på samma gång finns en risk att svarsfrekvensen inte ökar utan t.o.m. minskar.
- Det är värt att påpeka att väldigt lite av bortfallsforskningen har studerat de långsiktiga konsekvenserna av det de föreslår. Hur vi behandlar respondenterna idag påverkar om de vill delta i undersökningar i framtiden.

Sammanfattning

Enkätmetodik – sammanfattning

Avsnittsindelning:

- Frågorna ska inte komma huller om buller utan bör delas upp i avsnitt. Ett avsnitt bör innehålla minst två frågor. Oftast är en ämnesmässig uppdelning det allra bästa.
- Varje avsnitt bör ha en avsnittsrubrik som förbereder respondenten på innehållet.
- I intervjuer kan dock inte respondenten se när frågeformuläret byter ämne utan det måste intervjuaren tala om. Vanligt är att intervjuaren berättar det genom en kort introducerande text.

Ordningsföljden mellan och inom avsnitt:

- Den mest naturliga ordningsföljden är oftast att följa trattprincipen – att gå från generellt till specifikt.
- Frågeformulären ska helst inledas med frågor som är intressanta och lätta att besvara så att respondenterna känner att de kan hantera uppgiften.
- Känsliga frågor bör inte ligga först eftersom de kan ha en avskräckande effekt på hela frågeformuläret.
- Om frågeformuläret innehåller frågor som kan antas vara beroende av sammanhanget (att svaren på frågorna kan påverkas av placeringen i frågeformuläret) bör man behålla samma ordning som i tidigare frågeformulär eller som i andra undersökningar man vill jämföra med.

Instruktioner:

- Instruktioner är tilläggsinformation till frågeformuleringen. De kan utgöras av exempelvis definitioner eller regler för hur frågan ska besvaras.
- I pappersenkäter krävs det en avvägning av hur mycket instruktioner man kan ha utan att röra till layouten eller öka informationskravet för mycket.
- Man kan skilja på *nödvändiga* och *förtydligande* instruktioner. Nödvändiga instruktioner är instruktioner som respondenten verkligen bör läsa (eller höra) för att kunna besvara frågan korrekt. Förtydligande instruktioner är tänkta att användas mer "vid behov", till exempel om vissa respondenter skulle stöta på något speciellt problem.
- Om instruktioner verkligen har en avgörande betydelse bör de ligga så nära frågan som möjligt.
- I intervjuer måste det vara tydligt för intervjuaren vilka instruktioner som är nödvändiga och alltid ska läsas och vilka som endast ska läsas vid behov.
- I webbenkäter finns tekniska möjligheter att ge instruktioner vid behov. Exempelvis kan instruktioner finnas som klickbar länk eller bara vara synlig om respondenten markerar den med muspekaren.
- Det kan dock vara svårt att få respondenter att läsa instruktioner.
- För att stimulera användning av instruktioner ska de vara så enkla att aktivera som möjligt. Om instruktionen behöver aktiveras för att bli synlig är det viktigt att den tillför någon värdefull information för respondenten.

Längden på frågeformulär:

- Det kan tyckas självklart att längden på enkäten borde påverka bortfallet i riktningen att ju längre enkäten är, desto högre blir bortfallet. Forskningen har dock haft svårt att bekräfta att det faktiskt förhåller sig så.
- Förmodligen är upplevelsen av längden också viktig, vilket inte alltid beror på längden utan kan även påverkas av layouten och ämnesområdet.

- Den allmänna erfarenheten är ändå att längden har betydelse för bortfall och även för datakvaliteten. Den betydelsen är extra stor i undersökningar som inte har så många andra attraktionsfaktorer.
- Vid en telefonintervju var en tidigare rekommendation att den inte ska pågå längre än cirka 30 minuter (Bergman & Wärneryd, 1982). Idag är rekommendationen att de bör vara ännu kortare.
- Vid besöksintervjuer finns bättre möjligheter till längre intervjuer.

Förhandsinformation:

- Ett sätt att öka svarsfrekvensen är att skicka ut information till respondenterna innan insamlingen inleds – innan enkäten anländer eller intervjuaren ringer.
- Syfte: att annonsera intervjun eller enkäten, att ge information om undersökningen och att motivera respondenterna att delta i undersökningen, fungera som en "aptitretare".
- I en studie ökade förbrevet svarsfrekvensen med fyra-sex procent.
- En sammanställning av flera studier fann att introduktionsbrev ökar svarsfrekvensen med åtta procentenheter.
- De flesta introduktionsbrev innehåller namnet på undersökningen, namnet på organisationen som genomförde undersökningen, fördelarna med att delta och information om anonymitet och sekretess.
- Dillman rekommenderar att förbrevet ska vara personligt, positivt, kortfattat, förklara att undersökningen är viktig, bygga upp en förväntan och skickas ut relativt tätt in på insamlingen.

Missiv (följebrev):

- Missiv kallas det brev som medföljer enkäten. Missivet ska ge nödvändig information och motivera till deltagande.
- Respondenterna har rätt att få viss information om undersökningen innan de deltar. Sådan information måste alltid finnas med i missivet.
- För att motivera urvalspersonen att delta i undersökningen bör missivet på ett bra sätt förklara varför undersökningen och deltagandet är viktigt.

- Missivet ska ha en trevlig och lättillgänglig ton.

Påminnelser:

- Påminnelser leder generellt till fler svar även om effekten sällan är lika hög som man kanske hade hoppats på. Det gäller speciellt effekten av senare påminnelser.
- Det finns därför både kostnadseffektiva och etiska skäl till att begränsa antalet påminnelser.

Belöningar:

- Belöningar kan ha en positiv effekt på svarsfrekvensen, speciellt om de ges i förhand.

Andra utskicksfaktorer:

- En möjlighet att öka svarsfrekvensen är att personalisera kommunikationen (till exempel genom ett mer personligt tilltal). Effekterna är dock inte så övertygande.
- Om undersökningsorganisationen har urvalspersonernas förtroende kan det öka svarsfrekvensen att lyfta fram det.
- Att erbjuda flera datainsamlingsmetoder ökar svarsfrekvensen men om alla erbjuds på samma gång finns en risk att svarsfrekvensen inte ökar utan t.o.m. minskar.
- Det är värt att påpeka att väldigt lite av bortfallsforskningen har studerat de långsiktiga konsekvenserna av det de föreslår. Hur vi behandlar respondenterna idag påverkar om de vill delta i undersökningar i framtiden.

5 Layout

Författare: Pia Hartwig (SCB) och Andreas Persson (SCB)

Att kommunicera med grafik och symboler

Vi kommunicerar med respondenterna inte bara via bokstäver och ord utan även med grafik och symboler. Det kallas för layout och är framförallt relevant för pappersenkäter och webbenkäter där innehållet presenteras synligt för respondenten. I intervjuundersökningar presenteras frågeformuläret som ljud. Intervjuaren läser frågor och svarsalternativ och då finns det ingen layout för respondenten. Däremot finns det information synlig för intervjuaren i dennes datorstöd.

Layouten kan fylla flera funktioner i enkäter. Den kan underlätta eller visa hur respondenten ska besvara frågorna. Layouten kan också ha en stilistisk funktion genom att exempelvis ge enkäten ett mer trevligt intryck utan att påverka besvarandet i sig. Forskning har dock visat att även om avsikten är att detta ska vara två skilda funktioner är det inte säkert att respondenten uppfattar det så. Synlig information som är tänkt att vara stilistisk kan också påverka svaren (Schwarz, Hipler, Deutsch & Strack, 1985).

För enkäter riktade till individer och hushåll finns det en del konventioner för layouten, som att frågorna och svarsalternativen är vertikalt vänsterställda. Många enkäter följer detta format men det finns undantag. Enkäter riktade till företag och organisationer innehåller oftare tabeller med kolumner för olika typer av uppgifter. Generellt sett gäller att ju mer komplicerat upplägg på frågeinnehållet, desto viktigare blir layouten.

Layout kan alltså ha stor betydelse. Nedan redovisas flera forskningsresultat som bekräftar detta och några tips på god design. För att få en djupare förståelse kan det vara bra att inleda med hur layout uppfattas av respondenterna. För att förstå det är det bra att känna till lite om mänsklig perception – den process där vi tillgodogör oss information från våra sinnen. Kapitlet inleds därför med ett avsnitt om just detta. Sedan följer ett avsnitt om layout i allmänhet. Kapitlet avslutas med ett avsnitt om webbenkäter eftersom webben erbjuder en del unika möjligheter.

Sammanfattande punkter:

- Vi kommunicerar inte bara med språket utan även med grafik och symboler – layout.
- Layout har stor betydelse i enkäter. Den ökar respondentens uppmärksamhet, förståelse för enskilda frågor och visar vägen genom enkäten.

Perception – hur vi förstår information från våra sinnen

Perception är den process där information från sinnesorganen ges mening. Våra sinnesorgan matar oss ständigt med information. I layout är det ögonen som är det aktuella sinnesorganet. Ögonen gör inga urval av hur mycket information de ska skicka vidare eller hur informationen ska tolkas. Samtidigt är inte informationen från ögonen och det vi upplever att vi ser exakt detsamma. Olika människor kan uppfatta samma visuella stimuli på olika sätt. Perception är nämligen en aktiv och skapande process där informationen från sinnesorganen ges mening (Passer & Smith, 2007).

Perception är beroende av uppmärksamhet. Vi bombarderas hela tiden med massor av information, men det är bara en bråkdel av den informationen som vi uppmärksammar och tolkar. Var vi riktar vår uppmärksamhet är delvis under vår kontroll. Vi kan välja att fokusera på läraren vid katedern snarare än på vad som händer utanför fönstret, eller tvärtom. Vår uppmärksamhet kan dock dras till platser oavsett vårt aktiva beslut, till exempel vid en kraftig smäll fokuserar vi instinktivt på ljudets riktning. Men information kan också tas in via våra sinnesorgan utan att uppmärksammas. Vi kan titta på något utan att *se* det, för vår uppmärksamhet är riktad åt något annat håll (till exempel att vi tänker på ett problem vi måste lösa). Detta kallas för ouppmärksamhetsblindhet ("inattentional blindness" på engelska). Det är ett skäl till varför det kan finnas risker med att köra bil och tala i mobiltelefon samtidigt, oavsett handsfree. Om hjärnan ägnar stor uppmärksamhet åt telefonsamtalet förenklas syninformationen (Passer & Smith, 2007).

Att tolka visuell information är en komplicerad process. Att den ger så pass ofta ger ett självklart resultat är imponerande. Informationen från ögonen är egentligen bara olika varianter av ljus, men vi förstår det som separata objekt som på något sätt som blir en förståelig helhet. Denna

process kräver någon form av organisation där visuella stimuli kategoriseras.

Scheman – förkunskaper påverkar perceptionen

Perception utgår i regel från att vi har vissa förkunskaper. Vi upplever inte allting på nytt varje gång utan vi har så kallade perceptuella scheman – mentala representationer över distinkta egenskaper hos människor, händelser, objekt och så vidare. När vi träffar en arbetskamrat aktiveras det perceptuella schemat för den personen och vi känner igen personen som vår arbetskamrat. Det betyder att en del av perceptionen utgörs av att hitta den mest lämpliga tolkningen, det bästa schemat, för den visuella informationen. Det är inte alltid helt enkelt, speciellt i situationer som är otydliga. Var det en gren eller en människoarm som rörde sig vid trädet i den dunkla natten? Ögat bryr sig inte om att avgöra den saken utan det är en tolkning som sker inne i vårt huvud. Den kan påverkas av personlighetsdrag, kultur och förväntningar – de perceptuella schemana som ligger nära till hands för personen. Olika personer kan alltså tolka samma visuella information på olika sätt men samma person kan också tolka samma stimuli på olika sätt i olika situationer. Exempelvis kanske vi är mer benägna att se en arm än en gren om vi nyss läst en spökhistoria (Passer & Smith, 2007).

De perceptuella schemana kan generaliseras över olika situationer. Och det är tur, annars skulle vi behöva tolka objekt på nytt varje gång de uppträder i en ny situation eller ser lite annorlunda ut. När en bil åker mot horisonten blir den mindre till storleken, men vi uppfattar det som samma bil. Detta kallas för perceptionskonstans och möjliggör alltså att vi kan uppfatta kända stimuli under olika förutsättningar. Ett exempel är ljusstyrkakonstans, att objekt uppfattas som samma i olika ljusstyrkor som dagsljus eller dunkel. Ett annat exempel är formkonstans. Det betyder att vi kan uppfatta ett objekt som samma oavsett om vi betraktar det ur olika vinklar, varifrån objektet ser olika ut. Ett tredje exempel är storlekskonstans, som betyder att vi kan uppfatta det som samma objekt oavsett om storleken varierar på vår näthinna (exemplet med bilen ovan) (Passer & Smith, 2007).

Ovanstående är viktigt eftersom form, ljusstyrka och storlek tillsammans med färg är de fyra storheter vi har till vårt förfogande när det gäller en enkäts layout.

Gestaltlagarna

Förutom scheman finns det andra regler som styr vår perception. En samling sådana regler är de så kallade gestaltlagarna (Passer & Smith, 2007). "Gestalt" kommer från tyskan och ska i dessa sammanhang tolkas som form eller mönster. Ett sådant exempel är människors tendens att organisera stimuli i förgrund-bakgrund – att en figur ligger "ovanpå" en bakgrund. Vad som är förgrund-bakgrund är dock inte alltid helt enkelt att avgöra och vår upplevelse av samma stimuli kan ändras. På bilden nedan kan antingen en vas eller två ansikten i profil uppfattas vara i förgrunden. Här finns det ingen rätt tolkning.

Figur 5.1



Gestaltpsykologerna lanserade ett antal lagar för organisation av visuella stimuli. En sådan lag är *lagen om likhet* – att objekt som liknar varandra i till exempel färg, form och storlek upplevs höra ihop eller ha samma betydelse.

Figur 5.2



Personer brukar uppfatta ovanstående som en triangel, bestående av nio mindre trianglar, omsluten av cirklar. Det är ingen given tolkning. Figuren skulle lika gärna ha kunnat uppfattas som fem kolumner med

blandade objekt. Men de trekantiga objekten förväntas höra ihop enligt *lagen om likhet* och på så vis bilda en triangel. I enkätdesign innebär det till exempel att en symbol som används i enkäter alltid ska ha samma betydelse.

En annan lag är *lagen om närhet* – att delar som ligger nära varandra upplevs höra ihop.

Figur 5.3



Figuren till vänster upplevs av de flesta bestå av tre rader. Figuren till höger upplevs att bestå av tre kolumner. Den stora skillnaden mellan figurerna, och också orsaken till att de upplevs annorlunda, är att avståndet till nästa objekt är mindre horisontellt än vertikalt i den första figuren och tvärtom i den andra. Avstånd är alltså en viktig faktor för hur visuell information uppfattas. Det innebär bland annat att frågetext, instruktion och svar till en och samma fråga ska ligga nära varandra, medan det ska vara avstånd till nästa fråga. På det sättet blandar man inte ihop de olika frågorna. Det innebär också att svarsalternativ ska hållas ihop, så att avstånden inte varierar.

En tredje lag är *lagen om slutenhet* – att människor tenderar att "fylla i" tomrum så att visuella stimuli antar formen av kända figurer.

Figur 5.4



Ovanstående brukar uppfattas som en cirkel (i det här fallet ett runt ansikte) trots att det faktiskt saknas en bit för att vara en komplett cirkel.

Den fjärde lagen är *lagen om kontinuitet* – att människor tenderar att uppleva stimuli som att de fortsätter som man kan förvänta sig, till exempel i kontinuerliga mönster.

Figur 5.5



I ovanstående figur är det lättare att tänka sig att linjen i vägskälet fortsätter i samma riktning som den var på väg i (uppåt) än att den byter riktning och går neråt, trots att båda egentligen är lika troliga. I enkät-design betyder detta att en mjuk övergång mellan företeelser leder till att de tolkas som att de hör ihop. Grafiska verktyg kan användas för att göra ett tydligt flöde genom en enkät.

En annan lag är *lagen om pregnanz*. Enligt denna föredrar vi några få stora enkla objekt framför flera små och komplicerade.

Figur 5.6



Ovanstående figur skulle kunna uppfattas som väldigt komplicerad eller bestå av många mindre figurer. Men de flesta brukar uppfatta det som en stor kvadrat och en stor triangel. På grund av detta är det viktigt att den övergripande layouten är tydlig. De detaljerade delarna syns först när den övergripande visuella layouten tolkats. En snabb blick på enkäten

ska visa områden som är enkla, symmetriska och strukturerade – som ger ett tilltalande första intryck.

Den sista lagen är *lagen om gemensamt område*. Denna innebär att människor tolkar mindre objekt inom ett större som att de hör ihop. Det innebär till exempel att allt som står i ett sidhuvud eller på en bakgrund av specifik färg tolkas som sammanhörande. Ramar och färger kan visa vad som hör ihop.

Dessa lagar (Passer & Smith, 2007) ligger bakom en del av forskningen och rekommendationerna angående layout i enkäter.

Sammanfattande punkter:

- Perception är den process när information från sinnesorganen ges mening.
- Perceptionen är bland annat styrd av förkunskaper – vi förstår inte allting på nytt varje gång det uppträder.
- Gestaltlagarna är en samling lagar för hur människor sorterar och förstår visuella stimuli, till exempel lagen om likhet = att objekt som ser likadana ut förväntas höra ihop och betyda samma sak.

Allmänt om design

Kognitiv design

Innan vi går in på enkätlayout kan det vara bra att känna till en del om allmänna designprinciper som inte bara gäller enkäter, utan alla föremål där det finns en användare. Det handlar om att göra föremålet som ska användas mer förståeligt och intuitivt. Det finns en bred litteratur om detta som kallas användbarhet ("usability" på engelska). Litteraturen fokuserar inte främst på enkäter men en del av principerna kan tillämpas även där. Några principer är (Norman, 1988; Dillman, Gertseva & Mahon-Haft, 2005; Björnram, Erikson, Furubjelke, Hartwig, Henningsson & Lorenc, 2006):

Använd kända kunskaper

Dra nytta av respondentens tidigare kunskaper om hur det fungerar att fylla i en enkät. Var också konsekvent i enkäten. När respondenten väl lärt sig hur formatet fungerar i enkäten är det onödigt att tvinga dem att

lära om igen. Det finns inget egenvärde i att vara innovativ eller påhittig i designen av enkäten.

Gör saker synliga

Om saker inte syns kan de inte heller användas. Styr respondentens uppmärksamhet rätt utifrån uppgiftens krav – gör saker som ska användas synliga.

Använd begränsningar

Valmöjligheter eller för mycket frihet kan leda till otydlighet för respondenten. Begränsa därför antalet valmöjligheter, till exempel genom att svarsfält har lagom många positioner för den typ av svar som önskas. Ett annat exempel är att rutor för kryss inte ska förväxlas med rutor för siffror eller text.

Om inget annat funkar – standardisera

Det kan vara omöjligt att känna till eller tillgodogöra respondenternas tidigare erfarenheter av enkäter. Därför är det viktigt att enkäten är konsekvent designad – att det är standardiserat. Oavsett tidigare erfarenheter kan respondenterna i alla fall lära sig hur just denna enkät fungerar.

Sammanfattande punkter:

- Kognitiva aspekter av allmän design: använd kända kunskaper, gör saker synliga och använd begränsningar.
- Om det är svårt att få ovanstående att fungera – standardisera.

Layout i enkäter (pappers- och webbenkät)

Målet med layouten i en enkät är att underlätta besvarandet och ge ett positivt intryck. För att uppfylla målen finns ett antal verktyg till förfogande:

1. text – frågorna och svarsalternativen naturligtvis, men även introduktioner och instruktioner
2. font – det teckensnitt som används till texten
3. symboler – som pilar eller svarsrutor
4. bilder – som foton eller illustrationer

5. siffror – som numrering
6. linjer
7. färger – både gråskala och olika färger
8. det rumsliga – hur ovanstående (1–7) bäst ska placeras på den begränsade yta som finns till förfogande i enkäten.

Det finns specifika tips för punkt 1–7 men det är nog ändå 8 som kräver mest eftertanke. Resterande del av kapitlet handlar om hur man bäst tillämpar ovanstående verktyg för att få en bra layout på sin enkät.

Korta regler för layout i enkäter

Det finns inte så många vetenskapliga resultat att vila sig mot när det gäller layout i enkäter utan ofta handlar det om att försöka tillämpa mer generella regler om hur människor bäst tar till sig information. Utifrån regler om allmän design och bland annat gestaltlagarna har Dillman, Gertseva och Mahon-Haft (2005) formulerat tolv regler för layouten i företagsundersökningar. Reglerna är även relevanta för andra typer av enkäter. Några av dem är:

- **Avgränsa enkäten i ämnesmässiga avsnitt:** En enkät innehåller mycket information. För att respondenten ska kunna tillgodogöra sig innehållet så effektivt som möjligt behöver det struktureras på något sätt (läs mer i kapitel 4). Oftast är en ämnesindelning bäst. Detta minskar både respondentbördan och förbättrar mätningen. Avsnitten bör inledas med en rubrik där det tydligt framgår vad frågorna handlar om. Rubriken ska inte användas istället för frågetext och frågetexterna ska inte hänvisa tillbaka till rubriken. Frågan ska kunna förstås i sig självt utan avsnittsrubriken. Rubriken ska vara kort och koncis. Den ska inledas med versal bokstav men inte skrivas med enbart versaler eftersom det gör den svårsläst. Det är lättare för ögat att snabbt tolka gemena bokstäver eftersom de varierar mer i höjdlängd än versala bokstäver.
- **Gör information som inte gäller respondenten mindre synlig:** Enkäter innehåller ibland information som inte är till för respondenterna utan för undersökningsorganisationen. Det kan vara löpnummer eller annan kod som är tänkt att vara till hjälp vid registrering av svaren. Detta behöver inte respondenten belastas med och därför bör sådan information vara perifer och mindre synlig.

- **Använd symboler konsekvent:** Man bör vara lite försiktig med användandet av symboler, speciellt sådana som respondenten kanske inte känner igen sedan tidigare. Vissa symboler kan dock fylla viktiga funktioner (till exempel för att uppmärksamma informationsrutor) och då bör den funktionen vara densamma i hela enkäten. Annars måste respondenten lära om när symbolen ändrar betydelse.
- **Etablera visuella kopplingar mellan sidor, exempelvis med hjälp av pilsymboler:** Det är viktigt att respondenten enkelt kan se hur man går vidare i enkäten. Detta gäller framförallt i pappersenkäter som har mycket innehåll och en svår struktur. I webbenkäter går det istället att lösa problemet med hjälp av automatiska hopp.
- **Markera tydligt med hjälp av layout när respondenten förväntas ändra sitt svarsbeteende:** En av de största riskerna för fel är när respondenten förväntas bryta sitt beteende och göra något nytt. Därför bör sådana brott vara tydliga i layouten, till exempel om enkäten byter ämne eller om frågor byter svarsskala eller referensperiod.

Sammanfattande punkter:

- De verktyg som finns till förfogande i layout är: text, font, symboler, bilder, siffror, linjer och färger. Ett viktigt beslut är hur dessa bäst ska placeras på den begränsade yta som finns i enkäten.
- Se punktlista i texten för korta regler för layout i enkäter.

Layout för frågor

Frågorna och svarsalternativen är de viktigaste delarna i en enkät, därför är det viktigt att de har en bra layout. Det finns inte så mycket forskning på detta område, men med goda erfarenheter och kunskap om design i allmänhet går det dock att formulera några tumregler:

Tydliggör vad som är inledande text, frågeformulering och instruktion

En fråga kan bestå av olika delar, exempelvis inledande text, introduktion, frågeformulering eller instruktion. Dessa kan fylla olika funktioner och är också olika viktiga. Frågeformuleringen är till exempel

viktigare än en introduktion och därför kan det finnas ett värde i att särskilja dessa.

Ett annat syfte är att underlätta läsningen för respondenterna. Det kan till exempel hända att de behöver läsa om frågeformuleringen eller upprepa en instruktion. Om man då har separerat dessa delar i layouten blir det enklare för respondenten att hitta dem. Det konventionella för pappersenkäter är att skriva en instruktion i kursiv stil och frågeformuleringen i fet stil.

Inledande text brukar skrivas i normal stil och kan dessutom särskiljas från frågeformuleringen med en eller flera blankrader.

Använd vänsterställd text

Text är vanligtvis vänsterställd. Därför bör enkäter följa, snarare än bryta, den normen. Text som inte är vänsterställd blir mer svårläst och riskerar också att missas. Detta gäller även följdfrågor. Det händer att enkäter har följdfrågor inskjutna för att visa att de är en följd av den tidigare frågan. Det finns dock risker med detta eftersom följdfrågan kan missas vilket kan resultera i partiellt bortfall, se exempel A och B.

EXEMPEL A

3. Hur stor var månadshyran för lägenheten i december 2015?

kr

4. Ingår värme i hyran?

Ja
 Nej

5. Har lägenheten under 2016 genomgått ombyggnad eller större förändring, ombesörjd av fastighetsägaren?
Exempelvis ändrat rumsantal, total upprustning av kök/badrum, stamrening.

Typ av åtgärd:

Ja →
 Nej

EXEMPEL B

3. Hur stor var månadshyran för lägenheten i december 2015?

kr

4. Ingår värme i hyran?

Ja
 Nej

5. Har lägenheten under 2016 genomgått ombyggnad eller större förändring, ombesörjd av fastighetsägaren?
Exempelvis ändrat rumsantal, total upprustning av kök/badrum, stamreovering.

Typ av åtgärd:

Ja →
 Nej

Tänk på radlängden

Radlängden på en lång fråga får varken vara för kort eller för lång. Är frågan för lång före radbrytning blir den svår att överblicka, se exempel C. Är frågan uppdelad på alltför korta rader blir den istället svår att läsa, se exempel D. Frågelängden ska också anpassas till längden på övriga frågor på sidan, om sådana finns. Om radlängderna i en text eller på en sida skiljer sig för mycket åt kan det ge ett rörigt intryck.

En tumregel är att läsbarheten i en lång fråga försämras om man använder mer än 58 tecken eller mindre än 35 tecken per rad. Radbrytningen gör därför en lång fråga lättare att läsa. Maxlängden 60–70 tecken inklusive mellanslag är rekommendationen.

Om man utgår från den rekommenderade radbrytningen så blir skillnaden på en kort och en lång fråga inte så stor utan respondenten kan snabbt få en bra överblick över frågor som ligger på samma sida, se exemplet i E. Detta underlättar besvarandet och minskar respondentbördan.

EXEMPEL C

Lämna uppgifter	Instruktioner	Kontakta oss	
Start	Svara	Skicka in	Bekräftelse

4 När senast genomfördes en kvalitetsgranskning av journalföring?

- Mindre än 4 veckor sedan
- Mellan 1 och 3 månader
- Mellan 4 och 6 månader
- Mellan 7 och 12 månader
- Över 12 månader sedan
- Kommer inte ihåg

5 Har du någon gång tagit del av resultat från en kvalitetsgranskning av journalföring, i form av en muntlig eller skriftlig redovisning?

- Ja
- Nej
- Vet inte

EXEMPEL D

Lämna uppgifter	Instruktioner	Kontakta oss
Svara	Skicka in	Bekräftelse

3 Hur många år har du studerat vid universitet/högskola eller annan eftergymnasial utbildning i annat land än Sverige?

- Ingen eftergymnasial utbildning
- 1 år
- 2 år
- 3 år
- 4 år eller längre
- Doktorsexamen (motsvarande PhD)

4 Vilket år tog du din högsta examen?

År

[◀ Bakåt](#)

EXEMPEL E

SvaraSkicka inBekräftelse

1 När senast genomfördes en kvalitetsgranskning av journalföring?

- Mindre än 4 veckor sedan
- Mellan 1 och 3 månader
- Mellan 4 och 6 månader
- Mellan 7 och 12 månader
- Över 12 månader sedan
- Kommer inte ihåg

2 Har du någon gång tagit del av resultatet från en kvalitetsgranskning av journalföring, i form av en muntlig eller skriftlig redovisning?

- Ja
- Nej
- Vet inte

Använd olika avstånd för att visa vad som hör ihop eller inte

Lagen om närhet är ett viktigt verktyg i layout. Att använda avstånd är ett av de bästa sätten att visa vad som hör ihop eller inte (se även nästa underrubrik). Man kan använda avstånd på flera nivåer. Dels på enkätnivå, att frågor som handlar om samma ämne ska ligga samlade i samma avsnitt, men också på frågenivå. All information som hör till en fråga ska finnas nära varandra och det ska finnas avstånd till nästa fråga.

En följd av detta tankesätt är också att instruktioner bör finnas så nära deras användningsområden som möjligt. Instruktioner bör alltså finnas vid den aktuella frågan och inte till exempel i missivet. I förlängningen kan även tankesättet om närhet tillämpas inom frågor. Om en definition av ett begrepp är väldigt viktig kan det ibland vara bättre att bygga in den i frågeformuleringen än att ha det som en kursiv instruktion under frågeformuleringen, se exempel F och G nedan.

EXEMPEL F

Har du någon gång tagit del av resultatet från en kvalitetsgranskning av journalföring?

Resultat i form av en muntlig eller skriftlig redovisning.

- Ja
- Nej
- Vet inte

EXEMPEL G

Har du någon gång tagit del av resultatet från en kvalitetsgranskning av journalföring, i form av en muntlig eller skriftlig redovisning?

- Ja
 Nej
 Vet inte

Avgränsa innehåll med hjälp av linjer, ramar eller gråskala

Förutom avstånd (se ovan) är linjer det främsta verktyget för att visa vad som hör ihop eller inte. En linje har en tydligt särskiljande funktion, se exempel H. I många fall kan endera sätten, avstånd eller linjer, användas. Avstånd är dock mer behagligt för ögat än linjer, speciellt om det är många linjer som är asymmetriskt utplacerade.

Symmetriska linjer, som i en tabell, fungerar dock bra. De kan även kombineras med att till exempel varannan rad är gråmarkerad.

EXEMPEL H

Intäkter	Radnr	Belopp (tkr)
Nettoomsättning prästlönefastigheter	001	
Intäkter prästlönefond	003	
Resultat vid avyttring av prästlönefastigheter	005	
Övriga rörelseintäkter	007	
Förändring av lager, produkter i arbete, färdiga varor och pågående arbeten	009	
Summa rörelsens intäkter:	011	0
Kostnader		
Kostnader för varor, material och vissa köpta tjänster	013	
Kostnader för inhyrd personal	015	
Övriga externa kostnader	017	
Personalkostnader	019	
Av- och nedskrivningar av materiella anläggningstillgångar	021	
Övriga rörelsekostnader	023	
Summa rörelsens kostnader:	024	0

Kopplingen mellan huvud- och delfråga

Det är relativt vanligt att ha en huvudfråga och sedan ett antal delfrågor. I sådana upplägg kan det finnas möjlighet att förkorta delfrågorna. Om alla till exempel har samma inledning kan inledningen brytas ut till en huvudfråga och därefter kan respektive delfråga endast innehålla de unika delarna. Detta kan göras genom att till exempel avsluta huvudfrågan med tre punkter ("...") och inleda varje delfråga med detsamma. Punkterna visar då att delfrågorna hör ihop med huvudfrågan.

Layoutmässigt är ovanstående lösning bra. Det sparar plats samtidigt som det är lättare för respondenten att läsa korta delfrågor som inte upprepar samma sak. Det finns emellertid risker med den lösningen och med huvudfrågor i allmänhet. Speciellt om det är många delfrågor kan det hända att huvudfrågan mer och mer glider ur respondenternas minne och att de börjar svara för varje delfråga utan att ha huvudfrågan i åtanke. Detta kan förstås leda till problem, exempelvis om huvudfrågan tillför en viktig avgränsning (som en referensperiod).

Ibland kan alltså den bästa lösningen vad gäller tydlighet och respondentbörda hamna i konflikt med det frågetekniska. Uppdelningen i huvud- och delfrågor kan vara ett sådant exempel och det måste bedömas från fall till fall, bland annat beroende på hur många delfrågor det är och hur viktig huvudfrågan är för svaret på delfrågorna.

Tänk också på att i en webbenkät med många delfrågor kan huvudfrågan scrolas bort, vilket ytterligare förstärker ovanstående.

Förstärkta ord i frågor

I vissa fall behöver specifika ord framhävas. Det är oftast i frågeformuleringen men kan även gälla instruktioner eller i svarsalternativen. Till det används i första hand fet stil. Om texten redan är i fet stil, exempelvis i en fråga, fungerar det att använda kursiv stil som förstärkning, se exempel I. Understrykningar ska helst inte användas i webbenkäter eftersom de symboliserar länkar.

EXEMPEL I**Har företaget haft utgifter för utlagd FoU under år 2013?**

FoU som företaget givit i uppdrag till andra (i Sverige eller utomlands) att utföra, räkna också med bidrag som företaget lämnat för FoU t.ex. till universitet och högskolor.

- Ja
 Nej

Förstärkning av ord bör främst användas för att uppmärksamma respondenten på förändringar i texten som inte är förväntade. Exempelvis när något ska ingå som vanligtvis inte gör det: Vad är din lön efter skatt? Det kan även vara för att uppmärksamma respondenten på att det är olika referensperioder i enkäten.

Förstärkta ord bör användas sparsamt eftersom de drar till sig uppmärksamhet. För många förstärkta ord i en text eller på en sida kan påverka läsbarheten negativt istället för att förtydliga. Därför bör de endast användas där det är befogat och på ett konsekvent sätt genom hela enkäten.

Gör gömda frågor synliga

Gömda frågor löper större risk att missas och därmed ge partiellt bortfall. Det kan tyckas självklart, men ändå är det inte ovanligt att enkäter innehåller "gömda" frågor. Vad är då en gömd fråga? Det kan till exempel handla om en följdfråga som är inskjuten i marginalen. Det kan också handla om en fråga som ligger längst ner på en sida, eller under en fast sidfot om man använder webb. Den här typen av "gömda" frågor är de vanligaste – de som är perifera på papperet eller på sidvisningen. Frågor kan också vara layoutmässigt "gömda". Om exempelvis en sida domineras av frågor med en viss layout eller uppställning kan en fråga med en avvikande layout missas för att respondenten koncentrerar sig på det som är huvudsakligt och mest framträdande på sidan.

Vad ska man då göra åt saken? Det beror på varför frågan är gömd. Inskjutna följdfrågor eller fast sidfot bör till exempel inte användas alls.

I andra fall kan en lösning vara att flytta frågan till nästa sida. Om frågan ämnesmässigt däremot hör ihop med övriga frågor på sidan bör den inte ligga på en egen sida. Då är det bättre att med hjälp av en bakgrundslayout lyfta fram frågan, se fråga 2 i exempel J och K.

En rubrik kan lyfta fram frågan ytterligare, se exempel L. Frågor måste synas på lika villkor, därför är det viktigt att vara uppmärksam på "gömda" frågor vid formulärkonstruktionen.

EXEMPEL J

The screenshot shows a web form with a navigation bar at the top containing three tabs: "Lämna uppgifter" (highlighted in orange), "Instruktioner", and "Kontakta oss". Below the navigation bar is a progress indicator with three steps: "Svara" (highlighted in orange), "Skicka in", and "Bekräftelse". The main content area contains two questions:

2 Vilket år är Du född?

År: 19

3 Vilken är Din högsta avslutade utbildning?

- Grundskola eller liknande
- Gymnasieskola eller liknande
- Universitet, högskola eller liknande
- Har ännu inte avslutat grundskolan

EXEMPEL K

The screenshot shows a web form with a navigation bar at the top containing three tabs: "Lämna uppgifter" (highlighted in orange), "Instruktioner", and "Kontakta oss". Below the navigation bar is a progress indicator with three steps: "Svara" (highlighted in orange), "Skicka in", and "Bekräftelse". The main content area contains two questions:

2 Vilket år är du född?

År: 19

3 Vilken är din högsta avslutade utbildning?

- Grundskola eller liknande
- Gymnasieskola eller liknande
- Universitet, högskola eller liknande
- Har ännu inte avslutat grundskolan

EXEMPEL L

The screenshot shows a web form interface with a top navigation bar containing three tabs: "Lämna uppgifter", "Instruktioner", and "Kontakta oss". Below the navigation bar is a progress bar with three steps: "Svara" (highlighted in orange), "Skicka in", and "Bekräftelse". The main content area is titled "Bakgrundsfrågor, fortsättning" and contains two questions:

2 Vilket år är du född?
År: 19

3 Vilken är din högsta avslutade utbildning?

- Grundskola eller liknande
- Gymnasieskola eller liknande
- Universitet, högskola eller liknande
- Har ännu inte avslutat grundskolan

Förenkla tabeller

En tabell är ett effektivt sätt att ställa upp information. Hela sidan utnyttjas. Det upprepande formatet kan dock leda till slarvigt ifyllande.

Det finns inte så mycket forskning om tabeller och datakvalitet men en erfarenhetsbaserad rekommendation är att göra rader och kolumner så stora att all nödvändig information får plats utan att det blir trångt och otydligt. Det finns ändå alltid en risk för negativa svarsmönster, till exempel att respondenten markerar det första svarsalternativet oavsett frågan. Detta är svårt att göra så mycket åt, förutom genom att skära ner på antalet frågor. Alternativet, att ställa tabellfrågorna som separata frågor, leder nämligen till ett ökat antal sidor i enkäten. Detta kan i sin tur ge sämre mätkvalitet. Man bör dock vara medveten om problematiken och förenkla tabellen så långt det går.

På webben bör en tabell i möjligaste mån synas i sin helhet på skärmen. Den ska inte vara bredare än skärmbilden, men kan vid behov vara något längre. En stor tabell är inte bara komplicerad och svår att överblicka utan kan dessutom ta lång tid att ladda ner. Minska inte ner teckenstorleken för att få rum med en stor tabell – då blir den inte bara komplicerad utan även svårsläst. En stor tabell kan vinna på att delas upp två i tabeller.

Sammanfattande punkter:

- Tydliggör i layouten vad som är inledande text, frågeformulering och instruktion.
- Använd vänsterställd text.
- Radlängden på en lång fråga får varken vara för kort eller för lång. Maxlängden 60–70 tecken inklusive mellanslag är rekommendationen.
- Använd olika avstånd för att visa vad som hör ihop eller inte.
- Avgränsa innehåll med hjälp av linjer, ramar eller gråskala.
- Koppla ihop huvud- och delfråga t.ex. med tre punkter ("...").
- Använd främst fet stil för att förstärka ord i frågan.
- Gör gömda frågor synliga.
- Se bildexempel i texten.

Layout för svarsalternativ

Svarsalternativen är lika viktiga som frågorna i en enkät. Därför är det viktigt att de också har en bra layout. Tumreglerna är följande:

Använd kända svarsfält

Det säkraste är att använda svarsfält som respondenterna känner till sedan tidigare. Det innebär till exempel kryssrutor i en pappers- eller webbenkät eller radioknappar i en webbenkät. Det är de typiska svarsfälten och det finns massvis med beprövad erfarenhet av att de fungerar bra.

Det finns förstås många andra typer av svarsfält, allt från termometrar till glada och sura ansikten. Sådana alternativa svarsfält behöver naturligtvis inte orsaka problem, men det finns inte samma erfarenhet att luta sig mot. Dessutom kräver de mer engagemang och tankemöda från respondenterna. Att ha många olika typer av alternativa svarsfält torde öka respondentbördan.

Ibland används siffror, som att respondenten ska ringa in den siffra som passar. Dessa kräver mer möda att besvara och därför är svarsfält bättre.

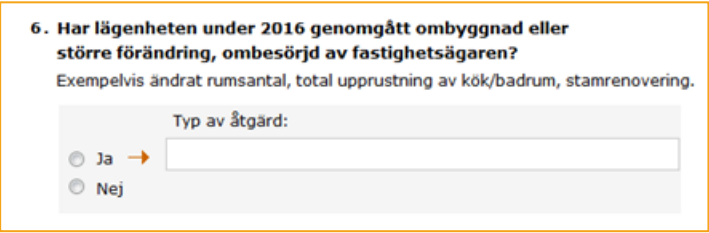
Ibland används också siffror i kombination med etiketter på svarsalternativen. Det vill säga att svarsalternativen beskrivs både med ord och med siffror. Om man använder denna lösning måste man tänka

på att både orden och siffrorna ska beskriva samma sak. Det kan till exempel bli förvirrande om man har en femgradig skala där etiketten "Väldigt ofta" kombineras med siffran ett och "Aldrig" med siffran fem. Då går riktningarna för etiketterna och siffrorna på tvärs (från högt till lågt och tvärtom).

Lyft fram svarsrutor

Svarsrutorna är oftast den enda plats i enkäten där respondenten behöver vara aktiv. Det är också markeringarna i svarsrutorna som sedan sammanställs i en datafil. För både respondenten och resultatet är det därför viktigt att svarsrutorna är synliga. I en konventionell enkätlayout krävs ingen speciell åtgärd för att uppfylla detta. Svarsrutan är tillräckligt tydlig som den är. I mer komplicerade enkäter kan det dock vara bra att lyfta fram den ytterligare, till exempel genom att placera den mot en grå bakgrund, se exempel M.

EXEMPEL M



6. Har lägenheten under 2016 genomgått ombyggnad eller större förändring, ombesörjd av fastighetsägaren?
Exempelvis ändrat rumsantal, total upprustning av kök/badrum, stamrening.

Typ av åtgärd:

Ja →

Nej

Svarsalternativens placering

Svarsrutorna brukar vanligtvis placeras till vänster om respektive svarsalternativ, se exempel N. Det kan tyckas lite konstigt att svarsrutan ligger före svarsalternativet, eftersom vi vill att respondenterna ska läsa igenom svarsalternativet innan de markerar svaret i rutan. Anledningen till detta är att om de placeras till höger om svarsalternativen blir det inte nödvändigtvis en rak kolumn av svarsrutor om svarsalternativen är olika långa. En sådan uppställning ger ett rörigt intryck och blir därför svår för respondenten att besvara, se exempel O.

EXEMPEL N**3 På vilket sätt fick du kontakt med ditt första arbete?**

Markera högst två alternativ.

- Examensarbetet
- Arbetsförmedlingen
- Rekryteringsföretag
- Platsannons
- Arbetsmarknadsdagar/ företagspresentationer på KTH
- Lade in CV i Internet-databas

EXEMPEL O**3 På vilket sätt fick du kontakt med ditt första arbete?**

Markera högst två alternativ.

- Examensarbetet
- Arbetsförmedlingen
- Rekryteringsföretag
- Platsannons
- Arbetsmarknadsdagar/ företagspresentationer på KTH
- Lade in CV i Internet-databas

En annan rekommendation är att ha svarsalternativen vertikalt, inte horisontellt. Horisontell uppställning sparar oftast plats men kan leda till otydligheter om vilket alternativ en svarsruta faktiskt tillhör, se exempel P.

EXEMPEL P**3 På vilket sätt fick du kontakt med ditt första arbete?**

Markera högst två alternativ.

- Examensarbetet
- Arbetsförmedlingen
- Arbetsmarknadsdagar
- Platsannons
- Rekryteringsföretag
- Lade in CV i Internet-databas

Placering av icke-svar

Svarsalternativen skrivs under varandra med samma avstånd med undantag av så kallade icke-svar. Icke-svar som "Vet inte", "Inte aktuellt" eller "Ingen åsikt" ska stå en bit ifrån övriga för att markera att de inte hör till det förväntade svaret, se exempel Q (läs mer i kapitel 3).

EXEMPEL Q

Enligt din bedömning, vilken möjlighet har de anställda till distansarbete?

	Arbetare	Tjänstemän	Chefer
Inga möjligheter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Små möjligheter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vissa möjligheter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ganska stora möjligheter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stora möjligheter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vet inte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Undvik otydligheter

En bra ledstjärna vid layout är att om något kan missuppfattas så kommer det att missuppfattas, åtminstone av någon. Om man har det i åtanke kan många otydligheter i layouten upptäckas och åtgärdas innan enkäten tas i bruk. Om ett värde till exempel ska lämnas i ett speciellt fält i en pappersenkät bör inte prickrader användas fram till fältet. Då kan det hända att respondenten fyller i på prickraden istället för i det avsedda fältet.

Sammanfattande punkter:

- Det säkraste är att använda svarsfält som respondenterna känner till sedan tidigare. Det innebär till exempel kryssrutor i en pappers- eller webbenkät eller radioknappar i en webbenkät.
- Lyft fram svarsrutorna i layouten.
- Svarsrutorna brukar vanligtvis placeras till vänster om respektive svarsalternativ.
- Svarsalternativen skrivs under varandra med samma avstånd med undantag av så kallade icke-svar. Icke-svar som "Vet inte", "Inte aktuellt" eller "Ingen åsikt" ska stå en bit ifrån övriga för att markera att de inte hör till det förväntade svaret.

Hoppänvisningar vid filterfrågor

Självadministrerade enkäter har ibland frågor som inte är relevanta för alla respondenter. En vanlig lösning på det problemet är att bygga in filter i enkäten så att de respondenter som frågorna inte är aktuella för

kan hoppa över dem (läs även mer i kapitel 2 om frågeteknik). I webb-enkäter programmeras hoppen i webbverktyget men för pappersenkäter byggs filter av frågor med hoppanvisningar. Där instrueras respondenter som nästkommande frågor inte är relevanta för att hoppa vidare i enkäten.

Enkäter med hoppanvisningar har högre partiellt bortfall än enkäter utan hoppanvisningar (se Redline, Dillman, Smiley, Carley-Baxter & Jackson, 1999). Partiellt bortfall är resultatet av det ena av de två fel som kan uppstå vid hoppanvisningar – att respondenterna hoppar över frågor de borde ha svarat på. Den andra typen av fel är att respondenterna svarar på frågor som de borde ha hoppat över. Hur vanliga dessa fel är varierar. Det beror på enkätens egenskaper, respondenterna och hur hoppanvisningarna är konstruerade.

Varför förekommer de här felen? En orsak till att respondenterna fyller i för få frågor (partiellt bortfall) i enkäter med hoppanvisningar kan vara att de missförstår och "hoppar" för långt. Den vanligaste orsaken till att respondenter fyller i för mycket är antagligen att de inte uppmärksammar hoppinstruktionen. Det kan finnas flera skäl till detta. Ett skäl har med förväntningar att göra. Respondenter förväntar sig nog att frågorna ska besvaras löpande utan avbrott, för det är så enkäter brukar besvaras. Även enkäter med hoppanvisningar bekräftar denna förväntning i en majoritet av frågorna eftersom hoppanvisningar är sällsynta lösningar. Hopp är alltså ett undantag, aldrig regel, och kan därför lätt missas. Ett annat skäl är att respondentens intresse kan påverka uppmärksamheten. Det kan leda till att fokus finns främst på frågeformuleringens innehåll och att hoppanvisningar därför missas. Denna risk förstärks av att hoppanvisningar ofta inte är konstruerade att dra uppmärksamhet till sig. Hur ska då hoppanvisningar konstrueras?

EXEMPEL

→ *Gå till fråga 6*

Hoppanvisningen ska ligga nära det aktuella svarsalternativet (lagen om närhet) och det måste vara tydligt vad respondenten ska göra. Här finns anledning att använda symboler, till exempel en pil, som förstärker budskapet att respondenten ska "hoppa" framåt i enkäten. Hoppanvisningen kan dessutom lyftas fram ytterligare genom att skrivas i fet stil.

Sammanfattande punkter:

- Hoppanvisningar kan leda till fel och bör därför vara tydligt konstruerade (se exempel).

Bilder i enkäter

Vi kommunicerar som sagt inte bara via språket med respondenterna, utan även med grafik och symboler. En ytterligare möjlighet som blivit alltmer relevant och vanlig i och med att webbenkäter ökat i popularitet, är att kommunicera med hjälp av bilder och fotografier. Vilken roll kan då bilder fylla i en enkät? De kan vara knutna till själva besvarandet, till exempel genom att illustrera ett ord eller en term i frågan. De kan även fylla mer perifera syften som att motivera respondenten, göra ifyllandet lättsammare eller förstärka undersökningsorganisationens image. Dessa syften är separata för frågekonstruktören men inte nödvändigtvis för respondenten (Couper, Tourangeau & Kenyon, 2004).

Bilder kan variera i komplexitet. Ju mer komplex en bild är, desto fler tolkningar kan den leda till. En bild säger mer än tusen ord sägs det ju, och det är inte säkert att alla dessa ord överensstämmer med det som ska mätas. Couper, Conrad och Tourangeau (2007) resonerar att bilder kan leda till liknande kontexteffekter som tidigare frågor i enkäten kan ha på senare frågor.

Exempel på bilder och förståelse av frågan

En studie undersökte hur bilder påverkar förståelsen av frågan. I studien graderade respondenterna sin hälsa som bättre om frågan inkluderade en bild på en sjuk kvinna i en sjukhussäng och sämre om frågan inkluderade en joggande kvinna. Frågan inkluderade vaga kvantifieringar som svarsalternativ. De har inte någon exakt betydelse utan kräver en referenspunkt, till exempel en svarsskala som går från "Mycket bra" till "Mycket dålig" (läs mer i kapitel 3). Sannolikt påverkar bilderna referenspunkterna för vad som är god respektive dålig hälsa och därmed också svaren. Bilden på den joggande kvinnan kan sägas höja kraven för vad som är god hälsa. Bilden på den sjuka kvinnan sänker dem (Couper, Conrad & Tourangeau, 2007).

Det kan alltså vara riskfyllt att använda bilder, speciellt om det inte finns ett givet eller genomtänkt syfte med bilden. De kan till exempel medföra att respondenten avgränsar ett koncept eller sin referenspunkt på

oönskade sätt. Det kan leda till att aktiviteter som borde ha ingått i svaret inte räknas med eller tvärtom. Detta kan bli extra ödesdigert vid kombinerad insamling ("mixed-mode" på engelska) om webbvarianten har bilder, för att det är så lätt och billigt att lägga till bilder på webben, men inte i pappersenkäten. Det finns dessutom inga indikationer på att bilder skulle höja motivationen hos respondenterna. I studien som nämndes ovan var respondenterna lika nöjda eller missnöjda, oavsett bild. De uppskattade tiden det tog att fylla i enkäten på samma sätt som respondenter som besvarat enkäten utan bilder (Couper, Tourangeau & Kenyon, 2004; Toepoel & Couper, 2011).

Respondenterna kan alltså använda information i bilderna när de besvarar frågorna. Hur tungt väger sådan information jämfört med verbala instruktioner? En studie (Toepoel & Couper, 2011) visade att både visuell och verbal information har oberoende- och interaktionseffekter, men att den verbala informationen är starkare. Verbal information verkar alltså vara viktigare än visuell. Eventuell påverkan av bilder borde därmed minskas med tydliga frågor (Toepoel & Couper, 2011).

Utifrån den här forskningen är rekommendationen därför att vara försiktig i användningen av bilder, speciellt i anslutning till frågorna, om man inte har ett genomtänkt syfte med bilderna (t.ex. att en bild behövs för att illustrera något). Om man vill ha bilder i enkäten men inte vill att bilderna i sig ska påverka mätningen kan det vara bra att använda bilder som ligger långt ifrån frågeområdet.

Sammanfattande punkter:

- Bilder ses ibland som något som kan "lätta upp stämningen" i en enkät.
- Bilder kan dock påverka svaren genom att begränsa respondentens förståelse eller bidra med en referenspunkt.
- Bilder bör därför användas med försiktighet i enkäter.

Layout i webbenkäter

Webbenkätens möjligheter

Webbenkäter blir vanligare och vanligare. Ett skäl är att de är billigare att använda än pappersenkäter eftersom det inte blir några tryck- eller portokostnader. Ett annat skäl är att webbenkäten som datoriserad metod erbjuder många andra möjligheter än pappersenkäten. Det är på grund av den egenskapen som webbenkäter har fått ett eget avsnitt i detta kapitel.

Pappersenkäter är fasta och givna till formen i och med att frågor och svarsalternativ är förtryckta. Webbenkäter kan däremot vara interaktiva. De kan programmeras så att de påverkas av vad respondenten gör. Utifrån det perspektivet kan webbenkäten ses som ett mellanting mellan pappersenkäten och telefonintervjun. Precis som pappersenkäten är webbenkäten självadministrerad och använder skriftlig information, men den kan vara interaktiv precis som intervjun. Denna interaktivitet kan användas på många olika sätt och begränsas egentligen endast av programvaran. Exempel på interaktiva funktioner är (Couper, 2008):

- granskningskontroller där respondenten kan upplysas om att ett svar är felaktigt eller misstänkt felaktigt
- automatiska summeringar och beräkningar
- automatiska hopp där nästa aktuella fråga automatiskt kommer upp på skärmen när respondenten trycker på knappen "Nästa/Fortsätt"
- dynamiska loopar som ger möjlighet att ställa samma frågor om till exempel olika personer där antalet personer till en början är okänt
- instruktioner som endast blir synliga när de behövs.

Det är lättare att arbeta med färg, form och format i webbenkäten än i pappersenkäten, där färgtryck till exempel är dyrt och sällan används. Det finns därför mer utrymme för layout i webbenkäten än i pappersenkäten.

Webbenkäten – design på sin spets

Det finns många beslut att ta och besluten kan ha konsekvenser. Det gäller därför att vara medveten om vilka möjligheter som finns. Utifrån styrkor och svagheter får man göra en bedömning om vad som är mest lämpligt i olika fall. Exempelvis kan man ju ha granskningskontroller vid varje fråga i syfte att förbättra kvaliteten på insamlad data, men man måste också ta hänsyn till respondentbördan.

Det är dessutom viktigt att påminna sig själv om att design inte främst handlar om att få utlopp för sin kreativitet, utan om att underlätta för respondenten att utföra sin uppgift (Couper, 2008).

Det behöver också tas hänsyn till att vi nu befinner oss i vad som kan vara en övergångsperiod från pappersenkät till webbenkät. Detta medför att pappersenkäter och webbenkäter ofta används parallellt i samma undersökning. I sådana fall bör man vara extra medveten om hur webbenkätens unika funktionalitet kan påverka svaren, eftersom pappersenkäten inte har någon motsvarighet.

Sammanfattande punkter:

- Webbenkäten har en del unika möjligheter jämfört med pappersenkäten.
- Webbenkäten kan vara interaktiv genom att påverkas av vad respondenten gör. På det sättet delar webbenkäten många egenskaper med telefonintervjun.
- Exempel på möjligheter i webbenkäten: automatiska hopp, beräkningar och summeringar, granskningskontroller och instruktioner "vid behov".
- Bara för att något *kan* göras betyder det inte att det *bör* göras – designen ska syfta till att förbättra mätningen och minska respondentbördan.

Övergripande upplägg – rullning eller sida-för-sida?

I en *scrollversion*, eller *rullversion* som det också kallas, av enkäten är det lätt för respondenten att gå fram och tillbaka bland frågorna. Det är bara att föra markören till höger på sidan upp eller ner eller att scrolla på mushjulet. Alla frågor finns på en och samma sida.

I upplägget *sida-för-sida* ("paging" på engelska) består webbenkäten av separata sidor med några frågor på varje sida. Respondenten tar sig framåt i enkäten genom att trycka på en fortsätt-knapp. Då laddas en ny sida.

Vad är då fördelarna respektive nackdelarna med rullning och sida-för-sida? Couper (2008) skriver att en fördel med scrollversionen är att det är lätt att hoppa fram och tillbaka i enkäten. Respondenten har också fler frågor synliga på skärmen samtidigt. Detta kan vara en styrka om det är

positivt att respondenten får en överblick. Kanske skulle respondenterna ha svarat annorlunda på filterfrågan om de hade kunnat se följdfrågorna? I sida-för-sida-versionen av en webbenkät kan det ibland vara ett problem att respondenten inte ser följdfrågorna efter en filterfråga. På motsatt sätt kan sida-för-sida vara att föredra när man *inte* vill att respondenten ska se efterkommande frågor, till exempel när ordningsföljden är viktig (Couper, 2008).

Filterfrågor fungerar inte särskilt bra i scrollversionen. Det går att programmera bokmärken som respondenten kan hoppa till, men om respondenten scollar försvinner kopplingen mellan filterfråga och målfråga (dit man ska "hoppa"). Denna risk finns inte om sida-för-sida används, vilken därför kan vara ett föredra om enkäten har många eller viktiga filter (Couper, 2008).

Vissa enkäter har mycket innehåll som ska laddas, till exempel många frågor, grafik eller bilder. Då är sida-för-sida att föredra, eftersom allt material laddas ner på en gång i scrollversionen. Beroende på mängd och respondentens uppkoppling kan det ta tid (Couper, 2008).

En annan viktig faktor att ta hänsyn till är *när* respondentens svar ska sparas. Detta är viktigt till exempel om ifyllandet avbryts av respondenten själv eller på grund av uppkopplingsproblem. I sida-för-sida sparas ofta svaren temporärt varje gång respondenten byter sida. Detta gör att respondenten kan återuppta ifyllandet utan att behöva besvara samma frågor en gång till. I scrollversionen finns ingen motsvarighet till detta och risken är större för förlorad data (Couper, 2008).

Enligt ovanstående resonemang har upplägget sida-för-sida fler fördelar än nackdelar och är numera mycket vanligare.

Sammanfattande punkter:

- Scrollning: bra om respondenten ska få en överblick, lätt att hoppa fram och tillbaka i enkäten.
- Sida-för-sida: bra om respondenten inte bör "läsa i förväg", bra för filterfrågor och har kortare laddningstid.

Format för webbenkät

Svarsfält

Sällan blir skillnaderna mellan pappers- och webbenkät så tydliga som just vid valet av svarsfält. I pappersenkäter används oftast kryssrutor. I webbenkäter finns det flera alternativ som kan används vid olika tillfällen. Gemensamt för de flesta är att svarsfält ska placeras före texten i svarsalternativet. Det blir då en rak kolumn med svarsfält, vilket gör det lättare för respondenten att fylla i sitt svar.

Radioknappar

Radioknappar ("radio buttons" på engelska) är runda cirklar som respondenten kan markera. Radioknappen har den specifika egenskapen att endast en knapp i en kolumn av knappar kan vara markerad samtidigt. Om respondenten markerar ett nytt svarsalternativ så försvinner den förra markeringen. Ett problem förr var att det inte gick att avmarkera radioknappar, det gick inte ångra sig och lämna frågan obesvarad när man väl fyllt i ett alternativ. Nuförtiden kan dock radioknappar avmarkeras genom att markera dem ytterligare en gång. Någon motsvarighet till radioknappar finns förstas inte i pappersenkäter. Radioknappar är lämpliga att använda i frågor där respondenten får ange *endast ett* svarsalternativ. Ett svarsalternativ får aldrig vara markerat i förväg utan radioknapparna måste vara blanka initialt. Annars går det inte att veta om respondenten har svarat eller hoppat över frågan (Couper, 2008).

Kryssrutor

Webbenkätens kryssrutor är det svarsfält som mest påminner om pappersenkätens svarsfält. Kryssrutor har ingen begränsning i antalet markeringar utan respondenten kan fylla i flera. Detta gör att kryssrutan är ett lämpligt svarsfält vid frågor där *fler än ett* svarsalternativ kan vara giltiga, exempelvis vid checklistor (Couper, 2008).

Ett relativt vanligt problem är att webbenkäter har kryssrutor istället för radioknappar, trots att endast ett svarsalternativ är giltigt. Detta kan bero på att enkäten överförts från papper till webb utan att man begrundat vilket svarsfält som är lämpligast. Om den som skapar webbenkäten inte är så insatt i enkätmetodik kan den typen av misstag lätt ske (kryssrutan liknar ju pappersenkätens svarsfält till formen). Om syftet är att webbenkäten så långt som möjligt ska efterlikna pappersenkäten bör de

visserligen ha samma typ av svarsfält, men då drar man inte nytta av webbens förbättringsmöjligheter (Couper, 2008).

Rullgardinsmenyer

En annan funktion webben erbjuder är rullgardinsmenyn där de olika svarsalternativen blir synliga först när menyn aktiveras. Då "rullas" de ner på skärmen. Rullgardinsmenyn har en del fördelar. Den främsta är att den sparar utrymme på skärmen. En annan är att rullgardinsmenyn kan innehålla fler svarsalternativ än vad som annars hade varit möjligt.

Men hur förhåller sig rullgardinsmenyn vad gäller ordningseffekter i jämförelse med en mer traditionell uppställning? I rullgardinsmenyer tillkommer två faktorer, förutom ordningsföljd och antalet svarsalternativ: 1) extra arbete med datormusen och 2) det visuella, i vilken grad svarsalternativen är synliga. I webbenkäternas begynnelse hade rullgardinsmenyer ibland ett defaultvärde som alltid var synligt. Detta är ingen bra lösning eftersom de som inte svarar eller aktiverar rullgardinsmenyn överhuvudtaget då tillskrivs detta defaultvärde trots att det egentligen är partiellt bortfall. Men även i rullgardinsmenyer utan defaultvärde kan synligheten variera – vissa alternativ är synliga på en gång medan andra måste scrollas fram för att de ligger längre ner i menyn. Har då synligheten någon betydelse? En studie hittade störst ordningseffekt ("primacy effect") i varianten där vissa svarsalternativ inte är synliga utan scrollas fram. Den här typen av svarsformat bör därför undvikas om det inte gäller väldigt välkända svarsalternativ, som exempelvis födelseland (Couper, Tourangeau & Conrad, 2004).

Man bör också ha i åtanke att rullgardinsmenyn innebär ett extra moment som respondenten behöver genomföra. Vissa kan ha svårare med detta, exempelvis äldre personer. Om listan inte är för lång går dock rullgardinsmenyn fortare än att fylla i svaret manuellt (Couper, Kennedy, Conrad & Tourangeau, 2011).

Vid väldigt långa listor kan en sökfunktion vara att föredra framför en rullgardinsmeny.

Textfält

Det kan finnas olika typer av textfält. En typ tillåter ett väldigt begränsat antal tecken knutet till frågans svarsenhet, exempelvis tre om svarsenheten är procent. En annan typ tillåter längre svar i fritext.

Respondenterna kan generellt använda formatet för att få information om vilken typ av svar som förväntas. Detta är även värt att tänka på när det gäller textfält. Vill man att respondenterna ska skriva utförliga fria svar bör textfältets storlek signalera detta. På motsvarande vis bör inte textfält heller signalera att respondenten ska skriva mer än vad som är önskvärt. Det ska inte finnas fler tecken än vad respondentens svar kan förväntas innehålla, annars kan det leda till fel eftersom det ger respondenten utrymme att hitta på egna lösningar (och till exempel svara "nästan 300"). En studie fann att ett stort textfält ledde till annan rapportering, jämfört med om en mindre textfält användes (Couper, Traugott & Lamias, 2001). Att hjälpa respondenten genom att dela upp fältet i "år/månad/dag" förbättrar också svaren jämfört med om respondenten ska fylla i allt i samma fält (Couper, Kennedy, Conrad & Tourangeau, 2011).

Textfältet bör därför vara anpassad efter format och det önskade svarets längd. Även vid textfält för fri text kan storleken på textfältet signalera hur pass utförlig respondenten förväntas vara. Textfältets storlek bör därför anpassas efter detta (Smyth, Dillman, Christian & McBride, 2009).

Sammanfattande punkter:

- Radioknappar är ett svarsfält som endast kan markeras ett åt gången.
- Kryssrutor påminner om svarsrutorna i pappersenkäter.
- Rullgardinsmenyn innebär möjligheter att ha mycket information i ett fält men också risker vad gäller synlighet och respondentens ansträngning.
- Textfält kan signalera eller styra hur mycket eller lite information som förväntas skrivas och bör därför anpassas efter detta.
- Vilket svarsfält som är mest lämpligt beror på frågetyp.

Annan webblayout och funktionalitet**Automatiska hopp**

Automatiska hopp används vid filterfrågor. Det innebär att hoppet är programmerat i webbenkäten och inte kräver att respondenten sköter hoppen på egen hand. Detta är generellt sett det bästa sättet att hantera

filter i webbenkäten och är en stor styrka jämfört med pappersenkätens hoppanvisningar.

Det är dock viktigt att tänka på att följdfrågorna inte får samma betydelse som i pappersenkäten. I pappersenkäten kan de till och med vara på samma sida som filterfrågan och det är då lätt för respondenten att bedöma vad filtret handlar om. I webbenkäten syns inte följdfrågorna till ett filter förrän respondenterna redan besvarat filterfrågan. De måste alltså besvara filterfrågan utan den "hjälpinformationen". Det kan få konsekvenser. Exempelvis att vissa som svarade "Nej" skulle ha svarat "Ja" om de hade sett följdfrågorna och fått en fördjupad förståelse för vad filtret handlade om.

Därför är det viktigt att kontrollera att filterfrågan ger tillräckligt med information för att respondenten ska kunna göra en korrekt bedömning.

Antal frågor på skärmen

Hur många frågor ska man ha per sida i ett sida-för-sida upplägg? Det finns en fördel om alla frågor på skärmen är synliga direkt och att respondenten därför inte behöver scrolla (det är ju en poäng med upplägget jämfört med en ren scrollversion av enkäten). Det bidrar med en naturlig avgränsning i hur många frågor som får plats på en skärmvisning. Utöver detta kan man generellt säga att många frågor på skärmen visserligen minskar tidsåtgången för besvarandet men det ökar det partiella bortfallet och försämrar respondenternas upplevelse av enkäten (Toepoel, Das & van Soest, 2009).

Sidbrytning

Generellt får det plats färre frågor på skärmen än på en papperssida. Detta gör att sidbrytningen blir annorlunda i en webbenkät jämfört med en pappersenkät. Man bör därför tänka på var sidbrytningarna bäst bör placeras. Ett synsätt är att betrakta varje sida i webbenkäten som ett sammanhang. Att bryta ut en eller några frågor ur dess sammanhang kan medföra att de uppfattas annorlunda än i motsvarande pappersenkät. Samtidigt, att ha för många frågor på sidan har också nackdelar (se ovan). Var sidbrytningarna bäst ska ligga är därför en bedömningsfråga som kräver lite eftertanke. Det gäller speciellt i undersökningar som använder kombinerad insamling (mixed-mode).

Skriva ut eller spara ner

Beroende på vad det är för typ av enkät kan respondenterna ges möjlighet att skriva ut eller spara ner en tom enkät för att kunna hämta information innan uppgifterna ska lämnas i webbenkäten. Detta är framförallt relevant för enkäter riktade till företag och organisationer. En sådan utskrift blir bäst för enkäter som inte är dynamiska, det vill säga som har en fast struktur och inte påverkas av bakgrundsinformation eller innehåller hopp eller loopar.

Numrering

I pappersenkäter fyller frågenumerering en viktig funktion för strukturen, speciellt vid hoppanvisningar. I en webbenkät med scrollning fyller den samma funktion. När sida-för-sida används försvinner dock den funktionen och frågenumerering är inte lika nödvändigt. Numrering kan till och med ha en förvirrande effekt när enkäten använder automatiska hopp. Eftersom hoppen är programmerade och inte redovisas för respondenten kan denne reagera på att frågenumren inte kommer i naturlig ordning. Respondenten kan uppfatta det som att någon fråga har missats. Samma gäller för sidnumrering. Numrering rekommenderas alltså inte vid webbenkäter.

Instruktioner

Läs mer i kapitel 4.

Förloppsindikator ("progress indicator" på engelska)

Rekommendationen är att endast använda en övergripande (mindre exakt) förloppsindikator för enkäter som inte är dynamiska. Om enkäten är dynamisk ger inte förloppsindikatorn en sann bild av enkätens omfattning (läs mer i kapitel 4).

Rullningslist ("scroll list" på engelska)

Vertikal rullningslist

Vertikal rullningslist är den list som oftast finns på högersidan av skärmen där man kan scrolla upp eller ned genom att dra med muspekaren. En vertikal rullningslist är vanligt förekommande till exempel i undersökningar som innehåller stora tabeller. Tänk dock på att en tabell är lättare att förstå och navigera i om det går att se hela tabellen på skärmbilden. Respondenten kan också ha behov av att förstå

tabellen för att öka läsbarheten. Tänk också på att kolumnrubrikerna försvinner ur bild om tabellen är lång. Därför bör det i första hand undersökas om en stor tabell kan delas upp i flera mindre. I vissa fall kräver sammanhanget stora tabeller och rullningslist blir då nödvändig.

Horisontell rullningslist

Undvik att använda en horisontell rullningslist (den list som kan finnas längst ner på skärmen och där man kan scrolla till höger eller vänster). Risken är stor att respondenten inte upptäcker den och missar delar av enkäten. Det är även lätt att förlora överblicken och det kan uppfattas som irriterande att behöva scrolla åt sidan. Om tabellen är för bred kan det vara bättre att dela upp den i flera mindre. Om det är en befintlig tabell hämtad från en pappersenkät kan det ibland räcka med att vrida tabellen så att kolumner blir rader och rader blir kolumner för att undvika en horisontell rullningslist.

Sammanfattande punkter:

- Automatiska hopp används vid filterfrågor. Det innebär att hoppet är programmerat i webbenkäten och inte kräver att respondenten sköter hoppet på egen hand. Det är viktigt att kontrollera att filterfrågan ger tillräckligt med information för att respondenten ska kunna göra en korrekt bedömning.
- Det finns många fördelar med att alla frågor på skärmen är synliga direkt och att respondenten därför inte behöver scrolla.
- Generellt får det plats färre frågor på skärmen än på en papperssida. Detta gör att sidbrytningen blir annorlunda i en webbenkät jämfört med en pappersenkät. Att bryta ut en eller några frågor ur dess sammanhang kan medföra att de uppfattas annorlunda än i motsvarande pappersenkät.
- Beroende på vad det är för typ av enkät kan respondenterna ges möjlighet att skriva ut eller spara ner en tom enkät för att kunna hämta information innan uppgifterna ska lämnas. Detta är framförallt relevant för enkäter riktade till företag och organisationer.

- I pappersenkäter fyller frågenummering en viktig funktion för strukturen, speciellt vid hoppanvisningar. I en webbenkät med scrollning fyller den samma funktion. När sida-för-sida används försvinner dock den funktionen och frågenummering är inte lika nödvändigt.
- Vertikal rullningslist är vanligt förekommande till exempel i undersökningar som innehåller stora tabeller. Tänk dock på att en tabell är lättare att förstå och navigera i om det går att se hela tabellen på skärmbilden.
- Undvik att använda en horisontell rullningslist. Risken är stor att respondenten inte upptäcker den och missar delar av enkäten.

Felmeddelanden

I undersökningar med uppgiftslämnarskyldighet eller i undersökningar där respondenten av andra skäl är extra motiverad att lämna korrekta uppgifter kan det finnas flera fördelar med att granska de ifyllda svaren i webbenkäten direkt vid besvarandet. Uppenbara fel som slarvfel och avvikelselfel i form av för stort eller för litet värde kan då omedelbart korrigeras av respondenten. Behovet av efterbearbetning kan då minskas väsentligt. Ifall respondenten har lämnat ett svar som är felaktigt eller misstänkt felaktigt bör felmeddelandet komma upp i anslutning till det givna svaret. Vid frivilliga undersökningar ska man generellt sett vara mycket försiktig med felmeddelanden, eftersom det ökar bördan på en uppgift som respondenten utför frivilligt.

Dessa kontroller i enkäten bör ha hög träffsäkerhet, vilket betyder att när en kontroll faller ut ska det vara stor sannolikhet att något är fel. Kontrollerna ska kännas meningsfulla och fylla en funktion för respondenten. Det kan vara bra att fokusera på vanligt förekommande fel eller sådant som har stora konsekvenser för undersökningen.

Ett felmeddelande ska vara trevligt formulerat och lätt att förstå. Det måste klart och tydligt framgå vad det är som är felaktigt eller misstänkt felaktigt. Det kan vara frustrerande att få ett felmeddelande men inte förstå hur felet ska kunna lösas. Undvik därför i möjligaste mån att använda tekniska beskrivningar, förkortningar eller sifferkoder i felmeddelandet.

Felmeddelandet bör även tala om ifall det är en så kallad hård eller mjuk granskningskontroll som utförts. Då vet respondenten ifall svaret måste

ändras för att det *är* felaktigt (hård), eller om det *kan vara* felaktigt (mjuk). Detta är främst relevant vid undersökningar där respondenterna är skyldiga att lämna uppgifterna. I sådana undersökningar kan det också ligga i respondentens intresse att göra rätt från början jämfört med att bli återkontaktad senare.

Eftersom hårda kontroller innebär att ett svar är felaktigt bör de vara tvingande, det vill säga att felet *måste* åtgärdas innan enkäten skickas in. Använd därför så få hårda kontroller som möjligt.

Mjuka kontroller är informativa men ska inte behöva åtgärdas för att enkäten ska kunna skickas in. Syftet är att informera respondenten om att något värde är avvikande och öka förståelsen av frågan. De mjuka kontrollerna ska respondenten kunna klicka sig förbi. De eventuella fel som fallit ut bör i kontrollen följas av ett kommentarsfält där respondenterna kan kommentera sina svar.

Sammanfattande punkter:

- Felmeddelanden kan leda till mer korrekta svar men bör bara användas om respondenten upplever det som något positivt eller om undersökningen har uppgiftslämnarskyldighet.
- Felmeddelanden kan vara mjuka (upplysande) eller hårda (tvingande).
- Felmeddelanden ska förklara vad som är fel och hur det kan lösas.

Färger

Färg är ett viktigt visuellt verktyg. Färgen kan påverka kontrasten mellan bakgrunden och förgrunden. Färg kan dock också påverka hur saker och ting förstås. Färger bör därför användas med eftertänksamhet. Man bör även ta hänsyn till att en del personer är färgblinda. Den vanligaste formen av färgblindhet är att inte kunna se skillnad på rött och grönt. Delvis på grund av den anledningen ska förståelsen av enkäten inte vara beroende av färgsättningen utan endast förstärka den.

Vissa färger har fått olika funktioner, speciellt i webbenkäter:

- **Blå text** betyder klickbar länk.
- **Röd text** betyder fel eller varningar.

Givetvis bör man utnyttja och inte gå emot sådant som respondenterna förväntar sig.

Sammanfattning

Layout – sammanfattning

Att kommunicera med grafik och symboler:

- Vi kommunicerar inte bara med språket utan även med grafik och symboler – layout.
- Layout har stor betydelse i enkäter. Den ökar respondentens uppmärksamhet, förståelse för enskilda frågor och visar vägen genom enkäten.

Perception – hur vi förstår information från våra sinnen:

- Perception är den process när information från sinnesorganen ges mening.
- Perceptionen är bland annat styrd av förkunskaper – vi förstår inte allting på nytt varje gång det uppträder.
- Gestaltlagarna är en samling lagar för hur människor sorterar och förstår visuella stimuli, till exempel lagen om likhet = att objekt som ser likadana ut förväntas höra ihop och betyda samma sak.

Allmänt om design:

- Kognitiva aspekter av allmän design: använd kända kunskaper, gör saker synliga och använd begränsningar.
- Om det är svårt att få ovanstående att fungera – standardisera.

Layout i enkäter (pappers- och webbenkät):

- De verktyg som finns till förfogande i layout är: text, font, symboler, bilder, siffror, linjer och färger. Ett viktigt beslut är hur dessa bäst ska placeras på den begränsade yta som finns i enkäten.
- Se punktlista i texten för korta regler för layout i enkäter.

Layout för frågor:

- Tydliggör i layouten vad som är inledande text, frågeformulering och instruktion.

- Använd vänsterställd text.
- Radlängden på en lång fråga får varken vara för kort eller för lång. Maxlängden 60–70 tecken inklusive mellanslag är rekommendationen.
- Använd olika avstånd för att visa vad som hör ihop eller inte.
- Avgränsa innehåll med hjälp av linjer, ramar eller gråskala.
- Koppla ihop huvud- och delfråga t.ex. med tre punkter ("...").
- Använd främst fet stil för att förstärka ord i frågan.
- Gör gömda frågor synliga.
- Se bildexempel i texten.

Layout för svarsalternativ:

- Det säkraste är att använda svarsfält som respondenterna känner till sedan tidigare. Det innebär till exempel kryssrutor i en pappers- eller webbenkät eller radioknappar i en webbenkät.
- Lyft fram svarsrutorna i layouten.
- Svarsrutorna brukar vanligtvis placeras till vänster om respektive svarsalternativ.
- Svarsalternativen skrivs under varandra med samma avstånd med undantag av så kallade icke-svar. Icke-svar som "Vet inte", "Inte aktuellt" eller "Ingen åsikt" ska stå en bit ifrån övriga för att markera att de inte hör till det förväntade svaret.
- Hoppanvisningar kan leda till fel och bör därför vara tydligt konstruerade (se exempel).

Bilder i enkäter:

- Bilder ses ibland som något som kan "lätta upp stämningen" i en enkät.
- Bilder kan dock påverka svaren genom att begränsa respondentens förståelse eller bidra med en referenspunkt.
- Bilder bör därför användas med försiktighet i enkäter.

Layout i webbenkäter:

- Webbenkäten har en del unika möjligheter jämfört med pappersenkäten.

- Webbenkäten kan vara interaktiv genom att påverkas av vad respondenten gör. På det sättet delar webbenkäten många egenskaper med telefonintervjun.
- Exempel på möjligheter i webbenkäten: automatiska hopp, beräkningar och summeringar, granskningskontroller och instruktioner "vid behov".
- Bara för att något *kan* göras betyder det inte att det *bör* göras – designen ska syfta till att förbättra mätningen och minska respondentbördan.

Övergripande upplägg - rullning eller sida-för-sida?

- Scrollning: bra om respondenten ska få en överblick, lätt att hoppa fram och tillbaka i enkäten.
- Sida-för-sida: bra om respondenten inte bör "läsa i förväg", bra för filterfrågor och har kortare laddningstid.

Format för webbenkät:

- Radioknappar är ett svarsfält som endast kan markeras ett åt gången.
- Krysrutor påminner om svarsrutorna i pappersenkäter.
- Rullgardinsmenyn innebär möjligheter att ha mycket information i ett fält men också risker vad gäller synlighet och respondentens ansträngning.
- Textfält kan signalera eller styra hur mycket eller lite information som förväntas skrivas och bör därför anpassas efter detta.
- Vilket svarsfält som är mest lämpligt beror på frågetyp.

Annan webblayout och funktionalitet:

- Automatiska hopp används vid filterfrågor. Det innebär att hoppet är programmerat i webbenkäten och inte kräver att respondenten sköter hoppen på egen hand. Det är viktigt att kontrollera att filterfrågan ger tillräckligt med information för att respondenten ska kunna göra en korrekt bedömning.
- Det finns många fördelar med att alla frågor på skärmen är synliga direkt och att respondenten därför inte behöver scrolla.

- Generellt får det plats färre frågor på skärmen än på en papperssida. Detta gör att sidbrytningen blir annorlunda i en webbenkät jämfört med en pappersenkät. Att bryta ut en eller några frågor ur dess sammanhang kan medföra att de uppfattas annorlunda än i motsvarande pappersenkät.
- Beroende på vad det är för typ av enkät kan respondenterna ges möjlighet att skriva ut eller spara ner en tom enkät för att kunna hämta information innan uppgifterna ska lämnas. Detta är framförallt relevant för enkäter riktade till företag och organisationer.
- I pappersenkäter fyller frågenummering en viktig funktion för strukturen, speciellt vid hoppanvisningar. I en webbenkät med scrollning fyller den samma funktion. När sida-för-sida används försvinner dock den funktionen och frågenummering är inte lika nödvändigt.
- Vertikal rullningslist är vanligt förekommande till exempel i undersökningar som innehåller stora tabeller. Tänk dock på att en tabell är lättare att förstå och navigera i om det går att se hela tabellen på skärmbilden.
- Undvik att använda en horisontell rullningslist. Risken är stor att respondenten inte upptäcker den och missar delar av enkäten.

Felmeddelanden:

- Felmeddelanden kan leda till mer korrekta svar men bör bara användas om respondenten upplever det som något positivt eller om undersökningen har uppgiftslämnarskyldighet.
- Felmeddelanden kan vara mjuka (upplysande) eller hårda (tvingande).
- Felmeddelanden ska förklara vad som är fel och hur det kan lösas.

6 Datainsamlingsmetod

Författare: Andreas Persson (SCB)

Olika datainsamlingsmetoder

Det finns många olika sätt att samla in data på och fler tillkommer allt eftersom tekniken utvecklas. Webbenkäter är den senaste metoden som blivit ett etablerat alternativ på bred front och nu för tiden besvaras de dessutom allt oftare via surfplattor och smarttelefoner.

Olika datainsamlingsmetoder har olika styrkor och svagheter. De varierar angående till exempel vem som administrerar frågorna, vilken informationstyp som används, vilket medium som används för kommunikation, hur respondenterna kontaktas, hur privata de är, hur mycket feedback de tillåter och vilken kontroll över datainsamlingen de möjliggör (Couper, 2011; Tourangeau, Rips & Rasinski, 2000). De varierar också i tidsåtgång, teknik, fältarbete och kostnad. Många av dessa variationer ligger utanför denna boks räckvidd. Nedan fokuseras på mätningen och vilken påverkan datainsamlingsmetoden kan ha på frågekonstruktionen.

Det är inte självklart hur man ska kategorisera datainsamlingsmetoder i surveyundersökningar. Metoder kan variera och kombineras på många olika sätt. En undersökning kan även innehålla flera av dem samtidigt (så kallad "mixed mode"). Nedan redovisas en traditionell uppdelning i intervjuer eller självadministrerade datainsamlingsätt (de Leeuw, 2008). Men det finns fler metoder och kombinationer än de som redovisas, till exempel självadministrerad enkät i intervjuares närvaro eller där respondenten "intervjuas" av en inspelad röst och väljer svarsalternativ med telefonens siffersats.

Besöksintervju: En intervjuare träffar respondenten i dess hem eller i neutral miljö efter att ha bestämt tid via telefonen först. Vanligtvis administrerar intervjuaren intervjuformuläret, läser frågorna och registrerar respondentens svar i ett datastöd. Eftersom metoden innebär ett fysiskt möte finns det även möjlighet att använda hjälpmedel. Det kan vara i form av svars kort med svarsalternativ som respondenterna kan välja mellan eller att en självadministrerad enkät används för vissa delar av intervjun.

Telefonintervju: Intervjuaren ringer respondenten. I en del undersökningar (till exempel SCB:s telefonintervjuer) har urvalspersonen i förväg fått ett introduktionsbrev som beskriver undersökningen skickat till sig via post. Vid telefonsamtalet ger intervjuaren en kort introduktion av undersökningen. Sedan administrerar intervjuaren intervjufrågeformuläret och läser frågorna. I de allra flesta fall registreras respondentens svar av intervjuaren i ett datastöd.

Självadministrerad pappersenkät: Respondenten får enkät och missiv hemskickat via post. Respondenten läser frågorna och svarsalternativen och markerar själv sina svar i enkäten. Enkäten skickas tillbaka via post.

Självadministrerad webbenkät: Respondenten får ett login till en webbplats skickat till sig via antingen vanlig post eller e-post (vid deltagande i till exempel en webbpanel). Respondenten loggar in på en webbplats, läser frågorna och svarsalternativen och markerar själv sina svar i enkäten. Svaren registreras i en databas.

De fyra ovanstående är de mest etablerade metoderna. Vilka metoder som är vanligast är något som förändras med tiden. Besöksintervjun används till exempel alltmer sällan i Sverige men är fortfarande vanlig i vissa andra länder. Webbpaneler har haft motsatt utveckling och blivit alltmer vanliga. Webbpaneler tas dock inte upp separat i denna bok eftersom de främst är en urvalsram, inte en datainsamlingsmetod, enligt denna boks synsätt (en webbpanel är ett register av e-postadresser till och övrig information om personer som uppgett att de är intresserade av att delta i webbundersökningar). Webbpaneler ingår därför i självadministrerade webbenkäter ovan. Detsamma gäller enkäter som besvaras på surfplattor eller i smarttelefoner, även om dessa fått ett eget kortare avsnitt i kapitlet.

Detta kapitel kommer att resonera utifrån de fyra vanligaste datainsamlingsmetoderna i listan ovan. Kapitlet innehåller också ett avsnitt om kombinerad insamling (så kallad "mixed mode"). Resonemang om konsekvenser av att blanda insamlingsmetoder är dock inte begränsat till det avsnittet utan genomsyrar hela kapitlet. Kapitlet avslutas med ett avsnitt med konkreta frågekonstruktionstips för de olika datainsamlingsmetoderna.

Sammanfattande punkter:

Det finns olika typer av datainsamlingsmetoder. De vanligaste är:

- självadministrerad enkät (papper eller webb)
- intervju (telefon eller besök).

Dessa skiljer sig åt på olika sätt.

Likheter och skillnader mellan olika datainsamlingsmetoder

Datainsamlingsmetoderna överlappar på flera olika sätt. Det mest naturliga är väl att som ovan dela upp dem i intervjuer och självadministrerade datainsamlingsätt. Men inte ens den grova indelningen är riktigt rättvisande. Eftersom besöksintervjun innebär ett fysiskt möte finns möjligheter att använda material som annars förknippas med de självadministrerade sätten (till exempel skriftlig information på svarskort eller separat enkät). På samma sätt kan webbenkäten sägas "administrera" frågornas ordningsföljd precis som en intervjuares datorstöd. Det kan till exempel ske genom att hopp i enkäten programmeras så att respondenten inte ens behöver bli varse om att frågor hoppats över. Det finns alltså likheter mellan metoderna på flera olika sätt. Redovisningen nedan fokuserar på två skillnader mellan metoder: det sociala sammanhanget och vilken informationstyp som används (ljud eller skrift) (de Leeuw, 2008).

Självadministrerat eller socialt samspel

När respondenten besvarar frågeformuläret på egen hand

I självadministrerade enkäter fyller respondenten, precis som det låter, i enkäten på egen hand. Detta är kanske en sanning med viss modifikation. Bergman och Wärneryd (1982) redovisar att i en undersökning var det 82 procent som fyllde i enkäten på egen hand utan att diskutera eller få råd av någon annan. 13 procent fyllde i själva men efter viss diskussion. Man får förmoda att diskussionen främst fördes med någon annan hushållsmedlem. I jämförelse med andra datainsamlingsmetoder är dock det sociala samspelet svagt i självadministrerade enkäter. Respondenten besvarar i hög grad enkäten själv. Detta har vissa konsekvenser. Eftersom respondenten själv sköter besvarandet så kan enkäten besvaras när det passar och till och med vid flera olika tillfällen.

Bergman och Wärneryd (1982) konstaterar att styra över besvarandet själv kan ha en kvalitetshöjande verkan.

Att respondenten fyller i på egen hand medför också att mätsituationen blir mer privat. Detta kan medföra att respondenterna svarar mer ärligt vid känsliga frågeområden (läs mer nedan). Den privata mätsituationen innebär också att respondenten, i den mån tillräckligt hög motivation finns, kan lägga mer tid på att formulera bra svar, till exempel genom att grundligare genomsöka minnet.

Vid självadministrerade enkäter tolkar och förstår respondenterna enkätfrågorna på egen hand. Respondenternas förmåga att tillgodogöra sig enkätens skriftliga information är därmed avgörande. Det finns ingen intervjuare som kan förklara, besvara motfrågor eller upptäcka och korrigera missförstånd. Omnibussundersökningen ovan (Bergman & Wärneryd, 1982) visar att respondenterna i vissa fall tar hjälp av andra men detta är förstås ingen garanti för rätt förståelse. I självadministrerade enkäter är vi alltid beroende av respondentens egen förståelse.

Intervjuaren kan hjälpa respondenten

I intervjuer har vi genom intervjuaren möjlighet att påverka respondenternas förståelse. Intervjuer möjliggör därför mer komplicerade frågor och upplägg. Svåra frågor är förstås något som bör undvikas oavsett datainsamlingsmetod, men vissa mätningar är av nödvändighet mer komplicerade än andra. Det kan till exempel gälla frågor där förklaringarna kan behöva anpassas efter respondenten eller undersökningar som kräver invecklade utredningar eller klassificeringar. Den typen av undersökningar, och de enskilda frågorna, är svåra att genomföra med självadministrerade datainsamlingsmetoder.

Styrkan med en intervjuare gäller inte bara frågornas innehåll och utformning utan även om frågeformuläret har en komplicerad struktur. Vissa frågeformulär har inte den traditionella formen att alla frågor ska besvaras i en följd utan avbrott, utan är mer av omfattande frågebatterier där olika respondenter kan ta väldigt olika rutter beroende på hur de svarar. Ett annat exempel är nätverksstudier där respondenten ska svara för flera olika personer. Komplicerade strukturer ställer krav på administrationen av frågorna. I intervjuundersökningar administrerar intervjuaren frågor med hjälp av ett datorstöd och då är komplicerade strukturer inget större problem – de programmeras in i datorstödet. I självadministrerade enkäter, där respondenten själv står för

administrationen, finns det risk för problem. Respondenterna kan uppfatta enkäten som rörig och avbryta sitt deltagande, besvara frågor som de borde hoppat över och hoppa över frågor de borde ha besvarat. Detta är dock en aspekt där det är en väsentlig skillnad mellan själv-administrerade pappers- och webbenkäter. Webbenkäter kan, precis som intervjuarnas datorstöd, programmeras och på så vis kontrollera administrationen.

Närheten till frågeställaren – skillnader mellan enkät och intervju

Närheten till frågeställaren (till exempel organisationen bakom undersökningen) varierar också mellan olika datainsamlingsmetoder. I en intervju får respondenten oftast ett introduktionsbrev per post men i övrigt är det intervjuaren som representerar frågeställaren. Det kan finnas ett värde i att skapa en viss distans mellan respondenternas upplevelse av intervjuaren och av frågeställaren, annars finns det en risk att respondenterna upplever att frågorna är intervjuarens egna och uttrycker hans eller hennes åsikter, vilket i sin tur kan påverka svaren.

I självadministrerade datainsamlingsätt begränsas frågeställarens utrymme till följebrevet (eller missiv som det brukar kallas för enkäter). I intervjuer finns alltså mer utrymme att beskriva undersökningsorganisationen, undersökningens bakgrund och syfte. Intervjuaren kan presentera detta och svara på eventuella frågor eller funderingar. Kanske är just detta anledningen till att telefonintervjuer traditionellt sett har haft högre svarsfrekvenser än självadministrerade datainsamlingsätt.

Ett sätt att minska det mänskliga avståndet i självadministrerade enkäter är att "personalisera" undersökningen (Bergman & Wärneryd, 1982). Detta kan innebära att man skriver under missivet med en riktig underskrift, tilltalar respondenten direkt ("Bäste Sven Svensson") och med ett mjukare och mer informellt språk, som om man skrev till en affärskollega (Dillman, 2007). Det finns en del forskning om effekten av personalisering på framförallt svarsfrekvensen och det är en integrerad del i Dillmans "tailored design" (Dillman, 2007). Ett SCB-experiment visade dock att en underskrift tryckt i blått, för att påminna om en signatur i bläck, inte hade någon effekt på svarsfrekvensen.

Det går alltså att personalisera på olika sätt. Det är viktigt att vara medveten om att vissa av dessa sätt kan stå i konflikt med andra mål, som känslan av anonymitet eller att undersökningsorganisationen ska uppfattas som seriös och förtroendeingivande (Bergman & Wärneryd,

1982; Groves, Cialdini & Couper, 1992; Wenemark, Persson, Noorlind-Brage, Svensson & Kristenson, 2010). Om undersökningen innehåller känsliga ämnen kan det därför finnas skäl att vara försiktig med för mycket personifiering, inte bara för svarsfrekvensens skull utan även för mätningens och respondenternas.

Sammanfattande punkter:

- Enkät: Respondenten kan besvara enkäten när det passar men vi är samtidigt beroende av att respondenten kan klara det på egen hand. Situationen är mer anonym vilket kan leda till mer ärliga svar men också att det är svårare att påverka respondenten, till exempel genom att beskriva undersökningen och öka motivationen.
- Intervju: Intervjuaren kan hjälpa respondenten med frågorna och sköter administrationen men besvarandet är inte privat. Intervjuaren kan motivera respondenten.

Intervju – ett socialt samspel

Den största skillnaden mellan datainsamlingsmetoderna vad gäller det sociala sammanhanget är att när en intervjuare är inblandad blir besvarandet ett samspel mellan två människor. Även om intervjun som situation har en del egna regler jämfört med vanlig konversation (att intervjuaren kontrollerar situationen och ställer frågorna) är den inte helt friskriven från social praxis. Det är, till syvende och sist, ett möte mellan två människor. Det är svårt att intervju i ett vakuum utan det är i grunden en feedback-situation. Till skillnad från i ett manus till en pjäs kan frågeställaren bara diktera intervjuarens repliker, inte respondentens. Enligt en studie utgjordes endast 37 procent av intervjuarnas beteende av frågeställande. 31 procent utgjordes av förtydliganden. Resten av intervjubeteendena bestod av svarskommentarer och vanlig konversation (Cannell, 1969; Bergman & Wärneryd, 1982). En intervju har alltså olika typer av innehåll.

Intervjuer kan dock variera i hur pass mycket feedback och avsteg de tillåter. De kan vara strikt standardiserade (inte tillåta avvikelser) eller mer flexibla (kallas flexibel eller konverserande intervju).

Två intervjutekniker – standardiserad eller konverserande

I standardiserade intervjuer är målet att alla respondenter ska få samma stimuli. Frågeformuleringarna ska vara låsta och intervjuarnas läsning ska inte avvika. Intervjuaren ska vara neutral under intervjun. Om respondenten ber om förtydligande eller dylikt ska intervjuaren endast uppmana respondenten att använda sin egen tolkning eller upprepa frågan igen, inte förklara för respondenten på egen hand. Målet med standardiserade intervjuer är att öka datakvaliteten genom att minska intervjuarvariansen (varians på grund av intervjuarnas beteende, till exempel att olika intervjuare ger olika förklaringar). Det är väldigt svårt att standardisera intervjuer fullt ut i praktiken även om man har den ambitionen. I praktiken är det vanligt att en intervju innehåller både standardiserade och konverserande inslag (Japac, 2005).

I dess rena form är den konverserande intervjun den standardiserade intervjus motsats. Utgångspunkten i den konverserande intervjun är idén att undersökningens koncept kan behöva förklaras för respondenterna mer utförligt än i standardiserade frågor och instruktioner och att det kan behövas olika förklaringar till olika respondenter. Conrad och Schober (2000) beskriver att båda intervjuformerna (standardiserad och konverserande) på sätt och vis har standardisering som mål. Den standardiserade intervjun syftar till att standardisera stimuli så att alla respondenter får exakt samma fråga. Den konverserande intervjun syftar till att standardisera *meningen*, innebörden, i frågorna. Synsätten skiljer sig alltså på hur de ser på ord och mening. Enligt den standardiserade intervjus synsätt är det säkrast att respondenterna exponeras för exakt samma ord för att undvika alternativa tolkningar. Den konverserande intervjus synsätt är snarare att en frågas mening kan framgå på olika sätt med olika ord och att det behöver anpassas efter respondentens förutsättningar (läs mer i kapitel 7).

Vilken intervjuform är då lämpligast? Detta kan givetvis bero på hur pass komplicerade undersökningens koncept är. En laboratoriestudie undersökte skillnader mellan standardiserad och konverserande intervju. Testpersonerna fick beskrivningar som de skulle svara utifrån (dessa beskrivningar gjorde det möjligt att undersöka om de insamlade svaren var korrekta eller inte). När frågorna var enkla var det ingen skillnad mellan intervjuteknikerna men vid komplexa frågor ökade den konverserande intervjutekniken svarens korrekthet med nästan 60 procent. En nackdel med den konverserande intervjutekniken är dock att

den tog nästan tre gånger så lång tid (Schober & Conrad, 1997). I komplicerade intervjusituationer kan alltså flexibilitet, det vill säga att intervjuarna kan förklara med egna ord, förbättra mätningen. En nackdel med konverserande intervjuer är dock att de kräver betydligt mer av intervjuarna, både vad gäller allmän intervjuteknik och kunskap om den specifika undersökningen. En annan nackdel är, som beskrivs ovan, att intervjutiden förlängs (Conrad & Schober, 2005). På grund av dessa nackdelar används sällan helt konverserande intervjuer i praktiken.

Svårt att standardisera fullt ut

Intervjuer kan ha alltså olika nivåer av standardisering. Det finns fördelar och nackdelar med graden av standardisering, beroende på syfte (se ovan). Men det är oerhört svårt att standardisera en intervju fullt ut. Dels kan intervjuaren som person påverka svaren (Bergman & Wärneryd, 1982).

Ovan beskrevs det som om frågeställaren kan diktera intervjuarens repliker i en intervju men inte ens detta är sant fullt ut, inte ens om undersökningen har som mål att vara fullt standardiserad. Det händer att intervjuare hoppar över instruktioner eller förändrar frågorna. Enligt en studie var det bara en tredjedel av frågorna som lästes korrekt, men det varierar förstås från undersökning till undersökning (Bradburn & Sudman, 1979; Bergman & Wärneryd, 1982). Detta motverkar standardiseringens syfte och kan leda till att svaren varierar på grund av intervjuaren, till exempel att de gett olika information eller läst frågan på olika sätt. Det kan leda till fel som kan vara systematiska, och leda till skevhet, eller osystematiska men öka variansen (läs mer om intervjuarvarians i Biemer & Lyberg, 2003; Bergman & Wärneryd, 1982). Det är dock viktigt att påpeka att läsavvikelse i sig inte behöver innebära fel utan tvärtom kan öka datakvaliteten (Conrad & Schober, 2005). Det beror på vad det är för slags avvikelse och hur den förhåller sig till frågans intention.

Skillnader i läsning och dylikt kan bero på oönskat intervjuarbeteende och kan minskas med utbildning. Men ofta beror det inte främst på intervjuaren utan på frågeformuläret. Frågorna kanske inte fungerar att läsa ordagrant i en intervjusituation. Situationer kan också uppstå där frågorna hamnar i konflikt med de konversationsregler som finns (läs mer i kapitel 7). Om respondenten tidigare spontant har gett information som utgör svaret på en viss fråga i frågeformuläret kanske intervjuaren

låter bli att ställa den frågan. Det bryter nämligen mot konversationsreglerna att fråga efter något som man uppenbart redan vet (regeln om kvantitet). Här finns det dock risk för bedömningsfel, till exempel att intervjuaren minns fel eller att informationen som respondenten gav kanske inte var svar på just den frågan.

Ett annat exempel är om respondenten inte förstår frågan. Om intervjun är helt standardiserad ska inte intervjuaren hjälpa respondenten i det läget utan endast uppmana respondenten att utgå från sin egen tolkning eller läsa om frågan (Conrad & Schober, 2005). Det bryter dock mot konversationsreglerna och kan uppfattas som oförskämt att återupprepa frågan ordagrant. Det kan vara mer naturligt för intervjuaren att agera som folk gör i andra situationer när motparten inte förstår – att man tillför någon ny information, till exempel genom att formulera frågan på ett annat sätt eller ge något exempel för att motparten ska förstå. I sådana situationer finns det förstas risk för variationer som bryter mot den ursprungliga intentionen. I viss mån kan ovanstående minskas med intervjuarinstruktioner och manus för hur de ska svara på vanliga motfrågor. Det viktigaste är dock att utarbeta ett bra frågeformulär och testa så att det fungerar i en intervjusituation. Men det är som sagt mycket svårt att standardisera en intervju fullt ut.

Läsavvikelser

Det finns en del forskning om vilka faktorer som kan öka risken för läsavvikelse. Vissa studier har funnit att öppna frågor leder till färre avvikelser än stängda. Andra studier har funnit motsatsen. En studie fann att kortare frågor var relaterat till färre läsavvikelser. I en annan studie (Presser & Zhao, 1992) undersöktes flera faktorer relation till läsavvikelse, bland annat frågornas längd. Författarna resonerade att läsavvikelsen borde vara högre för frågor som intervjuarna är mindre vana vid, till exempel frågor i ett ovanligt filter (som de flesta intervjuare därmed sällan läser). De resonerade också att läsavvikelserna borde vara fler i slutet av intervjun och att de borde vara vanligare för en serie av frågor i samma format. Resultatet innehöll inte så mycket läsavvikelse överhuvudtaget och de enda hypoteserna som var signifikant relaterade till läsavvikelse var frågans längd och om frågan ingick med flera i samma format (Presser & Zhao, 1992).

Även speciella intervjutekniker kan öka läsavvikelsen och intervjuarvariansen, men ha andra styrkor som mer korrekt rapportering. Hit

räknas till exempel kalenderintervju som utgår ifrån en kalender för att stödja respondentens minne (Sayles, Belli & Serrano, 2010) (läs mer i kapitel 11).

Sammanfattande punkter:

Det finns olika intervjutekniker:

- **Standardiserad intervju** följer frågeformuläret ord för ord och intervjuaren hjälper inte till.
- **Konverserande intervju** är mer i form av ett utredande samtal.

Båda teknikerna har sina för- och nackdelar. Den konverserande intervjun kan förbättra mätningen men kräver mycket mer resurser. Det är dock svårt att standardisera fullt ut. De flesta intervjuer innehåller i praktiken delar av de båda teknikerna.

Datainsamlingsmetoder och ärliga svar

Frågornas innehåll och att de ställs i ett socialt sammanhang kan, oavsett om de läses ordagrant eller inte, också hamna i konflikt med normer för hur människor vill presentera sig för en obekant motpart. Det är sällan vi avslöjar inför en främmande människa om vi till exempel begått brott. På liknande sätt kan människor klaga och vara negativa bland vänner men inför en främling vill många framstå som positiva och medgörliga personer. Dessa tendenser kan påverka svaren oavsett datainsamlingsmetod. Men det påverkar svaren i högre grad i intervjuer som sker i samspel med en annan människa (enkäter är mer privata men även där finns det en mänsklig mottagare av svaren). Exempelvis är hålla-med-tendensen (så kallad "acquiescence bias") starkare i telefonintervjuer än i självadministrerade enkäter (Ye, Fulton & Tourangeau, 2011; Biemer & Lyberg, 2003). Det kan vara enklare att hålla med än att ta avstånd ifrån något, speciellt om det sker inför en annan människa.

Personers sätt att presentera sig för en okänd motpart får dock störst utslag vid frågor som är känsliga till naturen. Hit räknas bland annat bruk av alkohol, droger och mediciner samt sexualliv, kriminella handlingar, sjukdomar och attityder om kontroversiella sakfrågor. Det varierar naturligtvis i viss mån från person till person vad som i praktiken uppfattas som känsligt. Att människor inte alltid är ärliga utan "förskönar" verkligheten när de besvarar frågor om känsliga ämnen

kallas för social önskvärdhet. Det är ett samlingsbegrepp för svars-tendensen att följa en norm snarare än att rapportera sanningen.

Känsliga frågor och social önskvärdhet behandlas utförligt i kapitel 2. Nedan tas det endast upp översiktligt och framförallt från perspektivet av datainsamlingsmetod. För att svara korrekt på känsliga frågor krävs det att respondenten är uppriktig både inför sig själv och inför en annan människa. Målsättningen med att svara örligt på känsliga frågor är framförallt att framstå i bättre dager men det kan också vara ett resultat av självförnekelse. Resultatet av sådana tendenser kan innebära under-rapportering av beteenden eller åsikter som anses negativa och över-rapportering av beteenden och åsikter som anses positiva. Vad som anses som positivt och negativt (eller önskvärt och icke-önskvärt) varierar både mellan personer och över tid. Men det finns trots allt många normer som de flesta människor delar. Normerna kan också vara av mer tillfällig eller specifik karaktär. De kan även påverkas av intervjuarens beteende, till exempel att intervjuaren reagerar negativt eller uppmuntrade på något (Bergman & Wärneryd, 1982).

Studier visar högre rapportering av droganvändning i självadministrerade enkäter än i intervjuer (Tourangeau, Rips & Rasinski, 2000). En studie fann att i en självadministrerad enkät var det sex procent som rapporterade att de använt kokain men bara fyra procent erkände det till en intervjuare (Biemer & Lyberg, 2003). Ett annat exempel visar signifikanta skillnader på frågan "Hur ofta kör du bil efter att ha druckit alkohol?" mellan pappersenkät ("Aldrig" = 52 %) och telefon ("Aldrig" = 63 %) (Dillman & Tarnai, 1991). En annan studie visade att i webbenkäter, jämfört med telefonintervjuer, ökade rapporteringen av känsliga upp-gifter och ledde till en bättre mätning enligt validerande data (Kreuter, Presser & Tourangeau, 2008). ACASI ("audio computer-assisted self-interviewing"), där respondenten intervjuas av en inspelad röst, är ett slags mellanting mellan intervju och enkät. Den ger högre rapportering av icke-önskvärda beteenden än en traditionell intervju. Dock har det visat sig att röstens egenskaper kan ha betydelse, åtminstone för män som rapporterade högre nivåer till en kvinnlig än till en manlig röst (Dykema, Diloreto, Price, White & Schaeffer, 2012). Även när det inte är en riktig människa som intervjuar utan en inspelad röst kan sociala effekter alltså förekomma.

Ett problem med att uppskatta social önskvärdhet i allmänhet och mellan olika datainsamlingsmetoder är avsaknaden av någon datakälla att

jämföra svaren med, ett "sant värde". Datainsamlingsmetoder kan skilja sig åt i svarsfördelning men vilken av dem ligger närmast sanningen? I litteraturen är ett vanligt antagande att högre andel av icke-önskvärda beteenden och lägre andel av önskvärda beteenden ligger närmast den sanna fördelningen. Detta är emellertid ett *antagande* eftersom validerande data ofta saknas (Kreuter, Presser & Tourangeau, 2007).

Sammanfattande punkter:

- Respondenterna svarar ibland oärligt p.g.a. självförnekelse eller för att de vill framstå i bättre dager (det senare kallas för social önskvärdhet).
- Denna tendens är svagare i datainsamlingsmetoder där respondenten besvarar frågeformuläret privat (pappers- och webbenkäter) jämfört med intervjuer.

Beror skillnaderna i svar på social närhet?

Det finns alltså en hel del forskning om social önskvärdhet och hur tendensen skiljer sig mellan olika datainsamlingsmetoder. En underliggande idé till skillnaderna i resultat är social närhet. Med det menas att ju närmare respondenten och intervjuaren är varandra och ju mer personlig intervjun är, desto större är riskerna för social önskvärdhet. Forskningen om självadministrerade insamlingssätt och telefonintervjuer bekräftar också den idén eftersom social önskvärdhet är högre i intervjuer än när respondenten besvarar frågeformuläret på egen hand (till exempel Chang & Krosnick, 2010). Idén blir dock problematisk om den appliceras på besöksintervjuer som innebär ett ännu kortare avstånd mellan respondent och intervjuare – ett fysiskt möte. Enligt idén om social närhet borde besöksintervjun vara mest utsatt för social önskvärdhet, men forskningen ger ingen sådan entydig bild. Telefonintervjuer verkar tvärtom vara förknippade med mer socialt önskvärda svar än besöksintervjuer (Holbrook, Green & Krosnick, 2003). Kanske beror det på att besöksintervjun ökar möjligheten att bygga upp ett förtroende så att respondenterna svarar ärligt. I besöksintervjun finns också möjligheter att använda självadministrerade enkäter som stöd för vissa delar av intervjun (Holbrook, Green & Krosnick, 2003).

Sammanfattande punkter:

- Beror skillnaderna mellan datainsamlingsmetoderna på social närhet? I sådana fall borde besöksintervjun visa störst problem med känsliga frågor men så är det inte. Effekterna av känsliga frågor är störst i telefonintervjun.

Sätt att minska eller mäta känsligheten

För känsliga frågor kan alltså självadministrerade datainsamlingsmetoder vara att föredra. Det finns olika sätt att försöka mäta social önskvärdhet (till exempel Marlowe-Crowner-skalan) för att få en uppfattning om problematiken och kunna vidta eventuella åtgärder. I besöksintervjuer finns även möjligheten att använda självadministrerade metoder för frågor med känsligt innehåll. Det finns även möjlighet att använda så kallade randomiserade svarstekniker där respondenten får svara på antingen en känslig eller en trivial fråga. Vilken av de två frågorna respondenten ska svara på är okänt för intervjuaren och avgörs av en slumpmekanism med känd sannolikhet för de två utfallen. Eftersom sannolikheten är känd kan man på makronivå räkna ut skattningar för känsliga frågor. Eftersom intervjuaren inte vet vilken av frågorna respondenten svarar på är svaren förhoppningsvis mer ärliga och det blir mindre känsligt för respondenten. Det finns också tips för att genom frågekonstruktionen motverka social önskvärdhet (läs mer om frågeteknik i kapitel 4).

Datainsamlingsmetoder och satisficing

En annan svarstendens som tas upp på många platser i den här boken är satisficing. Det betyder att respondenten svarar "tillräckligt bra" istället för optimalt, till exempel genom att ta olika genvägar fram till ett svar som också innebär att svaret får lägre kvalitet. Satisficing är relaterat till respondentens kognitiva förmåga, motivation och uppgiftens svårighetsgrad.

Hur påverkar olika intervjumetoder satisficing? Är den större eller mindre i telefonintervjuer jämfört med besöksintervjuer och självadministrerade enkäter? Den borde vara större i telefonintervjuer än i besöksintervjun eftersom kommunikationen är mer begränsad, respondentens investering lägre, tempot sannolikt högre och där respondenten kan hålla på med annat samtidigt som de deltar i intervjun.

Det finns forskning som också visar detta, att satisficing är mer utbrett i telefonintervjuer än i besöksintervjuer. Jämförelser mellan telefonintervjuer och självadministrerade enkäter visar samma tendens – mer satisficing i telefonintervjuer (Chang & Krosnick, 2010; Holbrook, Green & Krosnick, 2003).

Sammanfattande punkter:

- Risken för satisficing är större i telefonintervjun än i besöksintervjun.

Informationstyper – skrift eller ljud

Att läsa eller lyssna

Olika datainsamlingsmetoder använder olika informationstyper för att kommunicera med respondenterna. Dessa informationstyper är knutna till olika sinnen. Vid en intervju ställs frågorna av intervjuaren via samtal, antingen på telefon eller via besök. Där är informationen således ljudmässig och upptas av respondentens hörsel. Vid självadministrerade datainsamlingsmetoder ställs frågorna skriftligt i en enkät. Där är informationen visuell och upptas via synen. Det är en central skillnad mellan intervjuer och självadministrerade datainsamlingsmetoder (de Leeuw, 2008).

I självadministrerade enkäter läser respondenten frågorna. Detta möjliggör att respondenten kan stanna upp i sin läsning eller läsa om delar av frågan eller svarsalternativen efter behov. Om något skulle vara otydligt eller missas kan alltså respondenten gå tillbaka och läsa om frågan, eller skanna texten fram och tillbaka. I intervjuer finns inte denna möjlighet. Intervjuaren läser frågan från början till slut och sedan är det respondentens tur att svara. Givetvis kan respondenten be intervjuaren att förtydliga eller upprepa delar av eller hela frågan. För att detta ska inträffa krävs förmodligen ett visst mått av engagemang från respondenten samt en medvetenhet om att frågan har missförstås. Den här typen av avbrott är dessutom något som i sig bör undvikas i intervjuer eftersom det ökar både administrationstiden, bördan för respondenten och kan störa standardiseringen. Vid intervjuer är målet med frågorna att respondenten ska kunna förstå dem efter en rak uppläsning av intervjuaren. Det är givetvis också målet i självadministrere-

rade enkäter men i intervjuer finns en speciell begränsning som man måste ta hänsyn till – respondenternas arbetsminne.

Arbetsminnet är ett minnessystem där vi behåller information som vi behöver för tillfället eller "arbetar" med (läs mer om minnet i kapitel 11). Information kan behållas i arbetsminnet främst genom upprepning eller så kallad elaborering (Lundh, Montgomery & Waern, 2008). Arbetsminnet är begränsat både vad gäller utrymme och tid. Om ny information förs in kan äldre information falla bort. Det är alltså en väldigt annorlunda förutsättning jämfört med självadministrerade datainsamlings sätt.

Sammanfattande punkter:

Olika datainsamlingsmetoder riktar sig till olika sinnen hos respondenten:

- **Självadministrerad** innebär att respondenten läser med hjälp av sin syn.
- **Intervju** innebär att respondenten lyssnar med sin hörsel.

Vid intervjuer kan inte respondenten kontrollera informationsflödet, till exempel läsa om frågan vid behov. Därför är det av stor vikt hur mycket respondenten kan "hålla i huvudet" samtidigt i det så kallade arbetsminnet.

Layout – första intrycket

Vid självadministrerade postenkäter bidrar säkert enkäten i sig till beslutet att delta eller inte. Respondenten kan bläddra i enkäten för att skaffa sig en uppfattning om frågorna och undersökningen. Denna möjlighet finns förstås inte i en intervju och är betydligt omständligare även i en självadministrerad webbenkät. Layout och vilka frågor som inleder frågeformuläret är alltid viktigt men på grund av ovanstående resonemang borde det vara extra viktigt i en postenkät eftersom det där bidrar till det första intrycket.

Långa frågor och intervjuer

Längd är en viktig enhet för frågekonstruktion inte bara vad gäller frågeformuläret som helhet, utan även för de enskilda frågorna (läs mer i

kapitel 2 för en allmän diskussion om frågelängd). I detta avsnitt behandlas endast ett datainsamlingsperspektiv på frågelängden.

Som beskrivs ovan är vi i intervjuer beroende av arbetsminnet och måste beakta dess begränsningar. Det gäller speciellt i telefonintervjuer där det inte finns möjligheter till visuellt stöd. Frågorna kan därför inte vara för långa.

EXEMPEL:

Har du i någon utsträckning andra motionsvanor nu jämfört med innan det infördes ett friskvårdsbidrag för statligt anställda (2003)? Med motionsformer avses här den typ av motion som omfattas av friskvårdsbidraget.

Ovanstående fråga är för lång och innehåller för mycket information för att ställas i en intervju. Det finns uppenbara risker för missförstånd eller att information missas. Den här typen av fråga bör således delas upp i flera frågor om den ska ställas i en intervju. Det finns exempel på att långa frågor kan ge en bättre mätning eftersom respondenterna får mer betänketid (läs mer i kapitel 11). Men då handlar det främst om att förlänga frågan till exempel genom att introducera ämnet, inte att själva frågan innehåller mycket information.

I intervjuer bör alltså frågorna vara så korta som möjligt. Det är dock viktigt att påpeka att denna förkortning inte får ske på bekostnad av förståelsen. En kort men otydlig fråga kan få längre administrationstid på grund av motfrågor, krav på förtydliganden och upprepning av frågan, än en lite längre men tydligare fråga (Schaeffer, 2010).

Sammanfattande punkter:

- Intervjufrågor måste vara kortfattade annars kan information missas.

Svarsalternativen i intervjuer

Svarsalternativen är en viktig del av frågan (läs mer i kapitel 3). I en intervju tar de dock tid att läsa upp och ökar administrationstiden för frågan som helhet. De förlänger därmed intervjutiden vilket ökar både bördan för respondenten och kostnaden för undersökningsorganisationen. Svarsalternativen ökar också informationen som respondenten

måste behålla i arbetsminnet, vilket kan leda till att information missas eller att frågan som helhet missförstås.

Förgreningsteknik

På grund av ovanstående läses inte alltid svarsalternativen upp i sin helhet i intervjuundersökningar. Ibland läses inte svarsalternativen upp alls utan intervjuaren klassar själv respondentens svar och markerar det i datorstödet. Man kan även använda så kallad branchingteknik (ungefär "förgreningsteknik") – att vissa svarsalternativ först läses upp för att ringa in var respondenten befinner sig på svarsdimensionen, sedan läses mer specifika svarsalternativ upp inom det inringade området (se figur nedan).

Figur 6.1

Är din inställning till detta...

1) ... positiv

2) ... negativ?



Är den ...

1a) ... mycket positiv? 1b) ... ganska positiv? 2a) ... mycket negativ? 2b) ... ganska negativ?

Malhotra, Krosnick och Thomas (2011) undersökte vilken förgreningsteknik som ger bäst resultat vid bipolära skalor. De fann att sjugradiga skalor är bäst för bipolära skalor och att respondenten i första skedet ska svara för de två polerna samt ett neutralt alternativ, till exempel "bra", "dåligt" och "varken eller". Inledningsvis ska alltså endast tre alternativ presenteras. De som svarar att de är neutrala ska inte få fler frågor men de som väljer en av polerna "bra" eller "dåligt" ska presenteras tre nivåer av den pol de valt. Skalan blir då sjugradig – tre steg för respektive pol plus ett neutralt alternativ.

En ny studie visar dock att förgreningstekniken kan ge ett mer extremt svarsmönster och dessutom kostar mera att genomföra eftersom en fråga administreras i två steg istället för ett (Gilbert, 2015). Det behövs alltså mer forskning innan man säkert kan rekommendera om man ska använda förgreningsteknik eller inte.

Mot bakgrund av ovanstående exempel är det viktigt att inför en intervjuundersökning ta ställning till om och hur svarsalternativen ska läsas upp. Detta har nämligen inverkan på frågekonstruktionen och är enligt en del studier även relaterat till intervjuareffekter (Japac, 2005).

Frågor utan svarsalternativ

Om svarsalternativen inte ska läsas upp kommer respondenten inte att bli varse om dem. Det måste man ta hänsyn till när man konstruerar frågorna. Som tidigare skrivits innehåller svarsalternativen information som kan förtydliga och förbättra förståelsen av frågan. Om svarsalternativen inte läses försvinner den informationen från respondenten.

EXEMPEL:

När flyttade du hemifrån för första gången?

1986–1990

1991–1995

1996–2000

2001–2005

2006–2011

I självadministrerade enkäter är ovanstående fråga inget problem. Svarsalternativen visar att svaret ska anges i årtal. I en telefonintervju utan svarsalternativen är det däremot inte alls självklart. Frågan kan ge svar i stil med "när min mamma dog", "det var när jag tog studenten" eller "det var på sommaren" (Fowler, 1995). Dessa svar kan mycket väl vara korrekta svar på frågan – det är frågan det är fel på. "När" är inte tillräckligt specifikt utan den konkretiserande informationen i svarsalternativen. Om inte svarsalternativen ska läsas upp måste denna fråga därför formuleras om vid en telefonintervju, annars finns det risk för en ostrukturerad intervjusituation och även mätfel. Den enklaste lösningen är att specificera "när" till den svarsenhet som respondenten ska avge svaret i – "Vilket år flyttade du hemifrån?".

Ett annat exempel:

EXEMPEL:

Hur ofta besöker du vårdcentralen?

Ofta
Ibland
Sällan

Om inte svarsalternativen läses kan respondenten svara något annat än de svarsalternativ som finns i intervjuarens datorstöd. För att undvika detta kan frågan formuleras om så att svarsalternativen ingår i frågeformuleringen: *Besöker du vårdcentralen ofta, ibland eller sällan?*

Kraven på att vara specifik i frågan är alltså mycket högre om inte respondenten delges svarsalternativen. Det handlar om att inte möjliggöra några andra svar än de som tillåts. Detta är viktigt att tänka på speciellt när en enkät används som utgångspunkt för en telefonintervju (till exempel vid en telefonuppföljning). Frågor som tett sig som oproblematiske i en enkät kan behöva formuleras om för att passa intervjuformatet, till exempel om svarsalternativen ska läsas eller inte.

Kodning av svar

Ibland är det dock inte möjligt att specificera frågeformuleringen så pass långt att respondentens svar smidigt passar i något av ett fåtal svarsalternativ. Vissa frågor är så pass breda och omfattar så pass många tänkbara svar att de kräver att respondentens svar kodas utifrån en längre lista av alternativ. Ett exempel är frågor om respondentens yrke. Sådan kodning, eller klassificering som det också kallas, kan ske antingen i realtid av intervjuaren i fält (så kallad fältkodning) eller efter datainsamlingen. Den kan också ske datoriserat eller manuellt. Det finns en del saker som kan gå fel i en kodningsprocess. Allra viktigast är det att förebygga fel vid fältkodning, eftersom det då inte finns någon rådata att gå tillbaka till och korrigera felen utifrån. Det behövs bland annat tydliga kodlistor, användarvänligt datorstöd och även utbildning av intervjuarna (läs mer i Biemer & Lyberg, 2003).

Sammanfattande punkter:

- Svartalternativens vara eller icke vara är ett viktigt beslut i intervjuer eftersom de tar tid att läsa och det blir mer för respondentens att "hålla i huvudet".
- Man kan använda förgreningsteknik för att förenkla läsningen och förståelsen. Dock är inte forskningen entydigt positiv till detta.
- Om svartalternativen *inte* ska läsas måste frågan i sig vara tillräckligt specifik så att den leder respondenterna till de giltiga svaren.
- Om frågan är öppen och intervjuaren ska klassa respondentens svar behövs det bra stöd, speciellt om det sker i fält.

Att använda annan information än frågeformuläret**Externa hjälpmedel**

Som tidigare beskrivits kontrollerar respondenten själv administrationen av frågorna i självadministrerade insamlings sätt. Respondenten kan då ägna så lång eller kort tid som han eller hon vill och behöver på de enskilda frågorna och på enkäten som helhet. I intervjuer är det intervjuaren som kontrollerar administrationen och tempot är relativt högt. Detta gör att vissa typer av frågor är mer lämpliga för självadministrerade insamlings sätt.

En sådan faktor är om respondenten behöver använda extern hjälpinformation som till exempel kvitton eller räkningar för att besvara frågorna. I en telefonintervju kan respondenten visserligen informeras om att ha nödvändig hjälpinformation till hands vid intervjutillfället, men generellt passar självadministrerade enkäter bättre vid förfarandet.

En annan faktor är om respondenten behöver beräkna fram svaret eller omvandla det från andra uppgifter, till exempel från ett datasystem eller register. Den typen av undersökningar är vanliga till företag och organisationer och de genomförs i stort sett alltid självadministrerat.

En tredje faktor är om frågorna kräver djupa genomsökningar i minnet. Det finns forskning som antyder att längre svarstid kan resultera i högre kvalitet på svaren på sådana frågor (Tourangeau, Rips & Rasinski, 2000). I telefonintervjuer är svarstiden mer begränsad och inte under respondentens kontroll. Å andra sidan kan en intervjuare genom

följdfrågor eller annan feedback underlätta återerinringen för respondenten.

Sammanfattande punkter:

- Undersökningar som kräver externa hjälpmedel för eller beräkningar av respondenten genomförs bättre självadministrerat än över telefon.

Kalenderintervju

I besöksintervjuer är tempot ofta långsammare än i telefonintervjuer och det finns också möjligheter till hjälpmedel. Detta möjliggör andra intervjutekniker. En sådan är kalenderintervjun som sker med en kalender som stöd. Den är framförallt aktuell vid undersökningar som är beroende av det episodiska minnet hos respondenten (respondentens egna erfarenheter och upplevelser) snarare än faktakunskaper. Till skillnad från standardiserade intervjuer är det en konverserande intervjuform. Den utgår ifrån respondentens egen historia av händelser för att underlätta återerinring från minnet och i slutänden få bättre rapportering. Forskning visar att intervjuare med kalendermetoden arbetar mer med att underlätta återerinring men också använder mer följdfrågor som kan påverka respondentens svar jämfört med mer standardiserade intervjuer (Bilgen & Belli, 2010).

Det finns även fler intervjutekniker som, liksom kalenderintervjun, använder underlag för att främja respondentens minne. I en undersökning om utgifter kan förslagsvis en exempelbok med bilder på olika typer av varor användas.

Sammanfattande punkter:

- Att utgå från en kalender och respondentens historia kallas för kalenderintervju och kan förbättra minnet. Det finns även andra, liknande tekniker.

Beroende intervju

Eftersom intervjun är flexibel kan den utnyttja information på ett annat sätt än pappersenkäten. I undersökningar där samma respondenter får svara på samma eller liknande frågor vid flera olika tillfällen, kan information från tidigare tillfällen användas, till exempel för att förenkla

frågeformuleringar och leda förbi följdfrågor. Ett exempel är att fråga om respondenten fortfarande har samma sysselsättning som vid förra tillfället, istället för att ställa alla frågorna för att utreda sysselsättningen igen. Detta kallas för "beroende intervju" ("dependent interviewing" på engelska). Syftet med beroende intervjuer är att förbättra datainsamlingen, datakvaliteten och att minska respondentbördan. Tekniken anses leda till besparingar i tid och pengar (Jäckle, 2008).

En studie undersökte om olika tekniker av beroende intervju ledde till högre effektivitet, lägre respondentbörda och om det skilde sig mellan olika typer av frågeområden. Resultaten visade att effektivitetsvinsten framförallt ligger i minskad tid för efterarbete i form av kodning. Det är dock många faktorer som påverkar och som man behöver ta hänsyn till, till exempel graden av stabilitet i undersökningsgruppen. Om förändringar skett mellan mättillfällena kan beroende intervju till och med vara mer tidskrävande och omständigt, än att ställa huvudfrågorna igen. Detta beror i sin tur på referensperiodens längd och undersökningsämnet (Jäckle, 2008). Beroende intervjuteknik uppskattas dock av intervjuarna (Sala, Uhrig & Lynn, 2011).

En motsvarighet till beroende intervju är förtryckta uppgifter i en enkät där tidigare lämnade svar är tryckta på enkäten. Respondentens roll är då att avgöra om de förtryckta uppgifterna fortfarande gäller eller inte. En skillnad är dock att den beroende intervjun ger mer kontroll eftersom den leds av en intervjuare.

Sammanfattande punkter:

- Beroende intervju är ett sätt att utnyttja tidigare insamlad information för att minska administrationstiden och bördan.

Positionseffekter och datainsamlingsmetod

Informationstypen är olika i intervjuer och självadministrerade datainsamlingsätt (ljudmässig jämfört med visuell). Det är relaterat till så kallade positionseffekter. Det betyder att svarsalternativens position i sig påverkar svaren. I självadministrerade datainsamlingsätt finns en högre tendens att välja tidiga svarsalternativ. I intervjuer är det tvärtom – där är tendensen att i högre grad välja de sista svarsalternativen. Dessa fenomen diskuteras och förklaras mer djupgående i kapitel 13.

Enkäter i smarttelefoner och surfplattor

Smarttelefon

Smarttelefoner har blivit alltmer aktuella som datainsamlingsmedium. Oavsett om man önskar det eller har planerat för det kan man räkna med att en del respondenter kommer att svara på en webbenkät genom sin smarttelefon. Detta gäller speciellt om undersökningsinbjudan och påminnelser sker via kanaler som ligger inom smarttelefonens användningsområden (t.ex. e-post eller sms). Dock bör man ha förekomsten av smarttelefon i undersökningsgruppen i åtanke innan man beslutar att använda den som huvudsaklig datainsamlingsmetod. Av de som har en smarttelefon så varierar också erfarenheten och användningen. Man kan äga en smarttelefon men ändå vara främmande för att svara på en enkät via den.

Så vad innebär smarttelefonens entré för datainsamlingen? Först kan man konstatera att den teknologiska utvecklingen går mycket fort på det här området. Det kan innebära att forskning och praktiska erfarenheter har ett kort "bäst-före datum" och snabbt kan bli inaktuella. Det bör man ha i åtanke om man ska designa en undersökning för smarttelefoner.

Smarttelefonen delar egenskaper både med den traditionella webbenkäten och med telefonintervjun. Urvalspersonerna kan i smarttelefonen kontaktas via telefonsamtal, sms eller e-post. De kan även själva kontakta undersökningsorganisationen via dessa kanaler. Enkätverktyget kan vara detsamma som i en webbenkät fast den besvaras via smarttelefonens webbläsare istället för datorns (Buskirk, 2015).

Det finns dessutom både möjligheter och begränsningar som är unika för smarttelefonen. En unik möjlighet är att samla in annan typ av data än vad som är möjligt med till exempel en webbenkät som besvaras via dator. Det kan handla om att använda GPS, skanning av streck- och QR-koder eller foton. Om sådan information ska användas i undersökningen måste man dock överväga hur man ska få in informationen från respondenter som inte har någon smarttelefon, eftersom de andra datainsamlingsmetoderna inte har någon direkt motsvarighet till detta. Ett alternativ kan vara att "låna ut" eller ge smarttelefoner till urvalspersonerna. Vid större undersökningar är detta sannolikt en orimlig kostnad men för mer avgränsade undersökningar som är i stort behov av sådan information kan det vara en möjlighet. Ett annat och

troligen mer rimligt alternativ är att erbjuda kompletterande metoder för de som inte kan svara via smarttelefon. De skulle till exempel kunna få skicka in sina kvitton som sedan skannas av undersökningsorganisationen (Buskirk, 2015).

En annan unik möjlighet med smarttelefonen är möjligheten att kommunicera med urvalspersonerna via en applikation (app). Läs mer om detta nedan.

En unik begränsning är att skärmen på en smarttelefon är mycket mindre än skärmen på en dator. Samma sak gäller smarttelefonens tangentbord. Detta ställer krav på enkätdesignen eftersom risken för fel generellt sett är större än i andra insamlingsverktyg. Exempel på sådana fel kan vara att felaktiga värden eftersom respondenten tryckt på fel tangent eller att frågor missas för att de ligger "utanför" skärmbilden. Det begränsande formatet medför också att enkäter som ska besvaras via smarttelefoner bör vara relativt korta.

Att smarttelefonen är bärbar (urvalspersonen har den med sig) är en styrka om till exempel streckkoder ska samlas in. Men det kan också vara en svaghet. Exempelvis kan respondentens uppkoppling försvinna under besvarandet. Batteriet kan också ta slut. Man bör också ha i åtanke att urvalspersoner kan ha ett annat abonnemang för internetuppkoppling till sin smarttelefon än vad de har hemma. Eftersom nedladdad information kan kosta för respondenten bör banners, logotyper och annat material som inte behövs för besvarandet begränsas. Det tar också extra tid att ladda sådant material och det ökar därför respondentbördan (Buskirk, 2015).

Olika sätt att besvara en enkät i en smarttelefon

Enligt Buskirk & Andrus (2012) kan en enkät i en smarttelefon besvaras via en vanlig webbsida, en mobilanpassad webbsida eller via en applikation (app).

Enkät via vanlig webbsida har fördelen att det inte krävs några extra resurser för undersökningsorganisationen eftersom ingen anpassning till smarttelefonen sker. Nackdelen är att eftersom ingen anpassning sker kommer respondenten sannolikt att behöva scrolla och zooma för att få formatet hanterligt i mobilen. Det innebär högre respondentbörda och troligtvis fler felmarkeringar, avbrutna sessioner och högre bortfall. Formatet kräver också konstant uppkoppling och kan innebära långa

laddningstider. Sammanfattningsvis är det ingen bra lösning för smarttelefoner, speciellt inte för längre enkäter (Buskirk & Andrus, 2012).

Mobilanpassade webbsidor har flera fördelar. Med denna lösning är det visuella proportionellt och som enkätdesigner har man större kontroll över vad som ska synas på skärmen. Det blir mindre scrolling och zoomning, men kräver större utvecklingskostnader. Även om formatet innebär större kontroll än vid en vanlig webbsida så kan det likväl vara svårt att styra hur enkäten kommer att se ut i olika typer av smarttelefoner med olika webbläsare och operativsystem (Buskirk & Andrus, 2012).

Enkät via applikation är det i särklass bästa formatet ur mätteknisk synvinkel. I en app är layout och orientering anpassad. Det går även att kommunicera med urvalspersonen utanför själva enkäten, som till exempel att skicka påminnelser via appen. Detta är en stor styrka vid dagbokenkäter eller när respondenten ska svara vid specifika tidpunkter (t.ex. var de är eller vad de gör vid ett slumpmässigt klockslag). Förutom utvecklingskostnader är den stora utmaningen att få urvalspersonen att ladda ner och lära sig appen. För vissa undersökningar med speciella urval kan det ändå vara ett alternativ (Buskirk & Andrus, 2012).

Surfplatta

Surfplattan är också ett relativt nytt teknologiskt medium som kan användas för att besvara undersökningar. Surfplattan delar vissa men inte alla av smarttelefonens egenskaper. Precis som för smarttelefonen får man räkna med att en del kommer att svara på en webbenkät via sin surfplatta.

- Surfplattan används inte primärt för telefoni. Det medför att kontakt via sms inte är möjlig.
- Surfplattans tangentbord och skärm är mindre än datorns men större än smarttelefonens. Det är alltså smarttelefonen och inte surfplattan som sätter gränserna för enkätdesignen.
- Surfplattan används i första hand i hemmet. Insamling av information utanför hemmet (t.ex. streckkoder) är därför nästan omöjlig men å andra sidan finns inte heller samma problem med avbrott och nedladdning som för smarttelefonen.

Sammanfattande punkter:

- Smarttelefonen är relativt ny som verktyg för att besvara enkäter. Teknologin går fort framåt. Forskningen kan därför ha ett kort "bäst föredatum".
- Smarttelefonen delar egenskaper både med den traditionella webbenkäten och med telefonintervjun.
- Urvalspersonerna kan i smarttelefonen kontaktas via telefonsamtal, sms eller e-post.
- En unik möjlighet är att samla in annan typ av data än vad som är möjligt med t.ex. webbenkäten. Det kan handla om att använda GPS, skanning av streck- och QR-koder eller foton.
- En unik begränsning är att skärmen på en smarttelefon är mycket mindre än skärmen på en dator. Samma sak gäller smarttelefonens tangentbord. Detta ställer krav på enkätdesignen eftersom risken för fel generellt sett är större än i andra insamlingsverktyg.
- En annan begränsning är att respondentens uppkoppling kan försvinna under besvarandet. Batteriet kan också ta slut.
- Det finns olika sätt att besvara en enkät i en smarttelefon: 1) via vanlig webbsida, 2) mobilanpassad webbsida eller 3) via app.
- Vanlig webbsida är minst krävande för undersökningsorganisationen men mest krävande för respondenten. För en app är det motsatta förhållanden.
- Surfplattan har större skärm och tangentbord än smarttelefonen.
- Till skillnad från smarttelefonen används surfplattan bara i hemmet vilket innebär att uppkoppling inte är något problem. Å andra sidan kan inte information utanför hemmet samlas in.

Frågetyper och datainsamlingsmetod

Ovanstående avsnitt beskriver de ramar som datainsamlingsmetoden ger för frågekonstruktionen. Detta och följande avsnitt är mer konkret inriktade på själva frågekonstruktionen och innehåller resonemang om olika frågetyper och en del praktiska tips.

Åsikts- och attitydfrågor

Attitydfrågor har ofta formen av påståenden med en instämmer-skala som svarsalternativ (läs mer i kapitel 3 och i kapitel 12). Denna uppställning innebär störst risk för hålla-med-tendenser och den tendensen är starkare på telefon än i självadministrerade datainsamlings-sätt. Orsaken är förmodligen den sociala situationen (läs mer i avsnittet Självadministrerat eller socialt samspel ovan). Den kan också påverka svaren om attitydobjektet är känsligt. Social önskvärdhet är i regel starkare på telefon än i enkäter.

Attitydfrågor kan vara den frågetyp som är mest känslig för ordalydel-sen. Det gäller speciellt om attitydobjektet är något okänt, oväsentligt eller otydligt för respondenten. Vid attitydfrågor borde därför läsavvikel-ser från intervjuare få störst konsekvenser. I intervjuundersökningar är det extra viktigt att påpeka det vid utbildningen av intervjuare. I dator-stödet av attitydfrågor är det viktigt att det framgår vad som är obligatorisk text som alltid ska läsas upp för varje respondent.

En annan stor skillnad mellan datainsamlingsmetoderna har att göra med möjligheten att erbjuda "icke-svar", till exempel "Ingen åsikt". Detta kan påverka svarsfördelningen på olika sätt (läs mer i kapitel 3). Intervjuer har möjligheten att utelämna "Ingen åsikt" när svarsalternativen läses upp. Det kan vara en fördel om de erbjuds först när respondenten tvekar.

Sammanfattande punkter:

- Hålla-med-tendenser är starkare i telefonintervjuer vilket innebär att metoden passar sämre för attitydfrågor.
- Attitydfrågor är känsliga för ordalydel-sen – intervjuarna bör inte avvika i läsningen.
- I intervjuer, och i viss mån i webbenkäter, kan svarsalternativ för icke-svar som till exempel "Vet inte" erbjudas vid behov.

Beteendefrågor

Beteendefrågor besvaras ofta inte bara med "Ja" eller "Nej", utan även med någon form av frekvensskala. Det finns i huvudsak två olika typer av frekvensskalor: subjektiv (till exempel mycket ofta, ganska ofta,

ganska sällan, mycket sällan) eller objektiv/faktisk (aldrig, 1–3 ggr, 4–7 ggr, 8 ggr eller fler).

I en telefonintervju kan intervjuaren instrueras att inte läsa upp svarsalternativen. För respondenten uppfattas frågan då som "öppen" (se nedan) och intervjuaren passar in svaret i de befintliga svarsalternativen. Man bör sträva efter att det framgår av frågan i vilken enhet respondenten förväntas svara. Om svaret till exempel ska ges i objektiva frekvenser bör man fråga "Ungefär hur många gånger under de senaste tolv månaderna..." och inte "Hur ofta under de senaste tolv månaderna...". Det beror på att det sistnämnda lättare leder till att respondenten svarar på en subjektiv skala. Sådana svar måste alltid följas upp av intervjuaren. Det leder till ökad intervjutid samt ökad risk för intervjuarvarians eftersom det är lätt hänt att intervjuaren tolkar respondentens subjektiva svar, till exempel "ofta", efter sina egna referensramar för hur många gånger som skulle kunna vara "ofta".

Slutsats: Om frågan mäter något i objektiva frekvenser, se till att frågan i en telefonintervju indikerar vad som avses och läs bara upp svarsalternativen vid behov.

Om frågan däremot har en subjektiv frekvensskala är det i de flesta fall mest lämpligt att svarsalternativen alltid läses upp som en del av frågan. Om det av någon anledning inte är möjligt bör frågan som läses upp vara formulerad på ett sätt som tydligt indikerar för respondenten att svaret ska avges i subjektiva och inte faktiska frekvenser. Vid subjektiva frekvenser är risken större att intervjuaren tolkar svaret och omedvetet styr sina följdfrågor mot vissa svarsalternativ.

Sammanfattande punkter:

- Det måste framgå vilken svarsenhet frågan har vid intervjuer (oavsett objektiv eller subjektiv svarsskala).

Faktafrågor

Faktafrågor som frågar efter faktauppgifter (till exempel yrke, typ av anställning, inkomst, boendeform, utbildning) fungerar som regel bra i både enkäter och intervjuer. Om faktauppgiften är komplicerad och kräver utredningar eller klassificering kan det dock vara en styrka att ha en intervjuare som kan förklara för respondenten, administrera frågorna och koda respondentens svar.

Kunskapsfrågor

I mätningar av kunskap har självadministrerade enkäter alltid haft bristen att respondenten kan leta fram rätt svar. Den bristen har ökat i och med internets intåg som lättillgänglig kunskapskälla. Därmed inte sagt att alla eller ens många respondenter utnyttjar den möjligheten, men risken finns.

Intervjun skulle alltså kunna vara en mer lämplig insamlingsmetod för kunskapsfrågor. Även där finns det en del problem. Respondenten kan uppfatta kunskapsfrågor ställda av en annan människa (intervjuaren) som ett obehagligt test (Bergman & Wärneryd, 1982). Vissa typer av kunskapsfrågor kan också innehålla väldigt mycket information (till exempel scenariofrågor) eller många svarsalternativ (till exempel "multiple choice"-frågor), vilket gör dem mindre lämpliga i intervjuer.

Kunskapsfrågor är relativt ovanliga i surveyundersökningar. Det är dock inte helt ovanligt att de förekommer som en mindre del av något större ämnesområde (läs mer i kapitel 10). En annan typ av kunskapsfrågor är de som syftar till att mäta vad respondenten har uppmärksammat till exempel i närområdet eller i nyhetsflödet – "Känner du till..." eller "Har du hört talas om...". I denna typ av fråga uppmanas respondenten att värdera snarare än att redovisa sin kunskap. Frågetypen borde därför inte upplevas som ett test i samma grad. Denna typ borde fungera i alla datainsamlingsätt. Intervjun har fördelen att intervjuaren kan ge mer information om företeelsen. Nackdelen är risken för socialt önskvärda svar, till exempel om frågan gäller prestigefyllda områden.

Sammanfattande punkter:

- Kunskapsfrågor där respondenten ska redovisa sin kunskap är ganska ovanliga i surveyundersökning. Frågor där respondenten ska värdera sin kunskap är vanligare (läs mer i kapitel 10).

Mixed-mode (kombinerad insamling)

Mixed-mode betyder att fler än en metod används i datainsamlingen till en undersökning. På svenska används ibland beteckningen "blandad insamling" eller "kombinerad insamling". Det är alltså kopplat till själva datainsamlingen. Att använda flera metoder i en undersökning (till exempel för introduktionsbrev eller påminnelser) brukar inte anses vara mixed-mode om datainsamlingen endast sker med en metod.

Mixed-mode är ett ganska komplicerat område. Man skulle kunna säga att det innefattar allt som är viktigt i en undersökning (till exempel bortfall, teknik, formulärdesign, mätning, kostnader) fast två gånger. Inför införandet av mixed-mode bör man, åtminstone i en större och viktig undersökning, resonera kring hur alla fel (mätfel, bortfallsfel etc.) påverkas om mixed-mode används jämfört med om endast en datainsamlingsmetod ingår. Dessutom måste man ta ställning till kostnaderna både kort- och långsiktigt, de praktiska förutsättningarna (som tekniska lösningar och planeringen av fältarbetet) och mjukare konsekvenser som till exempel goodwill-effekter av att erbjuda respondenterna fler än ett sätt att svara på. Det är en utmanande ekvation!

Bedömningen underlättas inte heller av att många faktorer i ekvationen är okända eller i alla fall inte självklara. Vi vet inte allt om hur urvals-personer resonerar när de beslutar sig för att delta eller inte delta i en undersökning och hur det varierar mellan olika datainsamlingsmetoder. Vi har inte heller full kunskap om de kognitiva processerna hos respondenterna och hur det kan skilja sig åt om frågan är skriftlig i en enkät eller om den läses upp av en intervjuare. Men vi vet en del och den kunskapen bör ingå i bedömningen.

Ett komplicerat ämne alltså men samtidigt viktigt för det finns potential i att använda mixed-mode, främst vad gäller att minska kostnader, höja svarsfrekvensen och minska bortfallsfel.

Mixed-mode – grundläggande begrepp

Multi-mode system

Med mixed-mode menas som sagt oftast att fler än en metod används i datainsamlingen, men en undersökning består av fler steg än datainsamling (de Leeuw, Hox & Dillman, 2008):

- inledande kommunikation med urvalspersonerna
 - screening
 - förhandsinformation
 - inbjudan att delta
- ordinarie datainsamling
- påminnelser
- uppföljande datainsamling
- återkoppling (till exempel feedback).

Dessa steg kan genomföras i olika medium vilket då ger ett så kallat "multi-mode survey system", även om datainsamlingen i sig endast använder en metod. Exempelvis kan den inledande kommunikationen ske med post men datainsamlingen via telefonintervjuer. Detta är standard för SCB:s intervjuundersökningar som inleds med ett postalt introduktionsbrev. Syftet med detta är bland annat att det är lättare att upprätta legitimitet med ett officiellt brev med logotyp, kontaktpersoner och signatur än via en främmande röst i telefonluren. Ett annat exempel är att använda en telefonpåminnelse för en postenkätsundersökning.

Att använda flera medium för kommunikationen med urvalspersonerna är väldigt vanligt och har en lång historia. Allmänna syften är, precis som i exemplet ovan, att öka motivationen, etablera legitimitet och minska bortfallsfelet. Det finns få risker med detta utan tvärtom stora fördelar så länge som datainsamlingen bara använder ett medium. De risker som finns med att blanda medium är alltså kopplat till själva datainsamlingen (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008).

Varför blanda insamlingsmetoder?

Biemer och Lyberg (2003) skrev för över tio år sedan att mixed-mode var normen i många länder – ett faktum som har stärkts sedan dess (de Leeuw, Hox & Dillman, 2008).

Vad är då den tänkta vinsten med att blanda datainsamlingsmetoder? Det kan finnas flera. De vanligaste uttryckta är (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008):

- **Minskat bortfallsfel** kan fås genom att höja datainsamlingens täckningsgrad (inte att förväxla med urvalsramens täckning).

Faktisk täckningsgrad: Med flera datainsamlingsmetoder kan fler urvalspersoner nås och ges möjlighet att delta (en person utan dator kan inte delta i en webbundersökning men kan svara om undersökningen också erbjuder en pappersenkät).

Psykologisk täckningsgrad: Urvalspersoner som inte kan tänka sig att svara via en datainsamlingsmetod kan fångas upp av en annan metod.

Genom att använda flera datainsamlingsmetoder kan man alltså täcka in och få svar från urvalspersoner som annars inte skulle ha svarat och som förmodligen är annorlunda (de som har tillgång till

internet skiljer sig till exempel åt från de som inte har det). Det är därför en möjlighet att minska bortfallsfelet.

- **Ökad svarsfrekvens kan uppnås** genom att använda flera datainsamlingsmetoder, vilka kan fånga upp fler respondenter och inhämta fler svar (vilket har ett egenvärde bortom minskat bortfallsfel).
- **Förbättrad mätning** kan erhållas genom att utnyttja olika datainsamlingsmetoders styrkor, till exempel använda intervju för komplicerade frågor och självadministrerade metoder för känsliga frågor.
- **Minskad kostnad** kan erhållas genom att erbjuda eller "styra" respondenterna till en billigare datainsamlingsmetod snarare än en dyr, till exempel erbjuda webb som ett komplement till telefonintervju.
- **Goodwill** innebär att tillmötesgå respondenternas önskemål om sätt att svara på kan undersökningsklimatet förbättras.
- **Snabbare datainsamling** kan uppnås genom att erbjuda snabbare metoder och alternativ som passar olika typer av respondenter.

Att öka svarsfrekvensen, minska bortfallsfel och minska kostnader anses vara de vanligaste skälen (de Leeuw, Hox & Dillman, 2008; de Leeuw, 2009). Ovanstående skäl är alltså de som oftast ligger bakom valet att använda mixed-mode i praktiken.

Hur blanda insamlingsmetoder

Datainsamlingsmetoder kan blandas på olika sätt (de Leeuw, Hox & Dillman, 2008):

- sekventiell uppdelning
- parallell uppdelning
- formuläruppdelning
- urvalsuppdelning
- omgångsuppdelning.

De tas upp under varsin rubrik nedan.

Sekventiell uppdelning

Med sekventiell uppdelning menas att de olika datainsamlingsmetoderna används eller gör entré vid olika faser i datainsamlingen. De är alltså inte aktuella samtidigt. Syftet med denna variant är framförallt

att minska bortfallet genom potentialen att fånga fler respondenter med mer än en datainsamlingsmetod. Syftet kan också vara ekonomiskt, att inleda med en billigare metod och använda dyrare metoder först i senare sekvenser av datainsamlingen. Ett exempel på sekventiell mixed-mode är att inleda med papper och använda telefonintervjuer i ett senare skede (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008).

Nackdelen är att mätningen kan skilja *mellan* respondenterna i de olika sekvenserna som då svarat via olika datainsamlingsmetoder (till exempel mellan de som svarat tidigt och sent).

Parallell uppdelning

Parallell uppdelning betyder att de olika datainsamlingsmetoderna används samtidigt. Ett exempel är en pappersenkät med webbalternativ som erbjuds redan i första kontakten. Syftet kan vara goodwill, öka svarsfrekvensen och minska bortfallsfel (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008).

I denna uppdelning varierar datainsamlingsmetoden *mellan* respondenter. Om respondenterna ges möjlighet att välja är även eventuella effekter av datainsamlingsmetod och av valet i sig (så kallade selektionseffekter) sammanblandade.

Formuläruppdelning

En annan typ av kombinerad insamling är att använda olika metoder för olika delar eller mått i frågeformuläret. Exempelvis kanske vissa delar samlas in via intervju och andra via enkät. I besöksintervjun är det enkelt att implementera eftersom den innebär ett fysiskt möte. I övrigt är denna uppdelning inte särskilt vanlig. Ett skäl är att det ofta kräver flera kontakter med respondenterna och det vill man helst undvika ur bortfallssynpunkt. Vid speciella typer av undersökningar, där respondenterna är väl förankrade, kan det dock fungera (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008).

Syftet med att använda formuläruppdelning är främst att förbättra mätningen men det kan också finnas praktiska skäl. I formuläruppdelning av mixed-mode varierar alltså inte datainsamlingsmetoden *mellan* utan *inom* respondenter.

Urvalsuppdelning

En fjärde variant av kombinerad insamling är att använda olika datainsamlingsmetoder för olika delar av urvalet. Det kan till exempel innebära att man använder webb för yngre och telefonintervju för äldre. På ett mer övergripande plan kan också internationella studier anses ha denna uppdelning, att data samlas in med olika metoder (till exempel telefon eller besök) i olika länder på grund av tradition eller praktiska omständigheter (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008).

Syftet med denna uppdelning kan vara att öka svarsfrekvensen genom att använda den lämpligaste insamlingsmetoden för respektive del av urvalet. Syftena kan även vara att förbättra mätningen (vissa kan till exempel behöva stöd av en intervjuare) eller rent praktiska.

I denna uppdelning varierar datainsamlingsmetoden *mellan* respondenter.

Omgångsuppdelning

En femte variant är att använda olika insamlingsmetoder för olika omgångar i en panelundersökning. Det kan till exempel innebära att första insamlingen sker via telefon men resterande via enkät. En tidigare vanlig variant var att inleda med besök men att gå över till telefon i senare omgångar. Syftet kan vara att först kontrollera eller införskaffa kontaktinformation (postadresser, telefonnummer) eller att etablera en personlig relation (kunna övertala, svara på funderingar etc.) vid den första omgången men sedan använda en billigare insamlingsmetod när respondenterna väl är förankrade i studien (de Leeuw, Hox & Dillman, 2008).

I denna variant av mixed-mode varierar inte datainsamlingsmetoden mellan utan *inom* respondenter (om inte panelen är longitudinell med rotation för då varierar det mellan första rotationsgruppen och övriga). Däremot varierar det inom individer över tid och skillnader på grund av tidsmässiga aspekter är sammanblandade med datainsamlingsmetoden.

Sammanfattande punkter:

- Mixed mode betyder att mer än en datainsamlingsmetod används.
- Skäl att kombinera datainsamlingsmetoder är främst att minska bortfall, bortfallsfel och kostnader.

- Det finns olika sätt att blanda insamlingsmetoder (till exempel för olika omgångar i en panelundersökning eller för olika delar av urvalet).
- De vanligaste och enda som kan betraktas som allmänna datainsamlingsstrategier är: parallell mixed-mode (metoderna är aktuella samtidigt) och sekventiell mixed-mode (metoderna är aktuella i olika sekvenser men inte samtidigt).

Skillnader mellan insamlingsmetoder

Påverkan på mätningen

Det kan finnas flera risker med att blanda insamlingsmetoder eller att tillföra en ny insamlingsmetod till en gammal etablerad. Om den nya metoden inte är tillräckligt planerad och testad kan det till exempel uppstå problem med tekniken, organisationen av fältarbetet och administrationen av själva frågeformuläret.

Förutom ovanstående exempel kan det också finnas metodeffekter ("mode effects" på engelska) i data (Biemer & Lyberg, 2003). Med det menas ibland allt som kan ge skillnader mellan datainsamlingsmetoder och inkluderar då till exempel skillnader i bortfall. Men det kan också syfta på att datainsamlingsmetoden i sig har en effekt på mätningen, allt annat lika. Olika datainsamlingsmetoder har sina egna felstrukturer och är inte nödvändigtvis kompatibla. En konsekvens om data från olika datainsamlingsmetoder sammanförs är att skillnader på grund av datainsamlingsmetod kan tolkas som faktiska skillnader. Det kan även finnas andra konsekvenser än att felaktigt tolka metodskillnader som faktiska. Om man till exempel har en ordinarie metod som är bra på att mäta något och sedan tillför en ny metod som är sämre på att mäta detta, försämras mätningen totalt sett om data från de två metoderna sammanförs jämfört med om endast den ordinarie metoden använts, allt annat lika (till exempel samma andel och typ av respondenter). Det omvända är förstås också möjligt men kanske mer ovanligt eftersom den ordinarie metoden sannolikt är ordinarie just för att den är särskilt lämplig för att mäta det som undersökningen syftar till.

En annan aspekt är att frågor inte kan ställas på samma sätt i olika datainsamlingsmetoder utan att man måste använda olika lösningar. Det är en utmaning om till exempel webb- eller pappersenkät ska kombineras med telefonintervju. Hur ska till exempel telefonfrågor administreras på

webben? Det är inte ovanligt att olika lösningar används, vilket förstås kan påverka mätningen. Detta är egentligen ingen metodeffekt, om man ska vara renlärig, eftersom det inte bara är datainsamlingsmetoden som skiljer sig åt i de fallen. Det är också två olika lösningar som jämförs. Konsekvensen blir dock densamma: att data inte nödvändigtvis är riktigt jämförbar.

Mätningen i kombinerad insamling kan alltså skilja sig åt på två sätt:

1. Datainsamlingsmetoden i sig har en påverkan, allt annat är lika (metodeffekt).
2. Datainsamlingsmetoderna kräver olika lösningar vilka kan leda till olika mätningar.

Ibland kallas både punkt 1 och 2 för metodeffekter eller mode-effekter. Det är en lite onödig förenkling eftersom distinktionen dem emellan kan vara viktig: punkt 1 är svår att påverka eller göra så mycket åt medan punkt 2 ligger mycket mer inom undersökningsorganisationens egen kontroll. Dessutom indikerar en del studier att skillnader i frågeställandet (2) kan ha en större effekt än datainsamlingsmetoden (1) (Revilla & Saris, 2012). Det kan alltså vara värt att hålla isär dessa (1 och 2).

Sammanfattande punkter:

- Det finns många faktorer som kan leda till skillnader mellan datainsamlingsmetoder (till exempel bortfall eller självselektering).
- Med metodeffekt (eller mode-effekt) menas oftast att datainsamlingsmetoden i sig har en effekt på mätningen, allt annat lika.
- Olika datainsamlingsmetoder kan kräva olika lösningar vilket kan leda till olika mätningar.

Övergripande strategier om formulärdesign i mixed-mode

Införandet av mixed-mode bör föregås av risk-vinst-överväganden om kostnader, olika typer av fel (till exempel bortfall och mätning) och hur de påverkas, andra vinster som till exempel goodwill och vilka praktiska möjligheter som finns. Om ekvationen hamnar på plus kan mixed-mode

vara ett bra alternativ och det finns då i huvudsak fyra möjliga utgångspunkter (de Leeuw, 2009; Dillman, Smyth & Christian, 2009):

1. **Unimode design:** I denna variant ses inte datainsamlingsmetoderna som separata utan som gemensamma. Samma (eller så likvärdiga som möjligt) frågeformulär ska administreras, oavsett datainsamlingsmetod. Frågeformuläret ska vara baserat på gemensamma nämnare mellan metoderna. I praktiken kan det innebära att alla datainsamlingsmetoder är suboptimerade, men istället så lika varandra som möjligt. Idén är att detta minimerar metodeffekter och då gör data mer jämförbar.
2. **Mode-specifik design:** Enligt detta synsätt tillåts förändringar av frågeformuleringar eller format om det krävs för att kommunicera samma sak i olika datainsamlingsmetoder.
3. **Mode-förstärkt design (kallas ibland för mode maximering):** I denna optimeras respektive metod för sig utan direkt hänsyn till de andra. Målet är att göra den bästa möjliga mätningen i respektive metod. Idén är att data då får minst fel, totalt sett. Det kan till exempel innebära att använda webbens tekniska funktionalitet (som logiska kontroller) trots att det inte kan motsvaras i pappersenkäten. Eller att ha många svarsalternativ i webbenkäten trots att det kräver en annan lösning i telefonintervjun.
4. **Generalized mode design:** Enligt denna utgångspunkt är frågeformuleringarna i sig inte i fokus utan den kognitiva reaktion de orsakar hos respondenten. Det är alltså den kognitiva reaktionen som är det viktiga, inte frågeformuleringen i sig, och det kan krävas olika frågeformuleringar i olika datainsamlingsmetoder för att få samma kognitiva reaktion.

Punkterna 2, 3 och 4 överlappar varandra och beskrivs lite olika av olika forskare. De uttrycker ändå en variation över hur stora skillnader som tillåts eller eftersträvas mellan mätinstrumenten.

Vilken design är lämpligast?

Vilken av de olika utgångspunkterna är lämpligast? Först kan man konstatera att generalized mode design kräver mycket djupa insikter i undersökningen och dess svarsprocess. Därför är det en strategi som man nog bör vara väldigt försiktig med att använda, speciellt vid ovana att designa frågeformulär för mixed-mode.

Annars beror valet av design delvis på vilken roll de olika datainsamlingsmetoderna har. Det finns framförallt två lägen (de Leeuw, 2009; de Leeuw, Hox & Dillman, 2008):

- a) **En datainsamlingsmetod är dominant:** I det ena läget är en av datainsamlingsmetoderna dominant i den betydelsen att den är viktigare än den andra datainsamlingsmetoden. Det kan till exempel vara så att den ena metoden är ordinarie och den andra mer av ett komplement. Detta är en vanlig situation om en undersökning som tidigare endast har använt en datainsamlingsmetod vill lägga till en metod som ett komplement, till exempel av goodwill-skäl.
- b) **Alla datainsamlingsmetoder är lika mycket värda:** Det andra läget är att alla metoder är lika viktiga. Exempelvis vid en engångsundersökning som ska använda mixed-mode eller om en undersökning ska tillföra en datainsamlingsmetod men inte som komplement utan som ett fullvärdigt alternativ.

Om (a) – en datainsamlingsmetod är dominant – är det ganska naturligt att målet är att mätningen i den dominanta metoden blir så bra som möjligt, att den är optimerad i stil med (3) mode-förstärkt design. Den är då prioriterad i formulärdesignen och påverkas inte av eventuella begränsningar i de andra metoderna. Vad gäller designen för de andra icke-prioriterade datainsamlingsmetoderna kan ett mål vara att efterlikna den dominanta metodens frågeformulär i så stor utsträckning som möjligt. Enligt denna design används då både mode-förstärkt design (den dominanta metoden) och unimode (fast utifrån den dominanta metodens villkor). Den dominanta metoden är alltså optimerad, men de andra kan vara suboptimerade. Ett annat mål för de icke-prioriterade metoderna kan vara att använda mode-specifik design för dem, till exempel om man bedömer att det behövs för att kommunicera samma mening som i den dominanta metoden.

Vad gäller (b) – alla datainsamlingsmetoder är lika mycket värda – så finns det inte samma givna utgångspunkt för frågeformuläret som när en metod är dominant. I detta läge rekommenderar Dillman, Smyth och Christian (2009) att man använder unimode. De Leeuw, Hox och Dillman (2008) skriver att mode-specifik design främst är relevant om man vill göra skattningar på hela svarsmängden men att unimode är att rekommendera om man planerar att göra subgruppsanalyser (till exempel skillnader mellan gamla och unga) eftersom olika grupper kan

välja att svara via olika datainsamlingsmetoder och att gruppskillnader då kan vara sammanblandade med metodeffekter.

Utgångspunkterna 1–4 ska ses som ledstjärnor. De kan vara svåra att tillämpa helt renlärt i praktiken. Om unimode är utgångspunkten för mixed-mode i webb och intervju kan det säkert ändå krävas en del mode-specifik design för att kompensera webben för intervjuarens roll i intervjun. USA:s censusbyrås guidelines (2007) är ett bra exempel på tillämpning av unimode i praktiken. I slutänden handlar det om bedömnings- och designfrågor. Helt klart är i alla fall att unimode förmodligen är det tryggaste tillvägagångssättet. Det kan krävas att man har på fötterna om man ska börja ändra frågeformuleringar alltför mycket.

Sammanfattande punkter:

- Det finns flera olika utgångspunkter för formulärdesignen i mixed-mode, till exempel unimode eller mode-förstärkt design.
- Utgångspunkterna varierar beroende på hur mycket skillnader mellan frågeformulären de tillåter.
- Om en datainsamlingsmetod är viktigare än de andra bör den vara styrande och optimeras.
- Om alla datainsamlingsmetoder är lika viktiga rekommenderas generellt unimode som utgångspunkt.

Varför skillnader?

Varför kan det finnas skillnader mellan datainsamlingsmetoder?

Varför kan det finnas skillnader mellan datainsamlingsmetoder? Skälen kan sammanfattas i några punkter (de Leeuw, 2008):

- **informationstyp i frågeformuläret** – visuell eller ljudmässig
- **intervjuaren** – närvaron eller frånvaron av en intervjuare
- **beteendemönster** förknippade med mediet i datainsamlingsmetoden (till exempel datorn/webben eller telefonen)
- **bortfall och datainsamlingsmetodens täckningsgrad**
- **självelektering.**

Med beteendemönster menas till exempel att personers erfarenheter och beteenden kan variera mellan olika medier. Kanske är personer mer okoncentrerade vid datorn (eller i telefonen) än när de besvarar en

pappersenkät? En annan aspekt är att en surveyförfrågan kan uppfattas som olika seriös i olika medier, till exempel på grund av konkurrens från spam (webben) eller telemarketing (telefon). Generellt kan man säga att denna punkt inte är särskilt väl utforskad och kanske inte innebär någon större källa till skillnader mellan datainsamlingsmetoder i praktiken (de Leeuw, 2009).

De två första punkterna har däremot stor betydelse för både eventuella metodeffekter och för hur frågeformuläret ska eller kan designas. I en intervju *lyssnar* respondenten på intervjuarens läsning av frågeformuläret. I en enkät *läser* respondenten själv frågeformuläret. I en intervju är dessutom en annan människa, intervjuaren, inblandad i besvarandet och administrationen av frågeformuläret närmar sig ett vanligt samtal. Detta är grundläggande skillnader mellan framförallt själv- och intervjuaradministrerade frågeformulär och anses utgöra grunden för de flesta metodeffekter.

Sammanfattande punkter:

- De främsta orsakerna till eventuella skillnader har med informationstyp (visuell eller ljudmässig) och intervjuaren att göra.
- Störst risk för metodskillnader finns alltså om man blandar självadministrerade och intervjuaradministrerade metoder.

Mixed-mode – datainsamlingsnivå

Bortfall

Svarsfrekvens

Svarsfrekvenserna är sjunkande i västvärlden, oavsett datainsamlingsmetod. Det är många saker som påverkar svarsfrekvensen och det kan därför vara svårt att dra generella slutsatser. Det kan ändå vara intressant hur mycket datainsamlingsmetoderna producerar jämfört med varandra som enskilda metoder (inte mixade). En översiktsartikel konstaterade att webbenkäter generellt hade 10 procent lägre svarsfrekvens än pappersenkät eller telefonintervjuer. (Fan & Yan, 2010). En annan studie visade att fungerande faktorer för att öka svarsfrekvensen på webben är ungefär desamma som för pappersenkät, t.ex. förbrev, personalisering och påminnelser (helst via post, inte e-post) (Sanchez-Fernandez, Munoz-Leiva & Montoro-Rios, 2012; Cook, Heath & Thompson, 2000). Webben

hade alltså generellt sett lägre svarsfrekvenser som ensam data-insamlingsmetod. Detta är dock något som kan förändras över tid. På SCB är en erfarenhet att om man inleder med webbenkät (i en sekventiell mixed-mode ansats, inte som ensam metod) är det alltfler som svarar via webben.

Telefonintervjuer var tidigare förknippat med höga svarsfrekvenser (förmodligen högre än enkät) men har stött på bland annat teknologiska hinder såsom nummerpresentatörer, telefonsvarare, IP-telefoni och samtalsblockering. Smartmobilernas entré kan också ha förändrat hur människor använder telefonen. En annan betydande förändring är konkurrensen från telemarketing. Dessa faktorer har troligen förändrat personers telefonbeteende, till exempel hur man förhåller sig till en påringning från ett främmande nummer, och bidragit till kraftigt sjunkande svarsfrekvenser för telefonintervjun. Svarsfrekvenserna för enkäter har också sjunkit under en längre tid (dock inte lika dramatiskt som för telefonintervjuer, sett ur ett internationellt perspektiv; de Leeuw, 2009).

Partiellt bortfall

Intervjuaradministrerade datainsamlingsätt har i regel lägre partiellt bortfall (de Leeuw, Hox & Huisman, 2003). Detta beror förstås på att intervjuaren kontrollerar frågeformuläret, kan hjälpa respondenten att förstå frågan, söka i minnet och så vidare.

Datainsamlingsmetodens täckningsgrad

Om inte alla i urvalet täcks av en datainsamlingsmetod finns det stor risk för bortfall och även bortfallsfel, eftersom de som inte täcks sannolikt är annorlunda än de som gör det. Detta handlar alltså om datainsamlingsmetodens täckning, inte urvalsramens (kallas bland annat för "web access rate", "internet coverage rate" eller "internet penetration rate" på engelska). De som inte har tillgång till internet är till exempel generellt sett äldre och lägre utbildade. Men det handlar inte bara om demografi utan det finns även andra skillnader. De som har tillgång till internet har ofta bättre hälsa och ekonomi (även om man kontrollerar för demografiska faktorer) (Couper, 2009).

Man bör alltså tänka på datainsamlingsmetodernas täckning. Ur ett svenskt perspektiv är det framförallt webben men i viss mån också telefon som är intressant vad gäller täckning. Hur många och vilka

grupper har tillgång till och använder internet? Vilka grupper använder mobiler eller smarttelefoner? Sådan kunskap är viktig vid val av datainsamlingsmetod och mixed-mode design.

Mode-preferenser

En mode-preferens är en attityd – att urvalspersoner kan ha en mer positiv inställning till, det vill säga föredra, att svara via en viss data-insamlingsmetod än via en annan. Redan på 1970-talet undersöktes mode-preferenser och studier visade att det varierade mellan individer. I Groves och Kahns (1979) studie föredrog till exempel 39 procent telefonintervju, 23 procent besöksintervju och 28 procent pappersenkät (Dillman, Phelps, Tortora, Swift, Kohrell, Berck & Messer, 2009).

Det är inte helt lätt att få kunskap om mode-preferenser. Informationen finns förstås inte i register utan först måste den samlas in. Det krävs alltså mer än en insamling för att undersöka vilken betydelse mode-preferensen kan ha på deltagande. Ett relaterat problem är att metoden i den första datainsamlingen kan bli styrande: personer tenderar att gilla den metod de deltar i och de som inte gillar den metoden är i högre grad bortfall. Det vill säga, de som svarar vilken mode-preferens de har är övervägande de som gillar den datainsamlingsmetod som används för att samla in just modepreferenser (Olson, Smyth & Wood, 2012).

Ett annat sätt är att anta att respondenterna visar sina mode-preferenser genom sina val av datainsamlingsmetod. De som väljer till exempel pappersenkät framför webb antas alltså föredra pappersenkät. Detta skulle då betyda att jämförelsevis få föredrar att svara på webben eftersom färre väljer att svara via den metoden (i alla fall enligt tidigare amerikanska studier; Diment & Garrett-Jones, 2007; Millar, O'Neill & Dillman, 2009). En studie visar dock att mode-preferensen (definierad som valen respondenterna gör) kan påverkas. Om man styr respondenterna till en viss datainsamlingsmetod är det fler som föredrar den i framtida insamlingar. Mode-preferens verkar alltså inte vara någon stabil egenskap eller åsikt utan kan påverkas (Millar, O'Neill & Dillman, 2009).

Det finns ett antal nya studier som studerat mode-preferens och deltagande ännu mer i detalj (Olson, Smyth & Wood, 2010; 2012). Mode-preferens är en attityd, att delta är ett beteende. Studierna ovan drar slutsatser om den bakomliggande attityden (mode-preferens) utifrån beteendet (deltagande i viss datainsamlingsmetod) men det är välkänt

inom psykologin att kopplingen mellan attityd och beteende kan vara svag i ett konkret fall där många andra faktorer kan vara relevanta. Dessutom kan mode-preferens bara förväntas ha en effekt på deltagargraden, inte kontaktbarheten. Även om en person föredrar telefonintervju är det inte säkert att personen är lättillgänglig på telefon. I intervjuer kan vi skilja på deltagande/nekande och "ej kontaktbar". I pappers- och webbenkäter går inte det, vilket kan röra till resultatet.

I studien från 2010 undersöktes om personer med en viss mode-preferens i högre grad deltar via den datainsamlingsmetoden än via andra. Resultatet indikerade att det finns en sådan tendens – personer deltar i högre grad om de föredrar datainsamlingsmetoden som används (har den som mode-preferens). För telefonintervju var till exempel svarsfrekvensen 60 för de som hade telefonintervju som mode-preferens jämfört med 49 procent för övriga. De största skillnaderna fanns för webben där svarsfrekvensen för webbpreferensgruppen var 40 procent men bara drygt hälften för övriga (22 %) (skillnaderna för pappersenkät var 58% jämfört med 53% men inte statistisk signifikant). Personer verkar alltså, åtminstone när det gäller telefonintervju och webbenkät, delta i högre grad i en datainsamlingsmetod de föredrar. Svarsfrekvensen på webben var dock lägre än i telefon och pappersenkät även hos dem som angett webben som favoritmetod (40 % på webben, jämfört med 54 % på papper och 47 % i telefonintervju). Sammanlagt hävdar författarna att svarsfrekvensen ökade med 41 procent (inte procentenheter) av att tillmötesgå urvalspersonernas mode-preferenser.

Ovanstående gällde undersökningar med endast en datainsamlingsmetod, inte mixed-mode. Därför undersökte författarna även om personer väljer i linje med sina mode-preferenser när de får möjlighet att välja mellan flera datainsamlingsmetoder och om det då har någon betydelse vilken metod som erbjuds först. Författarna undersökte bara detta för webben och de som hade webben som mode-preferens valde i högre grad webben men bara om det erbjöds först.

Slutsats: Ovanstående studie visar att mode-preferens är en förklaringsfaktor för deltagande i både pappersenkät, webbenkät och telefonintervju samt för valet av metod om fler än en erbjuds. Vad gäller webben bör man dock ta hänsyn till att metoden överlag har lägre svarsfrekvens än pappersenkät och telefonintervju. Detta kan dock förändras över tid.

Sammanfattande punkter:

- Studier visar att respondenterna kan ha mode-preferenser men det är klurigt att mäta dessa eftersom den metod som används för datainsamlingen styr resultatet.
- Mode-preferensen kan enligt viss forskning påverkas.
- Annan forskning visar att mode-preferens är en förklaringsfaktor för deltagande i en viss datainsamlingsmetod och för valet av metod om fler än en erbjuds.

Att låta respondenterna själva välja metod

Vid parallell mixed-mode kan man låta respondenterna själva välja datainsamlingsmetod. Även i sekventiell mixed-mode kan respondenterna implicit välja, om man upplyser dem om att en annan datainsamlingsmetod kommer att erbjudas senare. Om man låter respondenter välja datainsamlingsmetod kommer förstås metoden att variera *mellan* olika respondenter. Eventuella metodeffekter är också sammanblandade med eventuella selektionseffekter (läs mer nedan). Om undersökningen är longitudinell kan respondenterna emellertid komma att välja en annan datainsamlingsmetod vid ett senare tillfälle. Då varierar datainsamlingsmetoden också *inom* individer, över tid. Om man ska mäta förändringar över tid är detta värt att tänka på (de Leeuw, Dillman & Hox, 2008; Dillman, Phelps, Tortora, Swift, Kohrell, Berck & Messer, 2009).

Studier visar att urvalspersoner kan föredra vissa datainsamlingsmetoder framför andra (se ovan). En enkel slutsats kan då vara att om man erbjuder flera datainsamlingsmetoder så borde man kunna tilltala fler och få ett högre deltagande. Det finns dock en del forskning som visar att erbjuda både pappers- och webbenkät samtidigt kan vara förknippat med till och med lägre svarsfrekvens än att bara erbjuda pappersenkät. En meta-studie (utvärdering av flera enskilda studier) visade att erbjuda webb- och postenkät samtidigt är relaterat till en signifikant lägre svarsfrekvens än att endast erbjuda pappersenkät (Medway & Fulton, 2012).

Hur ska det förstås? Enligt ett psykologiskt perspektiv har varje möjlighet en psykologisk kostnad förknippad med sig. Om möjligheterna är fler än en måste individen dessutom jämföra för- och nackdelar mellan de olika möjligheterna. Detta är ett ansträngande moment i sig men kan

också få möjligheterna att framstå som mindre attraktiva än om de presenterats var och en för sig (Millar & Dillman, 2011). En annan anledning kan vara att en del av dem som i beslutssituationen väljer webben av olika anledningar sedan inte kommer sig för att faktiskt delta (slå på datorn, surfa till rätt webbplats och logga in) utan avbryts någonstans på vägen (Medway & Fulton, 2012). Ett tredje skäl kan vara att om man påtrycker att urvalspersonen har ett val så använder de den valmöjligheten men väljer att inte delta alls (Couper, 2009).

Sammanfattande punkter:

- Om man låter respondenterna välja datainsamlingsmetod så är eventuella mode-effekter sammanblandade med eventuella selektionseffekter.
- Att erbjuda flera metoder samtidigt är inte förknippat med ökad svarsfrekvens utan ibland till och med tvärtom.
- Att välja datainsamlingsmetod kan dock vara en bra lösning för täckningsproblem med en datainsamlingsmetod.

Skillnader i svarsmängd – selektionseffekter och mode-effekter

Om man erbjuder olika datainsamlingsmetoder kan det mycket väl hända att man får olika typer av svarsmängder och fördelningar för de olika metoderna (p.g.a. olika bortfall till exempel). Om man tillåter respondenterna att själva välja datainsamlingsmetod kan även det leda till skillnader i svar eftersom de personer som valt att svara via respektive datainsamlingsmetod kan vara olika. Detta kallas för en selektionseffekt. Det är inget problem i sig utan kan tvärtom vara ett av syftena med att kombinera metoder – att få svar från andra typer av personer. Men det kan försvåra möjligheterna att utvärdera eventuella metodeffekter i data. Eventuella metodeffekter är nämligen sammanblandade med eventuella selektionseffekter. Det går med andra ord alltså inte att avgöra om eventuella skillnader beror på datainsamlingsmetoderna eller på att olika typer av personer har svarat på respektive metod.

Hur kan man då särskilja sådana så kallade selektionseffekter från mode-effekter? Egentligen går det inte utan att ändra designen, men man kan göra det bästa möjliga av situationen. Det finns flera sätt.

Ett sätt är att låta respondenterna svara via båda datainsamlingsmetoderna (inom-individ design). Detta kan dock vara svårt att få till i en riktig undersökning utan är mer aktuellt för laboratoriestudier.

Ett andra sätt är att använda ett experimentellt upplägg och till exempel köra ordinarie insamlingsmetod parallellt med en mixed-mode variant. Olika bortfallstrukturer mellan grupperna kan dock ställa till med bekymmer.

Ett tredje sätt är att jämföra insamlad data med validerande data, någon form av gold standard som man har hög tilltro till. Tyvärr finns sällan sådana validerande data tillgängliga (de kan dock samlas in för just detta syfte, till exempel genom återkontakter).

Ett fjärde sätt är att inkludera datainsamlingsmetoden i analysen för att kontrollera dess påverkan på resultatet. Det är emellertid oftast ingen bra strategi eftersom valet av datainsamlingsmetod kan vara relaterat till det man vill analysera (Lorenc, 2010).

Ett femte sätt är att använda statistikmodellering för att göra svarmängderna jämförbara (till exempel viktning) eller för matchning (PSM – "propensity score matching") (Lugtig, Lensvelt-Mulders, Frerichs & Greven 2011; Lorenc, 2010). Det finns dock alltid en risk att grupperna skiljer sig åt bortom de variabler man använder som underlag.

Sammanfattande punkter:

- Selektionseffekt är att olika typer av personer väljer att svara via olika datainsamlingsmetoder och det därför kan bli skillnader i svaren.
- Selektionseffekter är inte ett problem i sig men försvårar möjligheterna att utreda eventuella metodeffekter i data. Det finns olika strategier för att hantera detta.

Svarsfrekvenser – mixed-mode

Även om svarsfrekvenserna har sjunkit över tid oavsett datainsamlingsmetod, kan det variera ganska kraftigt mellan olika undersökningar och inom samma datainsamlingsmetod. Det är svårt att fastställa exakt varför undersökningar får en viss svarsfrekvens. Det är många saker som påverkar, till exempel urvalsgruppen, förhandsmaterial, ämnesområdet, formulärdesignen, kontaktstrategier, påminnelser, belöningar,

tidpunkten för insamling och organisationen bakom undersökningen. Många av dessa är dessutom svåra att kvantifiera (till exempel formulärdesign, ämnesområdet eller en belönings värde).

Denna variation gör det lite svårt att förhålla sig till litteraturen om svarsfrekvenser, speciellt när det gäller mixed-mode som inkluderar ännu fler aspekter som kan variera jämfört med en undersökning som endast har en datainsamlingsmetod. Man ska inte dra för stora växlar av ett enskilt resultat och om något verkar intressant bör man titta på studiens utförande i detalj för att kunna dra säkrare slutsatser om generaliserbarheten.

Det finns dock några relativt tydliga linjer i litteraturen. En viktig sådan är skillnaderna i svarsfrekvens mellan parallell och sekventiell mixed-mode (redovisas nedan). Vad gäller studierna som redovisas under avsnittet om sekventiell mixed-mode nedan bör de ses mer som exempel än riktlinjer eftersom det är så många faktorer som kan variera.

Parallell mixed-mode

Att erbjuda pappers- och webbenkät samtidigt verkar inte öka svarsfrekvensen utan kan till och med minska den enligt vissa studier (Medway & Fulton, 2012) (läs mer i avsnittet Att låta respondenterna själva välja metod ovan).

Sekventiell mixed-mode

Att byta datainsamlingsmetod i en andra sekvens av datainsamlingen verkar öka svarsfrekvensen (Dillman, Phelps, Tortora, Swift, Kohrell, Berck & Messer, 2009). Flera studier visar att sekventiell mixed-mode är ett framgångsrikt sätt att minska bortfallet i både undersökningar riktade till allmänheten och till specifika populationer (de Leeuw, Hox & Dillman, 2008).

Telefon och webb

I en studie ökade en webbuppföljning till en telefonintervju svarsfrekvensen från 67 till 84 procent och uppföljningen ledde också till en mer representativ svarsmängd (Greene, Speizer & Wiitala, 2008). Även den motsatta sekvensen, att inleda med webbenkät och följa upp med telefonintervju, ökade svarsfrekvensen. Det är dock värt att påpeka att urvalet till experimentet gjordes bland dem som redan hade besvarat en webbundersökning, det vill säga ett webbvänligt urval.

Pappers- och webbenkät

Några studier visar att man genom att inleda med webbenkät och inkludera pappersenkät först i senare sekvenser kan nå samma slutliga svarsfrekvens som vid endast pappersenkät men få fler svarande på webben. Detta är alltså inte ett sätt att höja svarsfrekvensen men resurser kan sparas genom att styra urvalspersonerna till den billigare webbenkäten (Millar & Dillman, 2011; Holmberg, Lorenc & Werner, 2010).

En fördel med postkontakt är att man kan inkludera belöningar och enkäten är i samma medium som kontakten i sig (vilket inte gäller om login till en webbenkät skickas ut via post) (Millar & Dillman, 2011).

Hur kan man använda samma medium i kontakten till webbenkäter? Ett sätt är att kontakta via e-post. Länken till enkäten kan då finnas direkt i e-postmeddelandet vilket är mycket smidigare för respondenten än att behöva starta datorn, skriva in en webbadress och logga in. Det krävs dock att man har respondenternas e-postadresser. I en panelundersökning kan det vara möjligt att samla in sådana (Millar & Dillman, 2011).

Mode-effekter på datainsamlingsnivå

Metodeffekter – övergripande om skillnader

Hur stort problem är metodeffekter i allmänhet? De Leeuw, Hox och Dillman (2008, s. 300) skriver: "In short: mode effects do exist but tend to be small in well-conducted surveys." Hur stora eventuella problem blir beror förstås på i vilken grad man har beaktat och försökt att parera dem i designarbetet. De mest påtagliga effekterna är sociala till naturen (social önskvärdhet, hålla-med-tendenser etc.), knutna till intervjuadministrerade datainsamlingsmetoder och ger utslag vid framförallt känsliga frågor. Detta kan betyda att för andra typer av undersökningar som inte är känsliga, kan det faktum att datainsamlingsmetoderna kräver olika lösningar vara ett större "problem" än rena metodeffekter. En del studier visar även större skillnader på grund av olika frågeformat än av själva datainsamlingsmetoden (Revilla & Saris, 2012).

Metodeffekter självadministrerat – intervjuadministrerat

Det finns många studier som visar att personer i intervjuer ger mer positiva och nöjda svar (till exempel mindre ensamma, bättre hälsa) än dem i enkät (Dillman, Phelps, Tortora, Swift, Kohrell, Berck & Messer, 2009; Dillman & Mason, 1984; Tarni & Dillman, 1992) och att det verkar

gälla för många olika typer av svarsskalor (olika antal skalsteg och skalformat) (Christian, Dillman & Smyth, 2008).

Samma tendens finns för det som kallas "acquiescence", det vill säga en tendens att hålla med om något snarare än att ta avstånd från det. Den är starkare i telefonintervjuer än i självadministrerade metoder (Greene, Speizer & Wiitala, 2008; de Leeuw, 1992). Intressant nog verkar tendenserna att hålla med (Mu, 1999) eller ge positiva svar (Tourangeau, Miller-Steiger & Wilson, 2002) också vara starkare i telefonintervjuer än IVR ("interactive voice response"). Detta tyder på att det är just intervjuaren, att svaren ges till en riktig person, som är orsaken till tendenserna.

Många studier bekräftar också att telefonintervjun är förknippad med mer socialt önskvärda svar än självadministrerade sätt – att respondentens svar följer samhällsnormer snarare än sanningen. En studie visar att detta även kan gälla saker som inte nödvändigtvis ses som särskilt privata eller känsliga men där det likväl kan finnas normer, till exempel livsstilsfrågor om diet och motion (Greene, Speizer & Wiitala, 2008).

Intervjuer är också förknippade med mer effekter av kontext och frågornas ordningsföljd, kanske för att det inte går att gå tillbaka och ändra tidigare svar som i en pappersenkät.

Det finns även andra skillnader mellan intervju och självadministrerad enkät. En fråga om civilstånd ledde exempelvis till olika svar (flera procentenheters skillnad) på webben jämfört med telefon, sannolikt för att frågan är öppen på telefon medan svarsalternativen syns på webben (Dillman & Smyth, 2007). Detta är en stor och viktig skillnad. Dillman och Smyth (2007, s.94) skriver:

"It is simple to copy and paste questions from telephone survey software into web design software, but one cannot assume that such questions communicate the same information to respondents. In moving a question from an aurally- to a visually-based medium, the meaning and/or expectations communicated by the question may be altered".

Metodeffekter pappers- och webbenkäter

Pappers och webbenkäter är båda självadministrerade metoder och lika varandra, därför är risken för metodeffekter lägre. Några resultat som bekräftar detta: 1) Inga skillnader mellan pappers- och webbenkät vad

gäller negativa svarsmönster (hålla-med-tendens, välja första alternativet). 2) Inte mer satisficing i webb- än i pappersenkäter (Clement & Shamshiri-Petersen, 2013). 3) Inga skillnader mellan pappers- och webbenkät på känsliga frågor om alkoholkonsumtion till collestudenter (McCabe, Diez, Boyd, Nelson & Weitzman, 2006).

Sammanfattande punkter:

- De mest påtagliga metodeffekterna är sociala till naturen (social önskvärdhet, hålla med-tendenser etc.), knutna till intervjuadministrerade datainsamlingsmetoder och ger utslag vid framförallt känsliga frågor.

Nedanstående tabell visar övergripande vilka styrkor och svagheter som finns med de olika datainsamlingsmetoderna vad gäller mätning (andra aspekter som exempelvis kostnader och bortfall ingår inte). En utgångspunkt i tabellen är att både telefonintervjun och besöksintervjun genomförs med datorstöd. En annan utgångspunkt är att besöksintervjun kan använda svarskort.

Tabell 6.1

Kriterium	Pappers- enkät	Webb- enkät	Telefon- intervju	Besöks- intervju
Komplext ämne	-	-	+	++
Känsliga frågor	++	++	-	+
Många svarsalternativ	+	+	-	++
Komplicerad struktur, hopp & filter	-	+	++	++
Visuella hjälpmedel	+	++	-	+
Kunskapsfrågor	-	-	+	+
Betänketid	++	++	-	+
Uppgifter från register, kvitton etc.	+	++	-	-
Respondenter med olika behov (av t.ex. instruktioner)	-	+	++	++
Standardiserade frågor	++	++	-	-
Långt frågeformulär	+	+	-	++
Översättnings- eller tolkbehov i urvalsgruppen	-	-	+	++
Direkt kontroll av svaret	-	+	++	++
Öppna frågor	-	-	+	++

- Mindre bra förutsättningar med datainsamlingsmetoden
- + Bra förutsättningar med datainsamlingsmetoden
- ++ Mycket bra förutsättningar med datainsamlingsmetoden

Sammanfattning

Datainsamlingsmetod – sammanfattning

Olika datainsamlingsmetoder:

- Det finns olika typer av datainsamlingsmetoder. De vanligaste är
 - självadministrerad enkät (papper eller webb)
 - intervju (telefon eller besök).Dessa skiljer sig åt på olika sätt.

Självadministrerat eller socialt samspel:

- Enkät: Respondenten kan besvara enkäten när det passar men vi är samtidigt beroende av att respondenten kan klara det på egen hand. Situationen är mer anonym vilket kan leda till mer ärliga svar men också att det är svårare att påverka respondenten, till exempel genom att beskriva undersökningen och öka motivationen.
- Intervju: Intervjuaren kan hjälpa respondenten med frågorna och sköter administrationen men besvarandet är inte privat. Intervjuaren kan motivera respondenten.
- Det finns olika intervjutekniker:
 - standardiserad intervju följer frågeformuläret ord för ord och intervjuaren hjälper inte till
 - konverserande intervju är mer i form av ett utredande samtal. Båda teknikerna har sina för- och nackdelar. Den konverserande intervjun kan förbättra mätningen men kräver mycket mer resurser. Det är dock svårt att standardisera fullt ut. De flesta intervjuer innehåller i praktiken delar av båda teknikerna.
- Respondenterna svarar ibland oärligt p.g.a. självförnekelse eller för att de vill framstå i bättre dager (det senare kallas för social önskvärdhet).

- Denna tendens är svagare i datainsamlingsmetoder där respondenten besvarar frågeformuläret privat (pappers- och webbenkäter) jämfört med intervjuer.
- Beror skillnaderna mellan datainsamlingsmetoderna på social närhet? I sådana fall borde besöksintervjun visa störst problem med känsliga frågor men så är det inte. Effekterna av känsliga frågor är störst i telefonintervjun.
- Risken för satisficing är större i telefonintervjun än i besöksintervjun.

Informationstyper – skrift eller ljud:

- Olika datainsamlingsmetoder riktar sig till olika sinnen hos respondenten:
 - Självadministrerat innebär att respondenten läser med hjälp av sin syn.
 - Intervju innebär att respondenten lyssnar med sin hörsel.
- Vid intervjuer kan inte respondenten kontrollera informationsflödet, till exempel läsa om frågan vid behov. Därför är det av stor vikt hur mycket respondenten kan "hålla i huvudet" samtidigt i det så kallade arbetsminnet.

Långa frågor och intervjuer:

- Intervjufrågor måste vara kortfattade annars kan information missas.

Svarsalternativen i intervjuer:

- Svarsalternativen vara eller icke vara är ett viktigt beslut i intervjuer eftersom de tar tid att läsa och det blir mer för respondentens att "hålla i huvudet".
- Man kan använda förgreningsteknik för att förenkla läsningen och förståelsen. Dock är inte forskningen entydigt positiv till detta.
- Om svarsalternativen *inte* ska läsas måste frågan i sig vara tillräckligt specifik så att den leder respondenterna till de giltiga svaren.
- Om frågan är öppen och intervjuaren ska klassa respondentens svar behövs det bra stöd, speciellt om det sker i fält.

Att använda annan information än frågeformuläret:

- Undersökningar som kräver externa hjälpmedel för eller beräkningar av respondenten genomförs bättre självadministrerat än över telefon.
- Att utgå från en kalender och respondentens historia kallas för kalenderintervju och kan förbättra minnet. Det finns även andra, liknande tekniker.
- Beroende intervju är ett sätt att utnyttja tidigare insamlad information för att minska administrationstiden och bördan.

Enkäter i smarttelefoner och surfplattor:

- Smarttelefonen är relativt ny som verktyg för att besvara enkäter. Teknologin går fort framåt. Forskningen kan därför ha ett kort "bäst föredatum".
- Smarttelefonen delar egenskaper både med den traditionella webbenkäten och med telefonintervjun.
- Urvalspersonerna kan i smarttelefonen kontaktas via telefonsamtal, sms eller e-post.
- En unik möjlighet är att samla in annan typ av data än vad som är möjligt med t.ex. webbenkäten. Det kan handla om att använda GPS, skanning av streck- och QR-koder eller foton.
- En unik begränsning är att skärmen på en smarttelefon är mycket mindre än skärmen på en dator. Samma sak gäller smarttelefonens tangentbord. Detta ställer krav på enkätdesignen eftersom risken för fel generellt sett är större än i andra insamlingsverktyg.
- En annan begränsning är att respondentens uppkoppling kan försvinna under besvarandet. Batteriet kan också ta slut.
- Det finns olika sätt att besvara en enkät i en smarttelefon: 1) via vanlig webbsida, 2) mobilanpassad webbsida eller 3) via app.
- Vanlig webbsida är minst krävande för undersökningsorganisationen men mest krävande för respondenten. För en app är det motsatta förhållanden.
- Surfplattan har större skärm och tangentbord än smarttelefonen.
- Till skillnad från smarttelefonen används surfplattan bara i hemmet vilket innebär att uppkoppling inte är något problem. Å andra sidan kan inte information utanför hemmet samlas in.

Frågetyper och datainsamlingsmetod:

- Hålla-med-tendenser är starkare i telefonintervjuer vilket innebär att metoden passar sämre för attitydfrågor.
- Attitydfrågor är känsliga för ordalydelsen – intervjuarna bör inte avvika i läsningen.
- I intervjuer, och i viss mån i webbenkäter, kan svarsalternativ för icke-svar som till exempel "Vet inte" erbjudas vid behov.
- Det måste framgå vilken svarsenhet frågan har vid intervjuer (oavsett objektiv eller subjektiv svarsskala).
- Kunskapsfrågor där respondenten ska redovisa sin kunskap är ganska ovanliga i surveyundersökning. Frågor där respondenten ska värdera sin kunskap är vanligare (läs mer i kapitel 10).

Mixed-mode (kombinerad insamling):

- Mixed mode betyder att mer än en datainsamlingsmetod används.
- Skäl att kombinera datainsamlingsmetoder är främst att minska bortfall, bortfallsfel och kostnader.
- Det finns olika sätt att blanda insamlingsmetoder (till exempel för olika omgångar i en panelundersökning eller för olika delar av urvalet).
- De vanligaste och enda som kan betraktas som allmänna datainsamlingsstrategier är: parallell mixed-mode (metoderna är aktuella samtidigt) och sekventiell mixed-mode (metoderna är aktuella i olika sekvenser men inte samtidigt).

Skillnader mellan datainsamlingsmetoder:

- Det finns många faktorer som kan leda till skillnader mellan datainsamlingsmetoder (till exempel bortfall eller självselektering).
- Med metodeffekt (eller mode-effekt) menas oftast att datainsamlingsmetoden i sig har en effekt på mätningen, allt annat lika.
- Olika datainsamlingsmetoder kan kräva olika lösningar vilket kan leda till olika mätningar.

Övergripande strategier om formulärdesign i mixed-mode:

- Det finns flera olika utgångspunkter för formulärdesignen i mixed-mode, till exempel unimode eller mode-förstärkt design.
- Utgångspunkterna varierar beroende på hur mycket skillnader mellan frågeformulären de tillåter.
- Om en datainsamlingsmetod är viktigare än de andra bör den vara styrande och optimeras.
- Om alla datainsamlingsmetoder är lika viktiga rekommenderas generellt unimode som utgångspunkt.

Varför skillnader?

- De främsta orsakerna till eventuella skillnader har med informationstyp (visuell eller ljudmässig) och intervjuaren att göra.
- Störst risk för metodskillnader finns alltså om man blandar självadministrerade och intervjuaradministrerade metoder.

Mode-preferenser:

- Studier visar att respondenterna kan ha mode-preferenser men det är klurigt att mäta dessa eftersom den metod som används för datainsamlingen styr resultatet.
- Mode-preferensen kan enligt viss forskning påverkas.
- Annan forskning visar att mode-preferens är en förklaringsfaktor för deltagande i en viss datainsamlingsmetod och för valet av metod om fler än en erbjuds.

Att låta respondenterna själva välja metod:

- Om man låter respondenterna välja datainsamlingsmetod så är eventuella mode-effekter sammanblandade med eventuella selektionseffekter.
- Att erbjuda flera metoder samtidigt är inte förknippat med ökad svarsfrekvens utan ibland till och med tvärtom.
- Att välja datainsamlingsmetod kan dock vara en bra lösning för täckningsproblem med en datainsamlingsmetod.

Skillnader i svarsmängd - selektionseffekter och mode-effekter:

- Selektionseffekt är att olika typer av personer väljer att svara via olika datainsamlingsmetoder och det därför kan bli skillnader i svaren.
- Selektionseffekter är inte ett problem i sig men försvårar möjligheterna att utreda eventuella metodeffekter i data. Det finns olika strategier för att hantera detta.

Mode-effekter på datainsamlingsnivå:

- De mest påtagliga metodeffekterna är sociala till naturen (social önskvärdhet, hålla med-tendenser etc.), knutna till intervjuadministrerade datainsamlingsmetoder och ger utslag vid framförallt känsliga frågor.

7 Språket och kommunikation

Författare: *Andreas Persson (SCB)*

Språket som kommunikationsmedel

Språket är det främsta verktyget vi har för att kommunicera med respondenterna, oavsett om vi använder intervju eller enkät. Därför har det en nyckelroll i en undersökning i allmänhet och i frågekonstruktionen i synnerhet. Det är lätt hänt att bagatellisera den nyckelrollen. I en statistisk undersökning hamnar gärna siffrorna i fokus, men siffrorna baseras ofta på språkliga förfrågningar till respondenterna och de hamnar tyvärr i skymundan ibland.

Det kan finnas flera skäl till det. I stort sett alla behärskar språket, kan läsa och skriva. Den skriftliga kommunikationen med respondenterna kan därför framstå som given, naturlig och oproblematiske. Dessutom kommunicerar vi hela tiden med vänner, arbetskamrater och familj utan större problem. Varför ska det då vara svårt att kommunicera med just respondenterna? Frågan kan kännas berättigad men den innehåller två felaktigheter. För det första är kommunikationen med vänner, arbetskamrater och familj inte problemfri. Det förekommer språkförbistringar och missförstånd där också. Men de allra flesta uppdagas inte och får inte heller några konsekvenser värda någon uppmärksamhet. För det andra kan inte kommunikationen med vänner, arbetskamrater och familj likställas med kommunikationen med respondenterna. När vi kommunicerar med personer vi känner har vi förutom språket en massa annan information tillgänglig som kan kompensera för språkets brister. Vi har bakgrundsinformation om både personen och kanske ämnet ifråga. Därutöver har vi ett sammanhang och en situation, samt ett tonfall och icke-verbal information och möjligheten att genom dialog försäkra oss om att vi har förstått varandra rätt. I kommunikationen med respondenterna saknas ofta den hjälpinformationen och den är därför mer utsatt för språkets brister.

Ett av de vanligaste problemen i frågeformulär är att den som skrivit frågorna har överskattat språket som kommunikationsmedel och inte reflekterat tillräckligt över frågeformuleringarna. Man kan vara insatt i frågeteknik eller expert på hur minnesprocesser fungerar, men om man inte har insikter om språket riskerar man likväl problem med sina frågor.

Frågekonstruktören bör därför helst ha en hög språklig medvetandegrad och fundera över ordval, satsbyggnad, språkförståelse och alternativa tolkningar. Syftet med detta kapitel är att ge insikter om språket och språkförståelse och att höja medvetandegraden kring desamma.

Sammanfattande punkter:

- De allra flesta kan läsa och skriva men det behövs mer kunskap och förmåga än så för att skriva bra frågor.
- Språket har många brister som kanske inte märks så mycket i vardagen men som kan få stort utslag i ett frågeformulär.
- Frågekonstruktören bör ha hög språklig medvetandegrad och fundera över ordval, satsbyggnad, språkförståelse och alternativa tolkningar.

Vad är språket?

Ett gemensamt symboliskt system

Språket är alltså avgörande i frågekonstruktion. För att förstå varför behöver man först fastställa vad språket är och inte är. Även om föregående avsnitt varnade för språkets brister går det inte att bortse från att språket är ett fantastiskt verktyg. Det är svårt att hitta något mer mänskligt, som särskiljer oss människor från andra varelser, än förmågan till och förekomsten av språk.

Vad är då ett språk? Det kan definieras som ett *gemensamt, symboliskt system för kommunikation* (Ashcraft, 1994). Språket är alltså ett symboliskt system. Det betyder att ord används som representant för något, till exempel ett fysiskt objekt, en känsla eller en aktivitet. Oftast finns det ingen given koppling mellan det skrivna ordet, eller ordets ljud i det talade språket, och objektet som ordet representerar. Ordet "katt" varken låter eller ser ut som det djur det representerar. Men det finns undantag. Till exempel låter ordet "pang" ljudmässigt likt den företeelse ordet representerar (en stor smäll). Det förekommer också i barnspråk, till exempel att en hund kallas för "vovve". Men i regel är kopplingen mellan ord och det de representerar svag. Detta bidrar med en flexibilitet till systemet. Ord kan beskriva abstrakta företeelser eller objekt som inte har någon naturlig ljudmässig utgångspunkt. Nya ord för nya företeelser eller objekt kan lätt tillföras systemet och ords betydelse kan utvecklas

med tiden. Detta får anses vara styrkor med språket men flexibiliteten kan också utgöra en svaghet (Ashcraft, 1994).

En svaghet med det symboliska systemet är att ord inte har någon given betydelse utan att det är något vi måste lära oss. Här finns en källa till problem både för kommunikation och för frågekonstruktion. I definitionen ovan beskrevs symbolsystemet som *gemensamt*. I hög grad stämmer detta, annars skulle vi inte kunna kommunicera med varandra alls utan vi skulle bokstavligt tala olika språk, men det är en sanning med modifikation. Olika personer kan ge olika ord olika mening eller betydelse (Ashcraft, 1994). En ordlista är nog det närmaste man kommer definieringar av ords betydelser. Om man till exempel slår upp ordet "juvenil" i Svenska Akademiens ordlista över svenska språket får man förutom information om ordklass och böjningar även veta att ordet betyder ungdomlig eller barnslig (SAOL, 2013). Ordet definieras alltså med hjälp av andra ord som i sin tur definieras av andra ord och så vidare. På sätt och vis är en ordlista ett stort cirkelbevis. Ords betydelse kan alltså variera mellan personer och det finns sällan en "sann betydelse".

Grammatik – språkets regelverk

Hur kan vi då förstå alla dessa kombinationer av ord? Det krävs regler och dessa regler kallas grammatik, som är ett ytterligare kännetecken för språket. Många kanske känner sig osäkra på sina grammatikkunskaper men på ett praktiskt plan är det en helt ogrundad oro. De allra flesta följer reglerna med sin språkkänsla oavsett om man kan beskriva regelverket och alla dess benämningar (som ordklasser) eller inte. Om vi inte följde reglerna skulle det nämligen vara väldigt mycket svårare att kommunicera med varandra än vad det allt som oftast är.

Grammatik kan delas in i bland annat fonologi (språkets ljud), syntax (den språkliga uppbyggnaden, till exempel ordningsföljden av ord), semantik (ords betydelse) och pragmatism (intentionen hos den som säger eller skriver något) (Harley, 2008; Ashcraft, 1994).

En sats syntax påverkar förståelsen av satsen. Om ordföljden ändras kan betydelsen ändras, trots att satsen innehåller samma ord (Ashcraft, 1994).

EXEMPEL:

Mamman gav pengar till barnet.

Barnet gav pengar till mamman.

De två satserna ovan innehåller samma ord men de har olika betydelser eftersom ordningsföljden, och därmed ordens funktion i satsen, är olika. Syntaxen styr alltså hur satser ska förstås. För att undvika missförstånd är det därför viktigt att kontrollera att den text eller de frågor man skriver har en korrekt satsbyggnad.

Syftningsfel är ett exempel på ett annat fel av samma sort. Detta kan tyckas enkelt att uppmärksamma, men vid mer komplicerade satser kan det vara mer förrädiskt än i exemplen ovan. Vid komplicerade satser gäller det alltså att hålla ordning på orden. Ett annat alternativ är att undvika komplicerade satser genom att förenkla språket eller dela upp en lång mening på två kortare. Detta minskar risken för fel och kan också underlätta förståelsen för läsaren.

Språket är inte bara bokstavligt

Ibland framgår inte en sats innebörd i det bokstavliga. En sats pragmatiska betydelse (sändarens intention och mening) kan skilja sig från den bokstavliga (semantiska). Det är så vanligt att vi oftast inte ens reflekterar över att vi gör tolkningar och drar slutsatser som går bortom ordens faktiska betydelse. Detta är dock en central del i kommunikation och nästa avsnitt handlar om just detta.

Sammanfattande punkter:

- Språket är ett gemensamt, symboliskt system för kommunikation.
- Ord har ingen given betydelse utan det är något vi måste lära oss och det kan variera något från person till person.
- Grammatik är språkets regelverk – om det inte fanns regler skulle vi inte kunna förstå varandra.
- Syntax är regler för hur enskilda ord sätts samman i en meningssats, till exempel ordningsföljden av ord.
- Semantik har med ords och språkets betydelse att göra.
- Pragmatik har med intentionen bakom ett yttrande att göra. I det praktiska bruket av språket ligger ofta meningen i ett yttrande bortom det bokstavliga.

Konversationens logik och praktik

Språkets syfte – att kommunicera

Språket är ett kommunikationsverktyg. Clark & Schober (1992, s.15, översatt till svenska) skriver att "... ett vanligt missförstånd är att språket främst har att göra med ord och vad de betyder. Det har det inte. Det handlar främst om människor och vad de försöker säga". Språkets främsta syfte är att kommunicera något till en mottagare. Detta gäller oavsett om det är en formell, skriftlig text med en okänd läsare eller en muntlig konversation med en kompis. Det är meningen och intentionen hos den som säger eller skriver något som är det viktiga, inte orden i sig (Clark & Schober, 1992).

Gemensamma plattformar för tolkning

Sammanhanget som plattform

Människor har sociala mål oavsett om de berättar en historia eller ställer frågor i ett frågeformulär. Språket är ett verktyg för att uppfylla dessa mål. En nyckel till att förstå talarens intention är att talare och mottagare har en gemensam plattform ("common ground" på engelska) som de delar som utgångspunkt för förståelsen. En konversationsprincip är att mottagaren ska kunna förstå talarens yttrande utifrån en gemensam plattform (Clark & Schober, 1992).

Clark & Schober (1992) använder repliken "Två tack" som exempel. Den enkla repliken syftar på helt skilda saker beroende på om talaren står i hissen, i biljettkassan på bion eller pratar i telefonen med bokningsansvarig på vårdcentralen. Trots dessa variationer i möjliga betydelser innebär repliken "Två, tack" förmodligen inget problem i någon av situationerna. Talare och mottagare delar nämligen sannolikt en gemensam plattform utifrån vilken repliken "Två, tack" kan förstås i respektive situation. Finns det utrymme för tolkning i situationen kan det dock leda till missförstånd. För spärrvakten i tunnelbannan kan "två, tack" till exempel tolkas som antal önskade kuponger istället för antal resande, och tvärtom.

Att ha en gemensam plattform är grundläggande i frågekonstruktion. Om frågeställaren inte delar plattform med respondenten kan det leda till missförstånd i allt från nyanser av perifera ords betydelser till centrala begrepps innebörd (till exempel vad menas med "diskriminering"?).

Kultur som plattform

Vad utgörs då den gemensamma plattformen av? Delvis är den kulturellt betingad. Personer från olika kulturer kan förstå samma ordalydelse olika eftersom de tolkar innebörden utifrån olika plattformar. I en studie (Clark, Schreuder & Buttrick, 1983; Clark & Schober, 1992) fick testpersoner se ett fotografi med dåvarande amerikanske presidenten Ronald Reagan och budgetansvarige David Stockman. De fick en av följande två frågor:

1. *Du vet vem den här mannen är eller hur?*
2. *Har du någon aning om vem den här mannen är?*

Det är viktigt att uppmärksamma att varken frågan i sig eller intervjuarens beteende avslöjar vem av de två personerna på bilden som frågan åsyftar. "Den här mannen" är otydligt – det kan syfta på antingen Reagan eller Stockman på fotografiet. Trots denna otydlighet tolkade en stor majoritet frågan som att den syftade på Reagan om fråga 1 ställdes. Om fråga 2 ställdes trodde betydligt fler att den syftade på Stockman. Vad beror detta på? Förmodligen på att frågeställare och testperson delade den kulturella plattformen i vilken Ronald Reagan är en mycket känd person och David Stockman betydligt mindre känd och att frågorna anspelade på detta. "Du vet vem den här mannen är eller hur?" uttrycker att testpersonen borde känna till den person som åsyftas, vilket nog hade varit en orimlig förväntan om det syftat på David Stockman. "Har du någon aning om vem den här mannen är?" uttrycker tvärtom att det är tveksamt om testpersonen känner till den person som åsyftas, vilket på motsvarande sätt nog hade varit orimligt om det syftat på en så känd person som Ronald Reagan. Dessa tolkningar är ytterligare exempel på tolkningsprocesser som går bortom det bokstavliga. Tolkningarna handlar snarare om vad frågeställaren rimligtvis kan mena, givet situationen. Båda tolkningarna är också kulturellt betingande. För personer som inte känner till någon av männen går det inte att dra dessa kulturellt betingade slutsatser om vem "den här mannen" rimligen syftar på (Clark, Schreuder & Buttrick, 1983; Clark & Schober, 1992).

Inom samma övergripande kultur kan även olika personer ha olika plattformar beroende på sina personliga erfarenheter. En del av dessa utgångspunkter kan delas med vissa personer men inte med andra. En gemensam plattform kan även skapas i situationen genom att varje deltagare i konversationen bidrar till den genom sina inlägg (så kallad "grounding") (Schober, 1999). På motsvarande sätt kan den skapas i

frågeformulär. Föregående frågor påverkar hur efterföljande frågor tolkas och frågorna tolkas på ett visst sätt beroende på hur undersökningen lanseras (läs mer i kapitel 13).

Om plattformen inte är gemensam

Delar inte frågeställare och respondent plattform kan det alltså uppstå problem. Det tydligaste problemet är när respondenterna själva upplever att de inte förstår frågan. Mer förrädiskt, och sannolikt mycket vanligare, är snarare mer implicita problem där respondenten själv tycker sig förstå frågan men förstår den på ett annat sätt och utifrån en annan plattform än frågeställaren. Frågeställare och respondent förstår alltså frågan på olika sätt men utan att det ger några felsignaler som stoppar kommunikationen. Respondenter kan, utan att reflektera över det eller ifrågasätta, ge egna betydelser till vaga ord och termer. De upptäcker inte nödvändigtvis om frågeställaren använder ord i en annan betydelse än vad de själva gör. Inte heller frågeställaren upptäcker detta om inte svaren avviker på något avslöjande sätt. Generellt sett är intervjuer bättre på att fånga upp och korrigera problem med att frågeställare och respondent inte delar plattform. Enkäter saknar mänsklig interaktion och där undviks kommunikationsproblem bäst genom förarbete, till exempel utredningar om vilken terminologi respondenterna själva använder och hur de tänker kring sakfrågan.

Sammanfattande punkter:

- Språket är ett socialt verktyg med målet att förstå andra eller göra oss själva förstådda.
- En nyckel till förståelse är att de inblandade i konversationen har en gemensam plattform att "stå på".
- Plattformen kan bestå av det aktuella sammanhanget ("Två, tack" syftar på olika saker i hissen och tunnelbanekuren).
- Plattformen kan också bestå av erfarenheter vi delar med andra och, allra bredast, den kultur och samtid vi lever i.
- Om vi inte delar plattform finns det större risk för missförstånd.
- Frågekonstruktör och respondent bör dela plattform, annars kan det bli missförstånd som inte nödvändigtvis upptäcks.

Konversationens logik och regler

Konversationens regler

Konversation syftar i det här sammanhanget på sättet som människor kommunicerar med varandra oavsett om det sker i tal eller i skrift. Det är alltså en lite bredare betydelse än den vanliga ("prata"). Det är kanske inget vi reflekterar över men det är ändå lätt att tänka sig att konversation måste styras av regler, både för att fungera praktiskt och för att effektivt överföra information mellan talare och mottagare. Enligt Grice (1975, 1978), och Levinson (1983) (Schwarz, 1996) har konversationer sin egen logik:

Övergripande är samarbetsprincipen – att alla inblandade i konversationen implicit antar att alla talare följer reglerna och att varje yttrande är ärligt och relevant. Grundsatserna är (Schwarz, 1996):

1. **Uppförande:** Var tydlig och undvik oklarheter. Var kortfattad, hövlig och avbryt inte.
2. **Relevans:** Yttrandet (det som sägs eller skrivs) ska vara relevant. Byt inte ämne eller yttra sådant som inte är intresserant i konversationen. Yttrandet ska vara förståeligt utifrån den gemensamma plattform som finns mellan talare och lyssnare.
3. **Kvalitet:** Yttra det som är sant. Var inte vilseledande eller hävda sådant som saknar belegg. Ljug och överdriv inte.
4. **Kvantitet:** Yttrandet ska vara så informativt som krävs, givet konversationens syfte och mål, men inte innehålla överflödigt information. Alltså: Varken för mycket eller för lite information.
5. **Överträdelser:** Överträdelser av reglerna ska signaleras tydligt till de andra deltagarna, till exempel vid byte av samtalsämne.

Dessa satser styr alltså konversationer som kännetecknas av samarbetsprincipen. Hit hör de flesta konversationer som människor deltar i.

Att tolka konversationens ämne

Grundsatserna kräver tolkningar och slutledningar av deltagarna för att det ska fungera, till exempel vad som utgör konversationens syfte eller ämne. I nedanstående exempel säger A följande i en konversation mellan A och B:

A: Jag skickade in slutrapporten för undersökningen om hälsa igår.

B måste nu tolka vad som är A:s huvudsakliga poäng och konversationens ämne, för att sedan formulera ett yttrande som för detta vidare. Två exempel på repliker från B är:

B1: Jaså? Kom det fram några intressanta resultat?

B2: Jag skickade in min deklarationsblankett igår.

Om vi börjar med B2 så kan det tyckas som ett märkligt inlägg i konversationen, men det har faktiskt flera beröringspunkter med A:s inlägg. Det uttrycker vad som skickades in igår, vilket A:s yttrande också gjorde. A:s huvudpoäng var dock sannolikt inte vad som i allmänhet skickades in just igår, utan ämnet var troligare slutrapporten för undersökningen om hälsa. B2 uttrycker därför att B missförstått konversationens ämne och A:s intention med yttrandet. B2 bryter därmed mot grundsatsen om relevans (se ovan) och om detta hade varit en autentisk konversation skulle nog de flesta i A:s position reagera med förundran på en sådan replik. B1 uttrycker vad som sannolikt är en korrekt tolkning av huvudpoängen i A:s yttrande och följer därför reglerna (Ashcraft, 1994).

Tolkningar bortom det bokstavliga

A: Hur bedömde ni Johns jobbansökan till enkätgruppen?

B: Han har snygg handstil.

Om B:s replik i exemplet ovan tolkas bokstavligt ter det sig obegripligt men våra tolkningar och slutsatser går ofta bortom ordens faktiska innebörd. Faktum är att vardagskonversation är full av bokstavligt obegripliga yttranden i samma stil som ovanstående. Konversationsreglerna, och antagandet att deltagarna följer dessa och är samarbetsvilliga, hjälper oss att tolka innebörden bortom det strikt bokstavliga. Snygg handstil kan knappast betraktas som någon av de viktigaste kvalifikationerna för att arbeta med enkäter. Om B ansåg att snygg handstil var den enda av Johns egenskaper som var värda att nämna (regler om kvantitet och relevans) måste det innebära att John inte ligger särskilt bra till för tjänsten. Utifrån förutsättningen att B följer konversationsreglerna och är samarbetsvillig, kan alltså A dra den slutsatsen trots att det inte alls uttrycks i orden i B:s yttrande (Schwarz, 1996).

Att inte vara missledande

Våra tolkningar av språket kan alltså gå bortanför det bokstavliga men också det saklogiska:

A: Undersökningen har 13 000 urvalspersoner.

Logiskt sett är A:s yttrande sant även om undersökningen skulle ha mer än 13 000 urvalspersoner, till exempel 34 000 urvalspersoner. Mottagarna tolkar det förmodligen som om undersökningen har ganska exakt 13 000 urvalspersoner eftersom om undersökningen har 34 000 urvalspersoner borde A ha sagt det istället för 13 000 (satsen om kvantitet och även satsen om kvalitet). Detta är ett ytterligare exempel på språkinferens – att mottagaren gör tolkningar bortom det strikt bokstavliga eller logiska (Schwarz, 1996).

Att inte ge överflödigt information

Ytterligare en aspekt av konversationsreglerna är att man inte ska ge information som redan är given. Informationen kan vara korrekt i sig men eftersom den är given bryter den mot regeln om kvantitet och måste därför förstås på något annat sätt (Schwarz, 1996).

A: Pelle är inte försenad till jobbet varje dag.

De allra flesta är inte försenade varje dag så varför ger då A denna information om just Pelle? Ska det tolkas som att Pelle trots allt har problem att passa tiden, även om han inte kommer försent varje dag? Kanske är han försenad *nästan* varje dag? Något syfte måste väl A ha med sitt yttrande? Alltså, trots att A:s yttrande bokstavligt är ett icke-uttalande kan mottagarna tolka det som om det finns ett annat syfte. Annars skulle nämligen A bryta mot regeln om kvantitet genom att ge överflödigt information som redan är känd (Schwarz, 1996).

Konversationsreglerna och undersökningar

Ovanstående regler gäller alltså för konversation, men det finns många tecken på att de även är relevanta för kommunikationen mellan frågeställare och respondent i ett frågeformulär (Schwarz, 1996). Till skillnad från vanlig konversation är konversationen i undersökningar väldigt begränsad och standardiserad. Det finns inte samma utrymme för förhandling eller förklaring som i muntliga dialoger med vänner. Detta skulle kunna betyda att respondenterna förlitar sig på reglerna i ännu högre grad. Exempelvis att varje del av de enskilda frågorna är relevanta

och har genomtänkta syften eller att frågeställaren är ärlig och ger tillräckligt med information.

Forskningen om frågekonstruktion innehåller en del resultat som kan få besvarandet av enkätfrågor att te sig slumpmässigt. Dessa resultat redovisas mer utförligt på andra platser i boken och beskrivs endast övergripande här. Några exempel (Schwarz, 1996):

- 30 procent av respondenterna anger åsikter om påhittade ämnen som borde vara omöjliga att ha en åsikt om.
- Respondenternas svar påverkas av om numreringen till en svarsskala går från -5 till +5 jämfört med från 0 till 10.
- Respondenternas svar på frågan hur ofta de tittar på tv påverkas av hur intervallen ligger i svarsskalan.
- Hur lyckliga personer uppger att de är i sitt äktenskap respektive i allmänhet påverkas av i vilken ordning frågorna ställs.

Ovanstående resultat kan få respondenterna att framstå som omotiverade och slarviga. Folk borde väl titta lika mycket på tv oavsett hur frågekonstruktören lägger intervallen i svarsalternativen? Schwarz (1996) menar tvärtom att resultaten ovan kan vara ett utslag av att respondenterna försöker förstå frågeställarens intention och göra sitt bästa utifrån antagandet att frågeställaren är samarbetsvillig och följer reglerna. Varför skulle frågeställaren ljuga och ställa frågor om ämnen som inte existerar? Varför skulle svarsskalan vara numrerad från -5 till +5 om numreringen i sig inte betyder något? Varför ska man ha med äktenskapet i svaret på en fråga om hur lycklig man är i allmänhet, när man redan svarat för äktenskapet i en separat föregående fråga? Schwarz och kollegor har gjort många experiment som bekräftar att så är fallet – att respondenter tenderar att tolka meningen i frågor som om konversationsreglerna gällde.

Dessa regler tycks alltså vara viktiga för hur respondenter förhåller sig till och förstår frågor och därför bör frågekonstruktören beakta dem. Respondenternas tolkning kan gå bortom den bokstavliga betydelsen och bero på sammanhanget (respondentens egen plattform eller den plattform som frågeställaren tidigare etablerat). Även aspekter som frågeställaren kan uppfatta som rent stilistiska kan uppfattas som meningsbärande av respondenterna. Alla dessa tolkningar är förstas svåra att förutsäga för frågekonstruktören men om man är medveten om de övergripande mekanismerna ökar möjligheten för att man ska upptäcka dem på förhand.

Sammanfattande punkter:

- Mänsklig kommunikation (så kallad konversation) styrs av regler vad gäller till exempel relevans, uppförande och kvantitet.
- En konversation styrs av samarbetsprincipen – att vi antar att andra följer reglerna.
- Reglerna, och antagandet att talaren följer dem, möjliggör tolkningar både bortom det bokstavliga och saklogiska.
- Mycket tyder på att dessa regler även är relevanta i kommunikationen med respondenter i en undersökning. Exempelvis: Varför ska svarsskalan vara numrerad från -5 till +5 (snarare än från 0 till 10) om inte numreringen i sig betyder något? Varför ska man ha med äktenskapet i svaret till en fråga om hur lycklig man är i allmänhet när man redan har svarat för äktenskapet separat i en föregående fråga?
- Frågekonstruktören bör vara medveten om dessa regler.

Några övergripande språkprinciper

Några vanliga problem för förståelsen av frågor är

- att centrala begrepp i frågan är ospecifika och tillåter många möjliga tolkningar som inte nödvändigtvis överensstämmer med vad frågekonstruktören tänkt sig
- att frågan innehåller exempel eller beskrivningar som kan leda till andra tolkningar av frågan än vad frågekonstruktören avsett
- att frågan innehåller för svåra ord och termer
- att frågan kan vara enkel språkligt sett men att syftet ändå är otydlig, vilket gör att frågan trots enkla ord är svår att förstå.

Dessa förekommer oftare än vad man tror eftersom de inte alltid upptäcks. Därför bör frågekonstruktören tänka på dem i förhand. Nedan beskrivs några övergripande språkprinciper som frågekonstruktören kan tänka på för att undvika ovanstående problem. Dessa språkprinciper är användbara i en mängd olika situationer. Principerna är bland annat baserade på kognitiv psykologi.

Ospecifika ord och begrepp

Alternativa tolkningar är ett vanligt problem med enkätfrågor. Om respondenterna tolkar frågan på olika sätt, och annorlunda än vad vi som

frågekonstruktörer tänkt oss, påverkar det förstås kvaliteten på svaren. Hur uppstår då problem med alternativa tolkningar och hur kan man försöka begränsa dem?

EXEMPEL:

Tror du att barn tar skada av att titta på program med våld i ...?

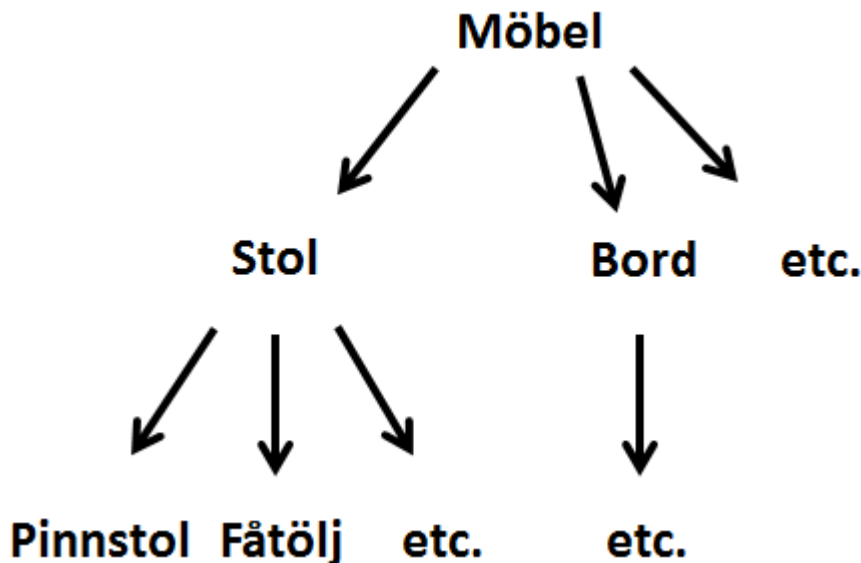
Utvärderingar av ovanstående fråga visade att begrepp såsom "barn" och "tar skada" uppfattades väldigt olika av olika respondenter (Belson, 1981). Barn är ett otydligt begrepp. Det syftar dels på ett släktled – vi är alla våra föräldrars barn. Det beskriver också ett utvecklingsstadium – barn är någonting man är när man är liten – och det finns ingen given gräns när barnstadiet övergår i nästa stadium. "Tar skada" är också otydligt. Tolkningarna kan variera från mildare oro till allvarigare psykologiska skador. Vad är det för tidsperspektiv som avses, tillfälliga skador eller långsiktiga konsekvenser? Frågan innehåller flertalet otydligheter. Det viktiga för frågekonstruktören är att förstå varför – vad är det som tillåter denna mängd av tolkningar? Vad är det för egenskaper i begreppen som ger en tolkningsrymd? Om man förstår detta kan man också åtgärda frågeformuleringen så att tolkningsrymden varken är för stor eller för liten utan väl anpassad till avsikten med frågan.

Begreppshierarkier – från generell till specifik

Converse och Presser (1986) skriver att ju mer generell en fråga är desto fler alternativa tolkningar möjliggörs. Vad menas då med "generell"? För att tränga lite djupare in i det tar vi hjälp av den kognitiva psykologin.

Kognition handlar om informationsprocesser, hur människan behandlar information. Vi bombarderas ständigt med information via våra sinnen. Hur strukturerar vi informationen så att omvärlden blir förståelig? Enligt ett kognitivt perspektiv sorterar människor informationen i hierarkiska kategorier som går från generell till specifik, så kallade semantiska nätverk (Lundh, Montgomery & Waern, 1992).

Figur 7.1



”Möbel” ligger på den mest generella nivån i nätverket ovan och har flera underkategorier, till exempel ”Stol” och ”Bord”. Dessa har i sin tur egna underkategorier. För varje nivå blir innehållet mer specifikt och avgränsat och alternativ utesluts. ”Byrå” är till exempel en möbel, men ryms inte inom underkategorin ”Stol”.

Ord/begrepp på en för generell nivå

Utifrån detta perspektiv – hierarkiska begreppsnivåer – är en källa till otydligheter att inkludera begrepp på fel nivå utifrån frågans syfte. Vanligast i praktiken är att begreppen, precis som Converse och Presser (1986) hävdar, är på en för generell nivå. Om vi syftar på ”hammare”, för att ta ett exempel, ska vi inte fråga efter ”verktyg” eftersom det ligger på en mer generell nivå i begreppshierarkin och därför inkluderar andra verktyg än just hammare, som en såg eller borr. Då finns det en risk att respondenten kommer att tänka på något av dem när de besvarar frågan, snarare än just det verktyg vi tänkte oss (Kerr, 2005). Om man överför detta perspektiv till frågan i Belsons studie ovan kan man konstatera att både ”barn” och ”tar skada” ligger på en för generell nivå för att minimera alternativa tolkningar. För att förtydliga den frågan bör man därför vandra neråt i begreppshierarkin tills det blir tillräckligt specifikt.

Det kan vara genom att till exempel ange vilken åldersgrupp det rör sig om eller vad för slags skadetyp som avses.

Ett annat exempel (från föregående kapitel):

EXEMPEL:

När flyttade du hemifrån?

När ovanstående fråga ställs är det oftast ett årtal som avses, men det finns inget i frågan som indikerar det. Frågeordet "när" tillåter tvärtom många olika tidsmässiga utgångspunkter. Denna fråga skulle kunna ge svar i stil med "när jag tog studenten" eller "det var när mina föräldrar skildes" (Fowler, 1995). Dessa svar kan mycket väl vara korrekta, det är frågan det är fel på. Tolkningsrymden är större än avsikten med frågan. Genom att vandra neråt i begreppshierarkin kan frågekonstruktören bli tillräckligt specifik.

Det är svårt att vara tillräckligt specifik

Ovanstående exempel kan kanske tyckas vara enkelt att undvika men det är det inte. Tvärtom är det ett av de vanligare problemen i frågekonstruktion. Dels kan det vara svårt att upptäcka. Som beskrevs i tidigare avsnitt är vi människor främst intresserade av mening och förståelse, inte av ord i sig. Detta kan göra att frågekonstruktören "läser" det som hammare trots att det faktiskt står verktyg, eftersom han eller hon vet att det är hammare som avses. Det är detta som brukar kallas att man är hemmablind för texter man själv skrivit. Det kan undvikas genom att låta någon annan granska frågeformuläret. Helst ska det vara någon som inte är insatt alls i undersökningen.

Ett annat skäl till att man ligger på en för generell nivå i sina frågor har med operationaliseringen att göra – den process där vi definierar vad vi vill mäta (läs mer i kapitel 8). Det vi gör när vi operationaliserar begrepp är på sätt och vis att vandra neråt i begreppshierarkin. Vi gör vägval, beslutar oss för vad vi menar och vad vi inte avser, tills vi har kommit ner till en tillräckligt specifik nivå. Att begrepp ligger på en generell nivå i frågeformuleringarna kan därför vara ett utslag av att man inte kommit tillräckligt långt i den processen – man har inte tillräckligt väl definierat exakt vad det är man vill mäta. Ibland kanske det därför kan ligga något i Esaias Tegnér's bevingade ord om att "det dunkelt sagda är det dunkelt tänkta".

Sammanfattande punkter:

- Ett vanligt problem i frågor är att viktiga ord och begrepp inte är tillräckligt specifika.
- Med ospecifikt menas att ett ord eller begrepp ligger på en generell nivå i begreppshierarkin.
- Att ligga på en för generell nivå med ord eller begrepp öppnar för andra tolkningar än det man egentligen avsett.
- Det är lätt hänt att vara ospecifik. Det kräver arbete och uppmärksamhet för att undvikas.

Missledande exempel

Tidigare beskrevs problemet med att ligga på en mer generell nivå i frågeformuleringen än vad man egentligen avsett. Men ibland kanske syftet med frågan är att ligga på en generell nivå i frågeformuleringen och tillåta alla tänkbara underkategorier. Som frågekonstruktör kan man vara osäker på om den generella nivån är tillräckligt tydlig och därför vilja tillföra något exempel för att hjälpa respondenten på traven.

EXEMPEL:

Hur många gånger i månaden motionerar du?

Ovanstående fråga är medvetet generell. Dess syfte är att inkludera alla tänkbara motionsformer snarare än någon specifik sort. Det kan dock behöva förtydligas vad som avses med "motion". Detta kan göras genom en definition eller genom att tillföra belysande exempel. Men det är viktigt att exemplen bidrar till *rätt* förståelse. De ska belysa den yta som frågan ska täcka, inte bidra till en bredare eller smalare tolkning. Detta är svårt i praktiken. Exempelen i frågor kan till exempel styra förståelsen i en viss snarare än i en annan riktning eller utesluta sådant som borde ingå. Att tillföra exempel i en fråga bör därför vara något som sker med eftertanke. Hur ska man då tänka vad gäller exempel? Vad är ett bra respektive dåligt exempel?

Bra och dåliga exempel ur ett kognitivt perspektiv

Återigen kan vi anknyta till den kognitiva psykologin som inspiration (Lundh, Montgomery & Waern 1992). Varje kategori borde definieras av en sammansättning av egenskaper. Det är sammansättningen som borde

avgränsa och särskilja kategorierna från varandra. Kategorin "fågel" kan till exempel definieras bland annat av nedanstående egenskaper:

- näbb
- fjädrar
- kan flyga.

Talgoxe är ett tydligt exempel på kategorin fågel. Den uppfyller många, kanske alla, av de egenskaper som kännetecknar kategorin. Talgoxe är därför en mycket god representant för kategorin. Pingvin är däremot inte en lika god representant för kategorin "fågel". Pingvinen uppfyller inte vissa av de mest centrala egenskaperna för kategorin. Pingvinen har till exempel inga fjädrar och kan inte flyga. Vad betyder då detta för konstruktionen av frågor? Varför kan dåliga exempel vara missledande?

Uppenbarligen kan saker och ting variera i hur pass goda representanter de är för en viss kategori. Eftersom pingvinens egenskaper bara delvis överlappar med kategorin "fågel" så innebär det samtidigt att pingvinen har en massa egenskaper som *inte* överlappar, och detta skulle kunna leda till andra problem. Vid frågekonstruktion kan det innebära att respondenten fokuserar sin förståelse på egenskaper som inte överlappar med kategorin vi vill exemplifiera. Detta borde kunna försvåra förståelsen eller till och med leda till att frågan helt eller delvis missförstås.

EXEMPEL:

Brukar du ha direktkontakt med andra på internet genom att till exempel chatta eller spela dataspel?

Ovanstående fråga syftar till att mäta direkt, social interaktion via internet. "Chatta" framstår som en mycket god representant på direkt, social interaktion eftersom det innebär att man pratar med en annan människa i realtid via internet. "Spela dataspel" är däremot mer tveksamt. Det är inget felaktigt exempel i sig. Det finns dataspel där spelarna har den interaktion som frågan syftar till att mäta, så kallade online-spel. Men det finns även en mängd dataspel som inte har det. "Spela dataspel" har därför en massa andra egenskaper som inte alls sammanfaller med frågans syfte. Att exemplifiera med "spela dataspel" kan därför försvåra förståelsen av frågans syfte och kan ha en missledande effekt.

Man bör sålunda undvika exempel som är otydliga representanter på den kategori de ska företräda i frågan. En tv är till exempel ett otydligt exempel på en möbel (Lundh, Montgomery & Waern, 1992). När man skriver frågor är det därför bra att överväga vilka egenskaper ord och begrepp har och i vilken mån de sammanfaller med vad det är man vill mäta eller exemplifiera.

Exempel kan dock fylla andra funktioner i en fråga än att illustrera en "högre kategori", till exempel påminna om saker som respondenten annars nog inte skulle ha tänkt på (läs mer i kapitel 2).

Sammanfattande punkter:

- Exempel kan fylla en viktig funktion i en fråga för att förtydliga vad som avses.
- Det är dock viktigt att exemplen främjar *rätt* förståelse. Att de har rätt "täckningsgrad" – det vill säga varken leder till för smal eller för bred förståelse.
- Ur ett kognitivt perspektiv ska ett exempel som ska illustrera en högre kategori dela viktiga egenskaper med den kategorin. Exempelvis är talgoxe ett bra exempel på fågel och pingvin ett dåligt exempel.
- Det är därför viktigt att överväga vilka egenskaper ord och begrepp har och i vilken mån de sammanfaller med vad det är man vill mäta.

Svårförståeligt innehåll

Svåra ord

Som beskrivs i ett tidigare avsnitt i det här kapitlet är språket ett gemensamt symboliskt system för kommunikation. Språket är dessutom inte gemensamt fullt ut och olika personer kan ge olika betydelser till olika ord. Det förekommer förstås också att personer inte alls känner till vissa ords innebörd, kanske för att de inte uppmärksammat eller stött på ordet tidigare. Ordets uppbyggnad kan också påverka om det är lätt eller svårt att läsa och förstå. Ord som har många betydelser, är okända för många eller som är svåra att läsa och förstå kan benämnas *svåra ord*.

Vad blir då konsekvensen av svåra ord? Det beror på av vilken anledning ordet är svårt. Om läsaren inte känner till ordet är förstås risken för

missförstånd hög. Om det är ett ord som är svårt för att det är vagt är risken kanske snarare att olika respondenter kommer att tolka det på olika sätt (se exemplen med "far illa" ovan). Det är också lätt att föreställa sig att respondenten kan bli fundersam över vad som avses med det mångtydiga ordet i just den aktuella frågan. Ord som är svåra för att de är krävande att förstå kan öka respondentbördan. De kan även öka andelen som svarar "Vet inte" (Sudman & Bradburn, 1974).

Vilka ord är det då som är svåra? Det beror bland annat på undersökningsgruppen. Om frågeformuläret exempelvis riktar sig till en viss yrkesgrupp kan facktermer som är kända inom den gruppen vara helt godtagbara att använda. Om man riktar sig till allmänheten bör man däremot inte använda facktermer utan att förklara dem. Man ska inte överskatta till vilken grad folk känner till ord eller att de ger samma betydelse till dem som frågekonstruktören (läs mer i avsnittet Praktiska språktips i kapitel 1). En bra grundregel är att vara medveten om respondenternas språkbruk och rätta sig efter det. Om undersökningen har ett specifikt undersökningsämne kan det vara värt att göra efterforskningar om vilken terminologi respondenterna använder för just det området.

Långa frågeformuleringar

Människor är begränsade i hur mycket information de kan hålla i huvudet samtidigt. En fråga som har mycket beskrivningar, bisatser och där det frågebärande innehållet kommer sent, kan vara svårare att förstå. Detta gäller speciellt om respondenten är okoncentrerad eller omotiverad samt vid telefonintervjuer där respondenten inte kan läsa om frågan vid behov. Därför ska frågor inte vara längre och mer komplicerade än nödvändigt. Om frågan har för mycket innehåll kan det vara värt att överväga att dela upp den i flera frågor. Den sammanlagda bördan blir förmodligen mindre med två lätta frågor än med en svår (Foddy, 1994). Frågor ska dock inte förkortas på bekostnad av förståelsen.

En annan risk med långa frågor är att ju längre och mer avancerad en fråga är, desto större är risken att frågekonstruktören begår språkliga misstag. Det kan till exempel handla om syftningsfel, att frågan innehåller mer än en fråga eller att inledningen och slutet på frågan inte hänger ihop. Att arbeta med att förkorta frågor kan därför vara ett sätt att renodla och förtydliga dem.

Sammanfattande punkter:

- Svåra ord försvårar förståelsen.
- Ord som har många betydelser, är ovanliga, är facktermer eller krångliga att läsa är *svåra ord*.
- Långa frågeformuleringar är både svårare att skriva och att läsa. De kan därför vara svårare att förstå.

Frågans intention

Hittills har avsnittet mest handlat om enskilda ord och begrepp och det är naturligtvis viktigt att de är tydliga. Men frågeformuleringar är sammansatta av flera ord som tillsammans utgör en helhet. Det är givetvis viktigt att helheten, inte bara de enskilda orden, förstås enligt den intention som finns. Frågans *mening*, dess innebörd och betydelse, måste framgå och respondenten måste förstå frågekonstruktörens intention. Med det menas inte att respondenten måste vara på det klara med frågans användningsområde – hur frågan ska analyseras eller vad den syftar till att mäta. Med "intention" menas snarare att respondenten måste förstå vad det är frågekonstruktören vill veta och vad respondenten ska ge för information.

Intentionen finns ibland bortom det bokstavliga

När man diskuterar intention är det naturligt att koppla till pragmatism som har just med sändarens intentioner att göra (läs mer i avsnittet Vad är språket?). Ibland ska språket inte tolkas bokstavligt utan meningen och sändarens intention är något annat bortom orden, eller kanske snarare mellan raderna.

EXEMPEL:

Bra jobbat!

Det drar från fönstret.

Det första exemplet ovan betyder ordagrant att någon har gjort något bra men sändarens intention kan vara ironisk och mena det motsatta – att någon har gjort dåligt ifrån sig. I det andra exemplet kan sändarens intention vara en uppmaning att stänga fönstret. Ingen av dessa intentioner uttrycks i satserna. De finns snarare i situationen, i vår samtalskultur och kan betonas med kroppsspråk och tonfall. Den här

typen av extremt sammanhangsberoende satser bör naturligtvis undvikas vid så pass begränsad kommunikation som i frågeformulär.

Otydlig intention = otydlig fråga

Men även i raka, enkla frågor kan frågekonstruktörens intention vara otydlig. Det försvårar förståelsen av frågan och hur den ska besvaras.

Schwarz, Knäuper, Oyserman och Stich (2008) ger frågan "Vad gjorde du igår?" som exempel. Den frågan är inte svår att förstå språkligt sett, men vad är syftet och meningen med frågan? Ska man räkna upp allting som hände under gårdagen? Det framstår förmodligen som orimligt. Ska man ta vissa, mer rutinmässiga aktiviteter som givna och därför utelämna dem? Ska man fokusera på det som gjorde gårdagen unik? Det är inte lätt att avgöra eftersom vi inte känner till frågekonstruktörens intention med frågan. Frågan är alltså svår trots att den är språkligt enkel.

EXEMPEL:

Arbetade du med utredning under 2009?

Ja

Nej → Tack för din medverkan. Var vänlig och skicka in frågeformuläret

Exemplet ovan är inte svårt att förstå språkligt men man kan tänka sig att svaren på frågan varierar mellan olika personer. Vissa kanske jobbade med utredningar till viss del eller under en särskild tidsperiod.

Svarsalternativen tillåter dock inga sådana nyanser utan endast "Ja" eller "Nej". Konsekvenserna av valet av svarsalternativ blir dessutom stora eftersom de som svarar "Nej" inte ska besvara fler frågor. Så ska de som endast till viss del arbetade med utredning svara "Ja" eller "Nej"? Det är svårt att avgöra eftersom frågan inte kommunicerar frågekonstruktörens intention, syfte med frågan och vad det är som är viktigt att tänka på.

När frågor skrivs eller granskas är det viktigt att reflektera över enskilda ord och begrepp men även frågan som helhet måste beaktas – framgår intentionen och är det tydligt vilken information respondenten ska lämna?

Sammanfattande punkter:

- Frågans mening (intentionen med frågan) måste framgå.
- Om frågans mening inte framgår är den otydlig trots att den kan vara språkligt enkel.

Sammanfattning

Språket och kommunikation – sammanfattning**Språket som kommunikationsmedel:**

- De allra flesta kan läsa och skriva men det behövs mer kunskap och förmåga än så för att skriva bra frågor.
- Språket har många brister som kanske inte märks så mycket i vardagen men som kan få stort utslag i ett frågeformulär.
- Frågekonstruktören bör ha hög språklig medvetandegrad och fundera över ordval, satsbyggnad, språkförståelse och alternativa tolkningar.

Vad är språket?

- Språket är ett gemensamt, symboliskt system för kommunikation.
- Ord har ingen given betydelse utan det är något vi måste lära oss och det kan variera något från person till person.
- Grammatik är språkets regelverk – om det inte fanns regler skulle vi inte kunna förstå varandra.
- Syntax är regler för hur enskilda ord sätts samman i en meningssats, till exempel ordningsföljden av ord.
- Semantik har med ords och språkets betydelse att göra.
- Pragmatik har med intentionen bakom ett yttrande att göra. I det praktiska bruket av språket ligger ofta meningen i ett yttrande bortom det bokstavliga.

Konversationens logik och praktik:

- Språket är ett socialt verktyg med målet att förstå andra eller göra oss själva förstådda.
- En nyckel till förståelse är att de inblandade i konversationen har en gemensam plattform att "stå på".

- Plattformen kan bestå av det aktuella sammanhanget ("Två, tack" syftar på olika saker i hissen och tunnelbanekuren).
- Plattformen kan också bestå av erfarenheter vi delar med andra och, allra bredast, den kultur och samtid vi lever i.
- Om vi inte delar plattform finns det större risk för missförstånd.
- Frågekonstruktör och respondent bör dela plattform, annars kan det bli missförstånd som inte nödvändigtvis upptäcks.

Konversationens logik och regler:

- Mänsklig kommunikation (så kallad konversation) styrs av regler vad gäller till exempel relevans, uppförande och kvantitet.
- En konversation styrs av samarbetsprincipen – att vi antar att andra följer reglerna.
- Reglerna, och antagandet att talaren följer dem, möjliggör tolkningar både bortom det bokstavliga och saklogiska.
- Mycket tyder på att dessa regler även är relevanta i kommunikationen med respondenter i en undersökning. Exempelvis: Varför ska svarsskalan vara numrerad från -5 till +5 (snarare än från 0 till 10) om inte numreringen i sig betyder något? Varför ska man ha med äktenskapet i svaret till en fråga om hur lycklig man är i allmänhet när man redan har svarat för äktenskapet separat i en föregående fråga?
- Frågekonstruktören bör vara medveten om dessa regler.

Några övergripande språkprinciper:

- Ett vanligt problem i frågor är att viktiga ord och begrepp inte är tillräckligt specifika.
- Med ospecifikt menas att ett ord eller begrepp ligger på en generell nivå i begreppshierarkin.
- Att ligga på en för generell nivå med ord eller begrepp öppnar för andra tolkningar än det man egentligen avsett.
- Det är lätt hänt att vara ospecifik. Det kräver arbete och uppmärksamhet för att undvikas.
- Exempel kan fylla en viktig funktion i en fråga för att förtydliga vad som avses.
- Det är dock viktigt att exemplen främjar *rätt* förståelse. Att de har rätt "täckningsgrad" – det vill säga varken leder till för smal eller för bred förståelse.

- Ur ett kognitivt perspektiv ska ett exempel som ska illustrera en högre kategori dela viktiga egenskaper med den kategorin. Exempelvis är talgoxe ett bra exempel på fågel och pingvin ett dåligt exempel.
- Det är därför viktigt att överväga vilka egenskaper ord och begrepp har och i vilken mån de sammanfaller med vad det är man vill mäta.
- Svåra ord försvårar förståelsen.
- Ord som har många betydelser, är ovanliga, är facktermer eller krångliga att läsa är *svåra ord*.
- Långa frågeformuleringar är både svårare att skriva och att läsa. De kan därför vara svårare att förstå.
- Frågans mening (intentionen med frågan) måste framgå.
- Om frågans mening inte framgår är den otydlig trots att den kan vara språkligt enkel.

8 Att mäta med frågor

Författare: Lina Fjelkegård (SCB)

Att göra en undersökning blir aktuellt när det finns ett område man vill studera. Det finns ett behov av att ta reda på någonting för att kunna förstå eller beskriva ett fenomen. Det här kapitlet handlar om denna startpunkt – vad man ska studera – och hur man sedan går vidare utifrån det.

Varför undersöka med hjälp av frågor?

När lämpar det sig att göra en frågeundersökning?

Varför undersöka med hjälp av frågor?

De typer av frågeundersökningar som hanteras i den här boken handlar i första hand om kvantitativa studier. Det är studier som ska ge numeriska data och där det undersökta fenomenet ska kunna beskrivas statistiskt. Kvantitativa frågeundersökningar ger oss en möjlighet att till exempel ta reda på hur stor andel av en undersökt grupp som äter godis varje dag, hur ofta man i genomsnitt träffar någon släkting eller hur många timmar i veckan man i genomsnitt lägger på transport till och från arbetet. Om vi använder statistiskt korrekta urvalsmetoder kan vi dessutom generalisera och dra slutsatser om den population vi dragit urvalet ifrån. Det är en lockande och i många fall utmärkt undersökningsmetod, men innan man bestämmer sig bör man fråga sig om det är den bäst lämpade strategin. Om det fenomen man ska undersöka är komplext och svåråtkomligt kan kvantitativa undersökningar bli begränsande. De fångar de aspekter av ett fenomen som går att beskriva med statistiska mått, men missar kanske annan viktig kunskap. De är också begränsade till att mäta det som vi på förhand identifierat som relevant och tagit med i vårt frågeformulär. Vi missar således sådant som vi inte kunnat förutse. Med exempelvis kvalitativa intervjuer och fokusgrupper eller textanalyser och deltagande observation kan vi istället få en djupare förståelse och även fånga upp processer eller det avvikande. Istället för att beskriva fenomenets generella drag kan vi lyfta fram det specifika. De kvalitativa metoderna kräver inte heller på samma sätt att vi på förhand formulerar våra frågor och svarsalternativ. Vi behöver inte heller ha lika utförliga kunskaper om vilka de relevanta aspekterna är (läs mer i till exempel Sjöberg, 1999).

Skillnaden mellan kvantitativa och kvalitativa metoder ska inte överdrivas, men innan man går vidare med att genomföra sin frågeundersökning bör man noggrant överväga vilken av de två metoderna som lämpar sig bäst. Kanske kan båda användas och integreras med varandra.

Exempel på ett område att studera:

EXEMPEL:

Hur ser livssituationen ut för barn till skilda föräldrar?

De olika strategierna lämpar sig olika bra utifrån hur vi ser på frågeställningen, vilken typ av kunskap det är vi är ute efter och hur mycket förkunskap som finns. Om vi vill kunna beskriva barnens situation i kvantitativa termer måste vi samla in kvantitativa data. Om vi med "livssituation" menar mätbara konsekvenser, som till exempel boendesituation, skolprestation och tid med föräldrarna, skulle kanske en kvantitativ frågeundersökning vara en fungerande metod. Men det är inte säkert att vi på förhand vet vilka omständigheter som är relevanta att fråga om, vilket krävs i ett strukturerat frågeformulär där samtliga deltagare ska få samma frågor. De mest betydande faktorerna är kanske inte heller sådant som går att få fram genom strukturerade frågor. Vi kan dessutom anta att föräldrarnas skilsmässa är ett känsligt ämne för många barn och att situationen kan ha upplevts på varierande sätt. Det kan därför vara både mer etiskt och givande att träffa barnen, för att då ställa anpassade och flexibla frågor för att sedan analysera deras berättelser kvalitativt.

Om det finns tillräckligt med resurser är ofta det allra bästa att kombinera kvalitativa och kvantitativa metoder. På det sättet kan vi både få en generell beskrivning och en djupare förståelse. I en kvantitativ frågeundersökning kan vi ställa enkla frågor till barnen om hur de har det. I kvalitativa intervjuer kan vi sedan samtala om skilsmässan, fråga hur de mår, vad som är bra och vad som är dåligt. Därefter analyseras hur barnens livssituation ser ut och hur de upplever att föräldrarna inte lever tillsammans.

I tabell 8.1 nedan framgår de viktigaste skillnaderna mellan kvalitativa och kvantitativa metoder utifrån fokus, mål, urval och metod.

Tabell 8.1 Översikt kvalitativa och kvantitativa metoder

	Kvalitativa undersökningar	Kvantitativa undersökningar
Fokus	kvaliteter djup intresserad av det specifika	kvantiteter hur många, hur mycket? bredd intresserad av det allmänna
Mål	förståelse, upptäckt, hypotesgenerering	förklaring, prediktion, hypotestestning, kvantitativ beskrivning
Urval	litet, icke slumpmässigt	stort, slumpmässigt
Metod	närhet (forskaren som instrument: djupintervjuer, ostrukturerade observationer)	distans (icke levande instrument – strukturerat, standardiserat: enkäter, strukturerade intervjuer, strukturerade observationer)

Sammanfattande punkter:

- Kvalitativa och kvantitativa undersökningar ger delvis olika typer av kunskap. Hur man väljer att inrikta sin undersökning metodologiskt måste styras av problem- och frågeställning.

Frågeundersökning eller annan kvantitativ metod?

Om vi nu efter noggrant övervägande kommer fram till att det lämpar sig bäst med en statistisk undersökning – vi vill kunna uttala oss om generella drag hos en större grupp individer (eller andra enheter) – är då en statistisk frågeundersökning den bästa metoden eller finns det andra sätt som bättre samlar in den kvantitativa informationen?

På ett vis är förutsättningarna för frågeundersökningar ganska självklara – vill man veta någonting så frågar man. Men det är inte så enkelt i praktiken. Det har redan framkommit i bokens tidigare kapitel att det finns mycket att tänka på och förhålla sig till när man ställer frågor för att

ta reda på någonting. Respondenter kan ha svårt att minnas det som efterfrågas. De kanske inte heller uppfattar frågorna på det sätt som avses eller de kanske inte alls har lust att delta i undersökningen. Kanske finns motsvarande data i register eller annan dokumentation som kan analyseras. Sveriges officiella statistik (SFS 2001:100) innehåller till exempel uppgifter om befolkningens utbildning, hushållens ekonomi, anmälda brott och företagets ekonomi. Annan dokumentation kan till exempel handla om media, offentliga dokument, domar och riksdagsprotokoll.

Ibland blir data mer tillförlitliga om den samlas in genom strukturerad observation. Det kan innebära att istället för att fråga personer om de brukar äta popcorn när de går på bio kan vi göra ett urval av filmvisningar och gå dit för att observera vad besökarna köper för snacks. Om man vill ha detaljerad information om någonting som man tror att respondenter lätt glömmer bort kan man be dem fylla i en loggbok eller kalender under en begränsad period. Ett exempel på en sådan datainsamling är SCB:s undersökning om hushållens utgifter. Om vi i en frågeundersökning skulle be respondenter ange hur mycket och vad de köper av olika varor skulle minnesproblematiken sannolikt göra att felet i mätningen blev förhållandevis stort. Istället låter man ett urval av hushåll löpande föra kassabok över samtliga utgifter under en begränsad period (Dahmström, 2005).

I väldigt många situationer är den kvantitativa frågeundersökningen den bästa metoden. Den kan ge oss information om väldigt mycket som vi annars inte hade kunnat få kunskap om. Det kan handla om attityder, erfarenheter och beteenden. Ibland kan man också se frågeundersökning som en mer etisk metod än att använda registerdata och vissa observationsstudier. Det beror på att respondenterna själva väljer om de ska delta och vilken information de ska dela med sig av.

Sammanfattande punkter:

- En kvantitativ frågeundersökning är ofta en bra metod, men inte alltid.

Från problem till fråga

Direkta och indirekta mätningar

I litteraturen om mätteori skiljer man mellan direkta och indirekta mätningar (läs mer i Carmines & Zeller, 1979; Magnusson, 2003). Med direkta mätningar avses när det som ska mätas är direkt observerbart. Det betyder att det går att se, höra eller känna på objektet och utifrån det dra slutsatser om dess egenskaper. Det kan handla om att mäta längd med ett måttband, att räkna antal personer i ett rum genom att gå in och titta eller att bedöma vilken dialekt en person har genom att lyssna när personen pratar.

När vi istället vill ta reda på någonting som vi inte kan observera genomför vi indirekta mätningar. Fenomen som exempelvis arbetslöshet, socioekonomisk status och attityder till privat sjukvård är relativt abstrakta, teoretiska begrepp. De går inte att se, höra och känna utan måste mätas på något annat sätt. Om fenomenen ska mätas genom en frågeundersökning krävs därför att de abstrakta teoretiska begreppen, här kallade för latenta begrepp eller koncept, konkretiseras (Björkman, 2000). Vi kan inte förvänta oss några tillförlitliga resultat om vi skickar ut en enkät med frågan: "Vad är din socioekonomiska status?" eller "Vad är din inställning till privat sjukvård?". Begreppen är för vaga och innehåller för många olika aspekter. Vi måste fråga på ett annat sätt. Vi behöver konkretisera begreppet och sedan formulera våra frågor utifrån denna konkretisering. Mätningen blir då indirekt eftersom det inte är det abstrakta begreppet, vilket egentligen är vårt intresse, som blir föremål för den faktiska mätningen. Konkretiseringen och frågorna fungerar då som mätbara representanter för det abstrakta begreppet. Man kan också säga att de är mätbara indikatorer för det abstrakta begreppet. Skillnaden mellan abstrakt begrepp, mätbar indikator och enkätfråga beskrivs mer utförligt i nästa avsnitt Begreppsanalys – det första steget.

För att det ska fungera med en indirekt mätning måste vi först definiera vad vi menar med de abstrakta latenta begreppen som ska mätas. Med dessa definitioner kan vi sedan få fram mätbara begrepp att formulera våra enkätfrågor utifrån. Att definiera vad begreppen innebär och vad vi menar med dem kallas vanligtvis för konceptualisering. Att sedan konkretisera det definierade konceptet till mätbara indikatorer och enkätfrågor är att göra en operationalisering. Båda dessa processer är avgörande för en mättnings kvalitet då de utgör grunden för att vi ska kunna mäta det vi avser. Frågorna ska representera vårt koncept. Det

kräver att vi har lyckats fastställa dess innebörd och utifrån det lyckats skapa mätbara indikatorer, för att utifrån det formulera frågor.

Direkta och indirekta mätningar bör dock inte ses som varandra uteslutande utan snarare som ändpunkterna på en dimension där olika mätningar innebär olika grader av direkt- respektive indirekthet. Att i en enkät ställa frågor om konkreta förhållanden, till exempel "Hur många barn har du?" är inte en direkt observation. Samtidigt är det betydligt mer direkt än att med ett stort antal indikatorer försöka fånga ett abstrakt latent begrepp som trivsel.

Sammanfattande punkter:

- När det som ska mätas är direkt observerbart – det går till exempel att se, höra eller känna på objektet – säger man att mätningen är direkt.
- När objektet är abstrakt och inte går att observera utarbetas indikatorer som är mätbara. Mätningen är då indirekt.

Begreppsanalys – det första steget

Ett vanligt problem vid frågeundersökningar är att man börjar formulera frågorna alldeles för tidigt, långt innan man vet vad det är man exempelvis menar med socioekonomisk status eller vad som ska ingå i sjukvård (Wärneryd m.fl., 1993). När data sedan är insamlad och analysen ska ta vid upptäcker man att det saknas centrala aspekter av begreppet. Det kan också vara att man inte vet hur man ska använda resultaten, eftersom man inte riktigt vet vad det var som avsågs att mätas. För att skapa bra förutsättningar för operationaliseringen behövs det därför först en begreppsanalys eller en konceptualisering.

Koncept och definitioner

Många begrepp uppfattas vid en första anblick som konkreta och okomplicerade. Man bryr sig då inte om att definiera deras betydelse.

Hus

Dator

Skada

Lärare

Använda

Begreppen ovan kan uppfattas som vardagliga och förhållandevis okomplicerade. Men även begrepp som vid en första anblick känns uppenbara kan leda till problem om de inte definieras. De kan ha olika betydelser i olika situationer eller så skiljer sig användningen åt mellan individer (Björkman, 2000). Vad är egentligen definitionen av en dator? En maskin som går att programmera? Är min kaffebryggare i sådana fall en dator? Någon kanske svarar ja på den frågan, någon annan svarar nej. Även om begreppen inte är så svåra att definiera måste en definition alltså göras. Om jag ska undersöka rektorers upplevelse av samarbete med lärare så måste jag ju veta vad jag menar med lärare och samarbete. Innan dess kan jag inte konstruera ett fungerande frågeformulär.

Rädsla

Lycka

Välbehållande

Levnadsförhållanden

Intolerans

Etnicitet

När det gäller ovanstående begrepp blir det lite mer uppenbart att det krävs en begreppsanalys för att kunna gå vidare med en undersökning. Begreppens betydelser är oklara, abstrakta och komplexa. De måste beskrivas och det ska framgå hur det relaterar till andra närliggande koncept (Hox, 1997). Är lycka en känsla, ett tillstånd eller en process? Hur förhåller det sig till välbefinnande och till levnadsförhållande – är något begrepp överordnat?

Tillvägagångssättet för en sådan begreppsanalys kan se ut på olika sätt (se Hox, 1997). En bra början är att se till konceptets teoretiska ramverk. Kanske finns det en fungerande och tydlig definition att arbeta vidare med eller så kan det teoretiska ramverket vara en grund för att utveckla definitionen. Om konceptet ingår i ett teoretiskt ramverk kan också övriga relaterade teoretiska begrepp bidra till en förståelse för konceptet. Genom att det relateras till andra begrepp framgår det unika och specifika med det aktuella konceptet och kärnan kan identifieras. Om den etablerade definitionen inte överensstämmer med det man själv avsett med konceptet bör man överväga att använda ett annat begrepp. Kanske finns det en väl utvecklad teoretisk diskussion kring konceptet lycka där delvis andra aspekter ingår än vad som är fokus i den egna studien (för olika definitioner av lycka se Brülde, 2007). Kanske ingår självkänsla och långsiktig tillförsikt i den etablerade definitionen, medan

vårt intresse snarare ligger vid en tillfällig känsla. I sådana situationer bör man se om det finns andra och kanske helt nya begrepp, som är bättre att använda (Björkman, 2000). Det är varken bra för oss eller för dem som tar del av våra resultat om de tror att vi har undersökt någonting annat än vi faktiskt har. Det är ju innehållet och betydelsen som är det centrala, inte vilken term vi använder för det latenta konceptet för att referera till dess innehåll. Eventuellt skulle *glädje* vara en bättre term i det här fallet. Om vi ser till hur begreppet används i litteratur och tidigare forskning kanske det visar sig att definitionerna av glädje och lycka skiljer sig åt och att den av glädje stämmer bättre överens med det vi är intresserade av. Om inte så blir det istället aktuellt att utveckla ett helt nytt begrepp. I vissa fall är det en betydligt bättre strategi än att använda ett begrepp som redan cirkulerar i det vardagliga eller ämnesspecifika språket. Ett etablerat begrepp utan gemensam definition blir ofta både svårt att definiera och operationalisera.

Att förhålla sig till olika definitioner av ett koncept är alltså en del av konceptualiseringen. Den egna definitionen kan antingen motsvara den minsta gemensamma nämnaren hos de tidigare definitionerna eller till en redan existerande definition. Om den egna definitionen blir en ny definition, trots omfattande tidigare forskning, måste detta kunna motiveras och argumenteras för i den dokumentation som blir aktuell för studien.

Ofta är det huvudsakliga och övergripande konceptet väldigt stort och brett. Den definition som erbjuds genom en analys av det teoretiska ramverket blir inte alltid tillräckligt tydlig för att arbeta vidare med. En möjlig lösning är då att specificera konceptets betydelse genom att identifiera och definiera centrala underkategorier eller komponenter (Frankfort-Nachmias & Nachmias, 1996). Vad ingår exempelvis i konceptet levnadsförhållanden – bostadssituation, inkomst och hälsa? (Se Wärneryd, 1993 och SCB Rapport 91, 1997, för en diskussion om konceptualisering av levnadsförhållanden). Genom att definiera dessa delar av begreppet klargör vi också vad vi menar med levnadsförhållanden. Dessa komponenter är också abstrakta, men inte i samma utsträckning som det övergripande konceptet. För att kunna definieras på ett tydligt sätt och fungera i den fortsatta processen kräver de därför sina egna begreppsanalyser (Björkman, 2000). Om vi fortsätter på det sättet finns det dock en risk för att begreppsanalysen övergår till operationalisering. I ett flertal steg sänks abstraktionsnivån och istället

för att definiera begreppet teoretiskt har vi då konkretiserat det och skapat mätbara indikatorer. Risken med ett sådant tillvägagångssätt är att konceptet aldrig definieras teoretiskt. Det blir istället en produkt av operationaliseringen – till exempel att levnadsförhållanden blir summan av våra indikatorer och inte ett definierat koncept. Det medför i sin tur att vi bland annat får problem att bedöma och förhålla oss till mätningens giltighet. Vi vet ju inte vad det är meningen att indikatorerna ska mäta (se Blumer, 1969 för en diskussion om riskerna med operationalism). Operationaliseringen är istället nästa steg. Efter att ha definierat begreppet ska det operationaliseras och det abstrakta konceptet ska bli konkreta mätbara indikatorer och frågor. Läs mer om det senare i kapitlet.

Konceptualisering av begrepp som inte är teoretiska

Ibland är det latenta begreppet inte särskilt teoretiskt utvecklat. Det kan ha en karaktär som i första hand varken uppfattas som teoretisk eller abstrakt, som i exemplen *hus*, *lärare* och *skada*. Då kan konceptets innehåll istället utredas utifrån den vardagliga användningen och förståelsen. Det bästa i sådana situationer är att lämna skrivbordet. Med hjälp av kvalitativa metoder går det att empiriskt utreda konceptets innebörd utifrån hur det används och definieras i samhället. Ostrukturerade eller halvstrukturerade intervjuer kan ge fördjupad kunskap om olika aspekter och delar av ett begrepp. Det kan sedan ligga till grund både för definitionen av huvudkonceptet (konceptualiseringen), men också för hur man utifrån definitionen ska kunna formulera indikatorer och konkreta frågor (operationaliseringen). På samma sätt kan fokusgrupper vara en bra metod för att empiriskt definiera ett koncept. Där får en grupp personer diskutera relevanta aspekter i förhållande till konceptet. Det ger dels en bred och nyanserad bild, dels en möjlighet att se huruvida det skapas en gemensam förståelse kring konceptet i gruppen. Ett sådant angreppssätt bygger på perspektivet att det genom interaktion och förhandlingar skapas en gemensam förståelse för olika aspekter och förhållanden i samhället. Människors definitioner är inte exakt överensstämmande med varandras och mellan olika kontexter, men det finns en övergripande gemensam tolkning och upplevelse av fenomen, språk och förhållanden i samhället (Mead, 1934; Parsons, 1951; Trost & Levin, 2004).

Skillnaden mellan den teoretiska konceptualiseringen och den mer empiriska kan delvis liknas vid skillnaden mellan ett deduktivt och ett induktivt förhållningssätt eller deduktiva och induktiva metoder.

Deduktion innebär att man utifrån teoretiska antaganden formulerar en tes som är logiskt nödvändig utifrån dessa antaganden. Tesen kan sedan testas empiriskt, men det centrala är att den har utvecklats ur ett teoretiskt resonemang (Rosing, 2002). Ett induktivt förhållningssätt innebär istället att man först gör sina empiriska observationer, utan att först formulera en testbar hypotes. Utifrån observationerna formuleras sedan teorin (Rosling, 2002; Sohlberg & Sohlberg, 2008). Inom forskning är dessa två förhållningssätt sällan helt renodlade utan det finns inslag av både deduktion och induktion. På samma sätt är det med konceptualiseringen. Det är ingenting som säger att den antingen ska göras teoretiskt eller empiriskt. Det bästa är om det finns tid och resurser för båda och ett samspel dem emellan.

Mer att definiera

Det är inte endast de koncept som är undersökningens fokus som behöver definieras, utan det kan också finnas andra nyckelbegrepp i mätningen. Om undersökningens syfte till exempel är att mäta hur studenter upplever sin studiesituation måste vi bland annat utreda vad vi menar med konceptet studiesituation och sedan operationalisera detta. Det kommer antagligen att resultera i att vi har indikatorer där begreppet lärare ingår, vars begrepp också måste definieras. En sådan definitionsprocess uppfattas kanske inte som en konceptualisering och blir sannolikt inte heller särskilt komplex. Men när begrepp finns med i indikatorer och enkätfrågor måste vi veta vad vi menar med dem. På det sättet kan vi ställa enkätfrågorna så att de som svarar uppfattar dem på det sätt vi avsåg.

Det kan också handla om att definiera den population som ska undersökas. Om syftet med en studie är att undersöka hur lärare hanterar problematiska elever måste det göras en konceptualisering av "problematiska elever" som sedan ska operationaliseras. Men det är lika viktigt att definiera vad som menas med lärare, det begrepp som definierar populationen. Vilka ingår egentligen i begreppet? De med lärarexamen, de som arbetar som lärare eller de som är med i Lärarnas riksförbund? I praktiken blir vi ofta begränsade av hur vi kan konstruera en urvalsram.

Sammanfattande punkter:

- Konceptualisering av undersökningens nyckelbegrepp bör vara det första steget i en undersökning. Innan vi kan gå vidare och mäta måste vi veta vad det är vi ska mäta.

Operationalisering – det andra steget: att konstruera indikatorer

Efter att vi tillräckligt väl definierat konceptet är nästa steg operationaliseringsprocessen. Det handlar om att bestämma vilka indikatorer vi ska ha för att fånga konceptet. Vi kan ta hjälp av de underkategorier eller komponenter som fanns med i konceptualiseringen, men även dessa måste konkretiseras till mätbara indikatorer. Ofta är det de mätbara indikatorerna som motsvaras av variabler. (I kapitlets inledande delar har skillnaden mellan indikatorer och enkätfrågor inte varit viktig. Här framgår det dock att det i många fall finns ett värde i att skilja dessa delar åt.)

Ta konceptet brott som exempel. Brottsförebyggande rådet genomför regelbundet frågeundersökningar om ungdomars delaktighet, rädsla och utsatthet för brott (Brå, 2010). I dessa mätningar måste bland annat konceptet brott definieras för att sedan operationaliseras till mätbara indikatorer. När det gäller brott finns det en etablerad definition av "antalet gärningar beskrivna i Brottsbalken och som är belagda med straff" (BrB 1:1). Är det denna formella definition man vill utgå från behöver konceptualiseringen inte bli så komplicerad.¹ Eftersom själva definitionen tydligt uttrycker att det finns delkomponenter blir det också uppenbart att vi måste använda dessa för att förstå vad vårt huvudbegrepp innebär. Delkomponenterna blir då de brottsrubriceringar som finns listade i brottsbalken, till exempel rån, grov misshandel och olovligt förfogande. Listan är lång men de är alla delkomponenter som ingår i vår definition av konceptet brott. Frågan blir sedan hur vi ska kunna mäta konceptet utifrån den utförliga definitionen. Ska vi ha indikatorer som representerar varje enskild brottstyp – är det möjligt att ställa så många frågor? Kanske måste man välja ut vissa brottstyper som

¹ Det finns alternativa definitioner och fördjupade diskussioner om vad begreppet brott innebär, innehåller och hur det kan definieras (se till exempel Andersson, 2011; Maguire, 1997 och Quinney, 1970). För att kunna använda brott som ett illustrativt exempel håller vi oss dock här till det något enklare begreppet.

får fungera som indikatorer för samtliga brott som beskrivs i Brottsbalken. Räcker det med en indikator per brottstyp? För att brotten ska gå att mäta krävs det sannolikt att de konkretiseras och att indikatorerna består av påtagliga händelser. Men hur många indikatorer med händelser måste vi då ha för att mäta den enskilda brottstypen? Ju konkretare vi väljer att ha våra indikatorer, desto fler indikatorer behövs det för att fånga konceptet. Det riskerar att göra enkäten lång och komplicerad och därmed öka respondentbördan. Låter vi däremot indikatorerna ha en lägre grad av konkretion skapar vi istället ett stort utrymme för olika tolkningar och har mindre kontroll över vad vi mäter. Om vi för att mäta delkomponenten bedrägeri formulerar en indikator om att *ha lurat folk på pengar eller något annat värdefullt* så är vår indikator mer konkret än delkomponenten, men det finns fortfarande förhållandevis stort utrymme för olika tolkningar. Vad innebär till exempel "folk" och var går gränsen för att någonting ska räknas som "värdefullt". Vi kan specificera vår indikator ytterligare för att begränsa tolkningsutrymmet och välja formuleringen *ha lurat föräldrar för att få köpa ny mobiltelefon*. En sådan indikator är mer konkret och begränsar tolkningsutrymmet, men problemet är att vi fångar betydligt färre händelser. Det kommer att krävas många fler indikatorer för att fånga in hela delkomponenten bedrägeri. Det handlar således om en balansgång där ett riktmärke bör vara att inte ha fler indikatorer än vad som är absolut nödvändigt och inte göra frågorna mer generella än vad som behövs. Om det vi vill veta handlar om det specifika brottet snatteri ska vi inte ha indikatorer som mäter det mer generella begreppet stöld.

Olika mätningar kräver olika antal nivåer av indikatorer innan man når den sista nivån utifrån vilken frågorna kan formuleras. Hur många olika nivåer det blir mellan det övergripande huvudkonceptet och de konkreta indikatorerna beror på hur brett och komplext konceptet är. Det beror även på vilken abstraktionsnivå det ligger på från början. Ett begrepp som "jämlighet" skulle sannolikt kräva en avancerad konceptualisering och innehålla många steg i operationaliseringsprocessen. Ett begrepp som till exempel alkoholkonsumtion är mer konkret från början. Både konceptualisering och operationalisering blir därmed mindre komplex och stegen från koncept till enkätfrågor blir inte lika många. Naturligtvis går det att konkretisera även förhållandevis konkreta koncept ytterligare. Alkoholkonsumtion kan specificeras och delas in i ett flertal olika aspekter som rör tid, rum och alkoholtyp. Utgångspunkten måste vara att konkretisera till den nivå som behövs för att mätningen ska fungera

tillfredsställande. Vi vill begränsa mängden indikatorer för att göra genomförandet av studien praktiskt möjligt och insamlad data möjlig att hantera och bearbeta. Samtidigt ska man komma ihåg att alkoholkonsumtion i ett annat sammanhang kan vara en indikator på ett koncept på en högre abstraktionsnivå, till exempel problembeteende eller hälsa. Vad som är ett koncept respektive en indikator beror således på vad som är den huvudsakliga problemformuleringen och vad som ska mätas.

Det är inte alltid helt tydligt var gränsen går och vad skillnaden är mellan de olika delarna här. När handlar det om underkategorier av konceptet eller konceptuella komponenter och när är det steg i operationaliseringen? Den viktigaste skillnaden är kronologin – först definierar vi vad det är som ska mätas och sedan definierar vi hur det blir mätbart.

Sammanfattande punkter:

- Efter att koncepten definierats är det dags för operationalisering. Koncepten görs mätbara genom att konkretiseras i indikatorer.

Konstruera frågor – det tredje steget

När vi har definierat vårt koncept, gjort en grundlig analys av vad det innebär och lyckats identifiera indikatorer som ska fånga konceptet är det dags att konstruera frågeformuläret. Det är i frågeformuläret som konceptet verkligen ska bli någonting mätbart. Genom att ställa konkreta frågor ska vi kunna mäta så pass abstrakta fenomen som jämlikhet, brottslighet, lycka, etnicitet och rädsla. De övriga kapitlen i den här boken handlar om olika aspekter av vad man ska tänka på i utformningen av enkäter. Därför berörs området endast översiktligt här och då främst i förhållande till konceptualisering och operationaliseringsprocessen.

Graden av konkretion – subjektiva eller objektiva mått

En aspekt att ta ställning till är hur specifika eller konkreta enkätfrågor vi ska ha. Det berördes delvis ovan då vi inte vill ha fler indikatorer än nödvändigt, eftersom svällande variabellistor och frågeformulär gör undersökningar svårhanterliga. Men graden av konkretion handlar också om balansgången mellan hur mycket utrymme vi vill ge till

respondentens varierande tolkningar och hur viktigt det är att det är vår definition som frågeställare som gäller. Man kan också tala om subjektiva och objektiva mått. Subjektiva mått ger utrymme för respondentens egen subjektiva tolkning av det som efterfrågas. Objektiva mått utgår istället från frågeställarens specifika definition. Vi kan ta en fråga om sömnsvärigheter som exempel. Säg att vi vill undersöka hur många som lider av sömnsvärigheter. Vi antar att sömnsvärigheter i det här fallet är en indikator på det övergripande teoretiska konceptet hälsa. I konceptualiseringen av hälsa och sömnsvärigheter har vi tagit ställning till att ha en konkret objektiv definition av sömnsvärigheter. Vi vill alltså ha ett objektiva mått. Har vi en sådan definition måste också frågan ställas så att den fångar sömnsvärigheter på det här sättet. För att minska respondenternas tolkningsutrymme kan det till exempel vara en idé att inte använda ordet sömnsvärigheter i frågan, utan istället använda en mer konkret formulering som stämmer med vår definition. På samma sätt konstrueras svarsalternativen i objektiva frekvenser för att vi ska kunna mäta enligt den konkreta definitionen och för att respondenterna inte ska variera i hur de uppfattar svarsalternativen. Det är inte relevant huruvida respondenten själv tycker sig lida av sömnsvärigheter eller inte. Det är vår definition som gäller och genom att ställa frågan på det här sättet kommer vi att kunna urskilja de individer som lider av sömnsvärigheter.

EXEMPEL:

Under den senaste månaden, hur många gånger har du, fast du har försökt sova, varit vaken mer än en timme?

Ingen gång

1–3 ggr

4–6 ggr

7–9 ggr

10 ggr eller fler

Om vi i konceptualiseringen däremot har definierat sömnsvärigheter som någonting subjektivt, det vill säga att de objektiva frekvenserna inte är relevanta utan istället är det respondentens upplevelse som har betydelse. Då vill vi istället mäta respondenternas subjektiva uppfattning och egen definition av sömnsvärigheter – om de upplever att de har sömnsvärigheter eller inte. Vi vill ha ett subjektivt mått. Då formuleras frågan inte lika konkret:

EXEMPEL:**Anser du dig lida av sömnsvärigheter?**

Ja, i stor utsträckning

Ja, i viss utsträckning

Ja, men endast i liten utsträckning

Nej

Den här uppställningen gör att vi inte kan få veta den konkreta innebörden av svaren. Någon respondent som ofrivilligt ligger vaken mer än en timme en gång i veckan kommer sannolikt att svara "Ja, i stor utsträckning" medan en annan med samma erfarenhet istället svarar "Ja, men endast i liten utsträckning". Det enda vi får veta är i vilken utsträckning respondenten själv upplever sig ha sömnsvärigheter, och i det här fallet var det också precis det vi ville veta. Vi ansåg det mer intressant hur respondenten själv uppfattar sin situation än att få kunskap om den konkreta frekvensen.

Användningen av subjektiva eller objektiva mått och frågeuppställningar hänger således ihop med hur vi definierar det övergripande konceptet och alltså vad det är vi vill mäta. I det här exemplet beror det på om sömnsvärigheter är en indikator på objektiv hälsa – där vi inte bara har definierat komponenter och indikatorer utan också vad som är bra respektive dålig hälsa utifrån på förhand uppställda kriterier. Eller om vi ska mäta subjektiv hälsa – där några kriterier inte är fastställda utan där det är respondenternas egen upplevelse som är av intresse.

Som framgått ovan är det även centralt att ta ställning till hur svarsalternativ ska ställas upp och hur svar och resultat ska hanteras och kategoriseras. Detta är en del av konceptualiseringen och operationaliseringen, men blir också väldigt påtagligt när vi ska formulera själva frågorna och svarsalternativen. Vad är exempelvis god hälsa och hur ska det passa in i det övergripande konceptet levnadsförhållanden? Vad är låg respektive hög alkoholkonsumtion? Genom att kategorisera våra svar definierar vi också begreppen. Hur vi ställer upp svarsalternativen avgör vilka kategoriseringar och analyser som är möjliga att göra på det insamlade materialet. Om vi delar in befolkningen i grupperna låg-, medel- och höginkomsttagare definierar vi samtidigt dessa kategorier (Sohlberg & Sohlberg, 2008). Ofta är det möjligt att redan i konceptualiseringen konstruera teoretiskt förankrade kategorier. Det kan till exempel vara att hög alkoholkonsumtion är att

dricka mer alkohol än motsvarande två flaskor vin i veckan. Frågor och svarsalternativ formuleras då på ett sätt så att vi i analysen kan särskilja grupperna utifrån våra teoretiska kategoriseringar. På motsvarande sätt innebär en subjektiv svarsskala en subjektiv kategorisering. I en sådan subjektiv skala, som i exemplet ovan om sömnsvårigheter, bör kategoriseringen fortfarande vara förankrad i konceptet och i studiens frågeställning. Om möjligt bör vi på förhand ha bestämt vilka svarsalternativ som ska motsvara allvarliga sömnsvårigheter och liknande (Dahmström, 2005). Skillnaden jämfört med en objektiv skala är att det är respondenternas subjektiva bedömningar vi är ute efter, inte att vid bearbetning av data kunna kategorisera värden subjektivt efter eget tycke och smak. Om frågan om sömnsvårigheter är en del av ett index där ett stort antal frågor ingår kommer ytterligare en kategorisering att göras i det sammanslagna indexet. Även hur den kategoriseringen görs ska vara bestämt på förhand och styras av konceptets definition.

Sammanfattande punkter:

- I det sista steget konstrueras frågorna.
- Det är med hjälp av frågorna man mäter konceptet.
- När frågan ska formuleras måste du ta ställning till graden av konkretion och hur specifikt frågans beståndsdelar ska definieras. Detta måste dock styras av konceptualiseringen.

Mätfel, mätträtt och mätningens kvalitet

En ständigt central och avgörande fråga i sammanhanget frågeundersökningar är den om undersökningarnas tillförlitlighet och giltighet. Vi vill självklart kunna lita på de undersökningar som görs, men frågeundersökningar innehåller alltid en varierande grad av osäkerhet (Wärneryd, 1989).

Inom mätteori pratar man om det sanna värdet (T), det observerade värdet (O) och mätfelet (E). Det sanna värdet är det värde man vill fånga med undersökningen. En del kanske skulle kalla det för det korrekta värdet på verkligheten. Det observerade värdet är det vi får fram i mätningen – vår observation. Det här värdet kan vara en sämre eller bättre representation av de sanna värdet. Mätfelet är skillnaden mellan det observerade och det sanna värdet ($O-T=E$). Det handlar alltså om de fel som uppstår i mätningen och som gör att det observerade värdet inte

motsvarar det sanna värdet. Det vi får fram i mätningen är alltså inte en exakt avspeglning av verkligheten (Spector, 1992; Magnusson, 2003). Mätfelets storlek handlar därmed om hur väl vi mäter konceptets sanna värde.² Ju mindre mätfelet är, desto större är tillförlitligheten.

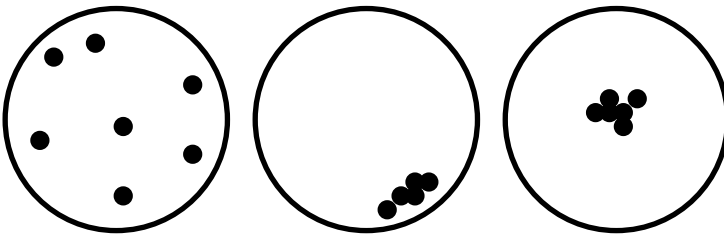
När man diskuterar fel i mätningar skiljer man ofta mellan slumpmässiga fel och systematiska fel. Slumpmässiga fel innebär att slumpen gjort att det observerade värdet inte motsvarar det sanna värdet och att vi därför får varierande resultat vid upprepade mätningar med samma mätinstrument. Det skulle till exempel kunna handla om att min kylskåpstermometer visar olika temperaturer varenda gång jag tittar på den, fast jag märker att mjölken är lika kall. I beskrivningar, diskussioner och resonemang om slumpmässiga fel används begreppet reliabilitet, som handlar om tillförlitlighet och stabilitet. Om skillnaden mellan det observerade värdet och det sanna är systematisk pratar vi istället om systematiska fel. Då är det inte slumpen som orsakat felet utan det finns ett regelmässigt fel i mätningen. Bara för att min kylskåpstermometer alltid visar samma temperatur behöver det inte innebära att den visar rätt. Kanske är det så att den vid upprepade kontroller visar på minusgrader i kylen, men jag märker ju att mjölken inte är djupfryst. När det gäller den här typen av fel brukar man prata om bristande validitet eller giltighet. Mätningen är stabil, men den är inte giltig. En mätning kan således vara reliabel utan att vara valid (Carmines & Zeller, 1979).

Förhållandet mellan reliabilitet och validitet kan beskrivas med hjälp av skissen nedan. Vi tänker oss att mitten av cirklarna är det sanna värdet, det vi vill fånga med våra mätningar: I den första cirkeln träffar vi olika hela tiden och resultaten ser ut att vara slumpmässiga. Vi kan därmed anta att felen är slumpmässiga – träffarna är inte i närheten av varandra och endast en gång prickar vi mitten (mätningen träffar sanningen). Det finns ingen stabilitet och reliabiliteten är således låg. I den andra cirkeln är reliabiliteten hög. De olika mätningarna ger resultat som ligger i närheten av varandra. Det verkar inte finnas några slumpmässiga fel utan vi mäter hela tiden ungefär samma sak. Problemet är dock att även om vi mäter samma sak med de olika mätningarna så mäter vi inte det vi avser. Klustret av mätresultat i cirkeln ligger långt ifrån mitten och felen

² I vilken mån "sanning" och "sant värde" är funktionella och meningsfulla begrepp beror dels på typ av mätning, dels på kunskaps- och verklighetssyn. Det kan till exempel handla om att det i vissa fall kan vara problematiskt att hävda att det existerar ett objektivt sant värde och definitionen av "mätfel" bör då modifieras (se bl.a. Andersson, 2011).

verkar snarare vara systematiska. Validiteten är således väldigt bristfällig och det blir uppenbart att reliabiliteten kan vara hög samtidigt som validiteten är låg. I den tredje cirkeln har samtliga mätningar träffat nära mitten och prickat det som vi avsåg att mäta. Därmed är validiteten hög. Tillsammans visar de tre cirklarna att om reliabiliteten är låg är validiteten det också. För att validiteten ska vara hög måste mätningen vara stabil och prickarna kan inte vara utspridda i cirkeln. Hög reliabilitet förutsätter dock inte hög validitet. Mätningen kan vara stabil och prickarna samlade utan att vi mäter det vi ska (se Wolming, 1998).³

Figur 8.1



Hög validitet som mål

Validitet och reliabilitet är i grund och botten en fråga om kvalitet. Man skulle kunna säga att hela den här boken handlar om validitet och reliabilitet eftersom syftet är att förmedla kunskap om hur frågeundersökningar håller så hög kvalitet som möjligt. Alla delar av planering, design och genomförande är avgörande för att en undersökning ska vara tillräckligt valid och reliabel. Men grunden för den fortsatta kvaliteten är koncept och operationalisering. Det kan ju aldrig gå att mäta det vi avser om vi inte vet vad det är vi avser att mäta – själva definitionen av konceptet. Om de indikatorer vi har formulerat inte motsvarar det definierade konceptet kommer vi inte att genomföra en giltig studie oavsett hur bra urval, insamlingsmetod och bearbetning av insamlad data fungerar (Wolming, 1998). Genom att utforma indikatorer som motsvarar konceptet och utifrån det formulera frågor som

³ Om man tittar på träffarna i just det här exemplet kan man också konstatera att en mätning med låg reliabilitet (den första cirkeln) sammantaget **kan** ge ett resultat med hög validitet än en mätning med hög reliabilitet. Spridningen i den första cirkeln är stor, men mätningarnas genomsnittliga värde ligger nära det sanna värdet. Detta till skillnad från cirkeln i mitten som får ett genomsnittligt värde långt från det sanna värdet.

respondenterna förstår, tolkar på det sätt vi avser och kan besvara, skapar vi grundläggande förutsättningar för en reliabel och valid studie.

En viktig aspekt att tänka på när det gäller validitet och reliabilitet är att en mätning aldrig kan vara helt utan fel och därmed heller aldrig fullständigt valid. Utifrån den mätteoretiska formeln skulle det innebära att det observerade värdet och det sanna värdet aldrig är helt överensstämmande och att det således alltid existerar en viss grad av fel. Både validitet och reliabilitet bör ses som kontinuum där graden av giltighet och tillförlitlighet helt enkelt varierar. Det är inte svart-vita begrepp som kan göra antingen tummen upp eller tummen ner gentemot en studie (Cronbach, 1988; Messick, 1989). Det handlar istället om att ständigt sträva efter en så bra mätning som möjligt.

När det gäller mätfel, mätkvalitet och validitet förekommer det en rad olika typer av validitet och validitetsbegrepp i litteraturen där validitetsbedömningen fokuserar på olika aspekter av giltighet. Benämningar och indelningar av de olika validitetstyperna skiljer sig något mellan olika författare, men i stora drag är hanteringen av begreppen detsamma. Gemensamt är fortfarande frågan huruvida mätningen mäter det som avses (Cronbach & Meehl, 1955; Carmines & Zeller, 1979; Kvale, 1997; Junger-Tas & Marshall, 1999; Björkman, 2000).

Björkman (2000) skiljer mellan inre (logisk) och yttre (empirisk) validitet. Inre validitet skulle kunna beskrivas som en skrivbordbedömning av kvaliteten där operationalisering och frågekonstruktion granskas av både inblandade och utomstående forskare och utredare. Vid en bedömning av den yttre validiteten gör man istället en empirisk granskning av mätningens kvalitet. Björkman delar in den yttre validiteten i kriterievaliditet och validitet utan kriterium. Kriterievaliditet fungerar när det finns ett annat mått att jämföra med och som kan fungera som ett kriterium. Exempelvis kan man på en begränsad del av urvalet ibland komplettera måttet med en mer kostnadskrävande metod som man vet ger giltiga resultat. Genom att jämföra de två måtten får man då en indikation på validiteten hos den billigare metoden. Oftast finns dock inte något sådant fungerande kriterium och man måste då använda andra sätt för att kunna göra en empirisk bedömning. En sådan strategi kan vara att kontrollera innehållsvaliditeten. Det innebär att man undersöker om mätningen fångar samtliga relevanta aspekter av det som ska mätas. Är indikatorer och enkätfrågor tillräckliga för att mäta konceptet? Den här bedömningen görs ofta av experter inom det område

som undersöks. En ytterligare strategi är att undersöka den så kallade begreppsvaliditeten ("construct validity" på engelska). Den handlar om relationen mellan det aktuella måttet och ett annat koncept där det antas finnas ett beroende. Om de båda koncepten korrelerar på det sätt man förväntar sig kan man anta att måtten är någorlunda valida (Cronbach & Meehl, 1955; Carmines & Zeller, 1979). Men begreppsvalidering behöver inte enbart handla om kvantitativa beräkningar av korrelationer. Det kan också handla om en kvalitativ bedömning av det logiska sambandet mellan olika koncept och mått. Det innebär att validering är en process med en ständigt pågående strävan efter att dels utveckla så bra mått som möjligt, men också att göra relevanta och valida tolkningar av de resultat som mätningarna producerar (se Kvale, 1997; Andersson, 2011).

Även reliabiliteten kan testas med empiriska metoder. Den kanske mest uppenbara är den så kallade test-retest-metoden. Där används samma mätinstrument på samma grupp två eller flera gånger med en viss tidsperiod emellan. Syftet är att därefter kunna jämföra resultaten och skaffa sig en uppfattning av stabilitet och frånvaro eller närvaro av slumpmässiga fel (Wärneryd, 1989; Eljertsson, 2005). En annan vanlig metod är att undersöka den interna konsistensen. Det är främst relevant när man har flera indikatorer eller frågor som är tänkta att tillsammans fånga ett abstrakt koncept. Man brukar då säga att man har konstruerat ett index med flera indikatorer för att mäta konceptet, och det är konsistensen inom detta index som undersöks. För att indexet ska fungera och för att man ska kunna anta att det tillsammans mäter konceptet bör det finnas en samstämmighet inom indexet. De olika indikatorerna ska korrelera med varandra och hur väl de gör det blir ett mått på reliabiliteten. Den interna korrelationen beräknas ofta genom Cronbach's Alpha som anges mellan 0 och 1. Ett högt värde visar på en hög internkorrelation och därmed en hög grad av stabilitet och reliabilitet (Carmines & Zeller, 1979; Eljertsson, 2005).

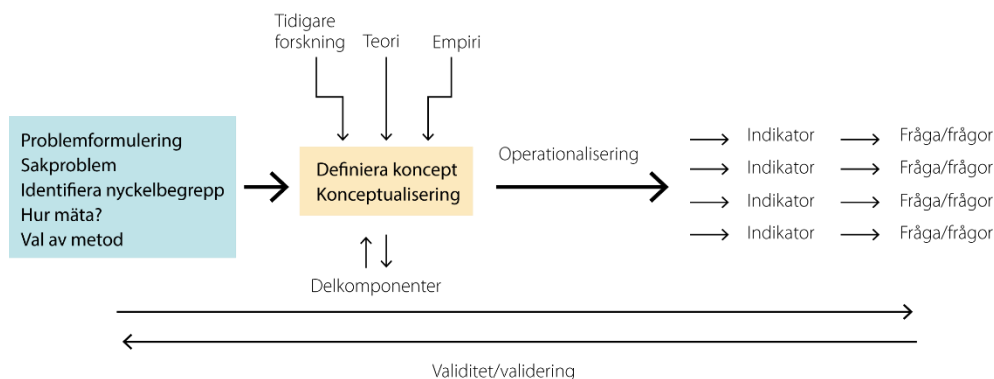
I förhållande till det här kapitlets tema, konceptualisering och operationalisering, är det bästa sättet att förbättra validitet och reliabilitet att testa sitt instrument i olika skeden av processen. Förutom att använda litteratur och tidigare forskning för att utreda konceptet kan det utvecklas empiriskt genom till exempel fokusgrupper, indikatorer och enkätfrågor. Det kan även testas genom kognitiva intervjuer och ett färdigt frågeformulär kan prövas i en pilotstudie (mer detaljerade beskrivningar om hur olika test fungerar finns i kapitel 14).

Koncept, operationell definition och enkätfrågor

Teoretiskt är de tre stegen som presenteras ovan olika delar (definiera konceptet, operationalisera det och konstruera enkätfrågor). I praktiken är det dock inte alltid möjligt att hantera dem separat. Till exempel kan det ibland vara aktuellt att integrera problemformulering i metodval och konceptualisering. En princip som dock inte ska överges är att hela tiden sträva efter att definiera konceptet innan man går vidare med mätningen. Om konceptualiseringen inte lyckas kommer inte heller mätningen att göra det (Antil, 1984). Kanske får man då helt enkelt förändra syftet med undersökningen. I den övergripande problemformuleringen frågas istället efter något som går att definiera och därmed också är möjligt att operationalisera. Vad som ska mätas får ibland styras av vad som går att mäta (Wärneryd, 1989).

Figuren nedan ska fungera som en sammanfattning av det förlopp som kapitlet har handlat om – att identifiera vad det är som ska studeras och hur, definiera de relevanta koncepten (konceptualiseringen), göra koncepten mätbara i indikatorer (operationaliseringen) och sedan formulera frågor. Validitet och validering löper som en pågående process i båda riktningarna utmed hela förloppet. Det är för att markera att validitet inte endast är en bedömning av hur väl slutprodukten fångat konceptet, utan berör samtliga val och överväganden som är aktuella i de olika stegen.

Figur 8.2



De olika stegen:

Steg 1) Problemformulering och val av metod

Steg 2) Konceptualisering

Steg 3) Operationalisering

Steg 4) Konstruktion av frågeformulär och frågor.

Man ska således *inte* tänka "De här frågorna vill jag ställa, vad mäter jag?", utan istället "Vad är det jag vill undersöka och hur ska jag göra det? Hur ska jag konstruera min enkät och ställa mina frågor för att undersöka det jag vill veta?". Om man följer det sistnämnda har man åtminstone uppfyllt de första grundförutsättningarna för en valid och meningsfull studie.

Sammanfattning

Att mäta med frågor - sammanfattning

- Kvalitativa och kvantitativa undersökningar ger delvis olika typer av kunskap. Hur man väljer att inrikta sin undersökning metodologiskt måste styras av problem- och frågeställning.
- En kvantitativ frågeundersökning är ofta en bra metod, men inte alltid.
- När det som ska mätas är direkt observerbart – det går till exempel att se, höra eller känna på objektet – säger man att mätningen är direkt.
- När objektet är abstrakt och inte går att observera utarbetas indikatorer som är mätbara. Mätningen är då indirekt.
- Konceptualisering av undersökningens nyckelbegrepp bör vara det första steget i en undersökning. Innan vi kan gå vidare och mäta måste vi veta vad det är vi ska mäta.
- Efter att koncepten definierats är det dags för operationalisering. Koncepten görs mätbara genom att konkretiseras i indikatorer.
- I det sista steget konstrueras frågorna.
- Det är med hjälp av frågorna man mäter konceptet.
- När frågan ska formuleras måste du ta ställning till graden av konkretion och hur specifikt frågans beståndsdelar ska definieras. Detta måste dock styras av konceptualiseringen.

9 Svartsprocessen och felkällor

Författare: Andreas Persson (SCB)

Inledning

Detta kapitel innehåller ligger främst på en övergripande och teoretisk nivå. Det beskriver hur surveyforskningen med olika modeller och teorier har försökt förklara vad som händer från det att respondenterna har fått en fråga tills att de har gett ett svar. Det är detta förlopp som kallas för svarsprocessen. Kapitlet beskriver även andra aspekter av mätsituationen, till exempel vilka källor som kan leda till fel i mätningen och vilka olika typer av fel som brukar förekomma i surveyundersökningar. Det finns viktiga skillnader i mätsituationen mellan undersökningar riktade till individer och hushåll och undersökningar riktade till företag och organisationer. Därför diskuteras dessa två varianter i separata avsnitt senare i kapitlet.

Forskning om frågekonstruktion – historik och teoriutveckling

Att frågekonstruktionen har betydelse för mätningen har varit känt länge. Även inom surveyforskningen uppmärksammades tidigt frågekonstruktionens betydelse och dess påverkan på undersökningsresultatet. På 1930- och 1940-talen utfördes mycket forskning, både om frågeformuleringar och om intervjuareffekter. Frågekonstruktion förblev dock ett hantverk som baserades mer på omdöme och erfarenhet än på vetenskapliga forskningsresultat. Förmodligen berodde detta på att forskningens teoriförankring var svag och att det därför var svårt att generalisera resultaten. Exempelvis kan ett experiment visa att fråga A är bättre än fråga B i något avseende. Men utan förklaringar om *varför* A är bättre än B är det svårt att generalisera resultatet till fråga C, D, E, F och så vidare. Kunskapen från forskningen blev väldigt begränsad och takten framåt långsam. Med tiden tynade därför intresset för och forskningen om frågekonstruktion bort något (Sudman, Bradburn & Schwarz, 1996).

Ett kognitivt skifte på 1970-talet

På 1970-talet växte intresset för frågekonstruktion igen. Bradburn och Sudman publicerade en metaanalys om svareffekter. I det arbetet lanserade de också en modell för fel i en intervju-situation (Sudman & Bradburn, 1974; Sudman, Bradburn & Schwarz, 1996). Parallellt med detta skedde ett skifte inom psykologin från behaviorismens stimuli-respons-förklaringar till ett fokus på informationsprocesser. Det var ett kognitivt skifte som medförde att språkförståelse, minnet och hur människor gör bedömningar och uppskattningar åter blev populära forskningsobjekt.

I slutet på samma årtionde började kognitiva psykologer och survey-forskare samverka. Detta skedde först genom ett antal konferenser, både i Nordamerika och i Europa. I och med detta samarbete etablerades utgångspunkter och modeller för fortsatt empiriskt arbete (Sudman, Bradburn & Schwarz, 1996). Sedan dess har kognitionsvetenskapen haft en given plats inom surveyområdet. Exempelvis har många statistikbyråer i Nordamerika och Västeuropa kognitiva laboratorier för frågetestning sedan några decennier tillbaka (Willis, 2005). Forskningen om och teoriutvecklingen för frågekonstruktion har också ökat kraftigt de senaste decennierna. Även om det fortfarande finns många frågetecken och outforskade områden är frågekonstruktionsområdet idag bättre rustat att erbjuda generella förklaringar och regler (Sudman, Bradburn & Schwarz, 1996).

Ett resultat av arbetet beskrivet ovan är bland annat ett antal modeller över mätsituationen eller svarsprocessen. Med svarsprocessen menas alltså den process som inleds med att respondenten exponeras för en fråga och avslutas med att respondenten lämnar ett svar. Vid individ- och hushållsundersökningar är processen en kognitiv resa som sker främst inne i respondentens huvud. Vid undersökningar riktade till företag och organisationer tillkommer organisatoriska faktorer. Med mätsituationen avses oftast något mer övergripande och sådana modeller inkluderar även annat i situationen, till exempel närvaron av en intervjuare. En del av dessa modeller redovisas och diskuteras i följande avsnitt. Modellerna har bland annat följande syften:

- att ge en överblick över processerna
- att visa på interna relationer i ett system
- att sammanfatta och integrera tidigare forskning
- att ge ett ramverk att tolka ny kunskap mot

- att ge uppslag till framtida forskning, till exempel att vara utgångspunkt för hypoteser.

Sammanfattande punkter:

- Det har länge varit känt att frågornas utformning påverkar svaren. Forskning genomfördes redan på 1930- och 1940-talen.
- På 1970-talet utvidgades det kognitiva perspektivet med tankeprocesser inom psykologin.
- Kognitiva psykologer och surveyforskare började samverka. Idag är kognitionspsykologin given inom surveyområdet.
- Samverkan resulterade bland annat i ett antal modeller som syftar till att beskriva mätsituationen eller svarsprocessen (från det att respondenten får en fråga tills ett svar har lämnats) i en undersökning.

Kognition och kommunikation i svarsprocessen

Modeller över mätsituationen

Bradburn och Sudman (1974) beskrev surveyintervjun som en social situation. Den sker i form av ett samtal med en given struktur fast med sina egna regler som gör att den skiljer sig från ett samtal mellan vänner. Utifrån denna utgångspunkt identifierade de tre källor till mätfel: 1) respondenten, 2) intervjuaren och 3) frågeformuläret. Faktorer förknippade med respondenten kan vara dennes kunskap, motivation eller attityd till undersökningar. Faktorer förknippade med intervjuaren kan vara dennes intervjuarbeteende. Bradburn och Sudmans metaanalys av svarsfel antydde dock att de flesta felen härstammade från frågeformuläret. Det kunde till exempel vara att frågan innehöll vaga ord eller begrepp som kan betyda olika saker i olika sammanhang (Sudman, Bradburn & Schwarz, 1996).

Bergman och Wärneryd (1982) presenterar också en modell som har ett stort inslag av kommunikation. Modellen inleds med att en fråga ställs och avslutas med en markering i frågeformuläret. Den innefattar en yttre ram som kan påverka systemet, till exempel tid, plats eller intervjuarens egenskaper. I centrum för modellen finns respondentens kognitiva aktivitet som exemplifieras med dennes kunskap, attityd och tidigare erfarenheter av undersökningar. I övrigt består modellen av

feedbackprocesser, både vad gäller förståelsen av och svaret på frågan. En intervjuare kan till exempel ge ytterligare information för att respondenten ska förstå frågan och reagera uppmuntrande eller inte på respondentens svar. Feedback förekommer även vid självadministrerade frågeformulär, till exempel att respondenten måste läsa om frågan vid missförstånd eller att ett visst svar ger särskilda följdfrågor.

Båda dessa modeller beskriver mätsituationen. De inkluderar de relevanta aktörerna och relationerna dem emellan. De har fokus på intervjuundersökningar. Som modeller över mätsituationen är dessa och andra liknande modeller värdefulla. Som utgångspunkt för frågekonstruktion är det dock värdefullt med modeller som sätter de kognitiva aspekterna av svarsprocessen (språkförståelse, minnesprocesser etc.) i centrum och i mer detalj (Tourangeau, Rips & Rasinski, 2000).

Modeller över den kognitiva processen

En tidig modell

En av de första modellerna som inkluderar kognitiva aspekter lanserades av Cannell, Miller och Oksenberg (1981) (Tourangeau, Rips & Rasinski, 2000). Modellen innehöll två möjliga vägar fram till ett svar – en noggrann och en ytlig. Den noggranna antas leda fram till adekvata svar. I den ytliga vägen är risken för inkorrekta svar betydligt större. Den noggranna vägen innehåller fem steg som alla innebär relativt djup tankeverksamhet hos respondenten. I den ytliga vägen baseras svaret däremot på signaler, frågeformuleringen och ledtrådar eller antydningar från intervjuaren eller respondentens egna värderingar och mål. Svar baserade på den ytliga vägen är därmed genvägar och inte lika genomtänkta. De löper därför större risk för effekter som till exempel social önskvärdhet eller hålla-med-tendenser (Tourangeau, Rips & Rasinski, 2000).

Fyrstegsmodellen – den mest spridda

Den kognitiva modell över svarsprocessen som fått mest spridning och som fokuserar just på hur respondenten kommer fram till ett svar är Tourangeaus fyrstegsmodell (Tourangeau, Rips & Rasinski, 2000; Willis, 2005:). Stegen är:

1. **Förståelse av frågan:** Vad betyder enskilda ord och termer? Vad är det som efterfrågas?
2. **Återerinring från minnet:** Vilken information måste återerinnras? Vilken återerinnringsstrategi använder respondenten?
3. **Bedömning:** Bearbetning av den återhämtade informationen utifrån hur frågan har förståtts.
4. **Lämnande av svar:** Anpassning av det individuellt formulerade svaret till de befintliga svarsalternativen i frågeformuläret.

Modellen fokuserar på de kognitiva delarna av svarsprocessen. För att ge ett korrekt svar bör oftast alla steg generellt gås igenom av respondenten. Det är dock inte alltid så är fallet. Respondenten kan till exempel hoppa över frågan och därmed alla de fyra stegen eller ge ett svar utan att ha gått igenom de tre första stegen (alltså bara kryssa något utan att ens läsa frågan ordentligt).

Satisficing-modellen

En annan modell som fått stor spridning i surveylitteraturen är Krosnick och Alwins satisficingmodell (Krosnick & Alwin, 1987; Tourangeau, Rips & Rasinski, 2000). Med satisficing menas en beslutsstrategi där individens mål är att prestera *tillräckligt* bra snarare än optimalt. Ibland kan människor istället för att söka den optimala lösningen på ett problem godta en lösning som är *tillräckligt* bra. Att söka den optimala lösningen kräver psykologisk ansträngning och alla beslut är inte värda den mödan. Denna strategi kan även tillämpas på frågeformulär och dess frågor. Enligt Krosnick och Alwin (1987) kan respondenter alltså försöka optimera sina svar, till exempel lägga kraft på att försöka förstå frågan korrekt och söka rätt svar i minnet. Andra respondenter kan ha en strategi där de tycker att det räcker med att hitta ett godtagbart men inte optimalt svar. Detta ska inte uppfattas som att respondenten nödvändigtvis svarar vad som helst slumpmässigt, utan att respondenten lägger mindre möda på att hitta det mest korrekta svaret. Tendensen kan därför vara svår att upptäcka, för den ger inte nödvändigtvis några tydliga felsignaler.

Krosnick och Alwin (1987) beskriver det som två alternativa rutter men förmodligen är det en dimension längs vilken personer kan variera (Tourangeau, Rips & Rasinski, 2000). Satisficing är relaterat till motivation, kognitiv kapacitet hos respondenten och kraven i frågeformuläret (Krosnick, 2000). Modellen har använts som förklaring

till bland annat ordningseffekter som primacy-effekten och återkommer på andra platser i denna bok (läs mer i kapitel 13).

Modellerna som helhet

Modellerna beskrivna ovan har olika inriktningar och ändamål. Sammantaget ger de en god bild över vilka faktorer som kan påverka resultatet i mätsituationen och som man bör ha i åtanke vid frågekonstruktion:

- den yttre ramen, till exempel tid och plats
- intervjuaren som förmedlare av frågan och genom feedback till respondenten
- respondentens förståelse av frågan
- respondentens återerinring av information från minnet
- respondentens bedömningar av återhämtad information
- respondentens anpassning av sitt eget svar till frågeformulärets svarsalternativ
- respondentens inställning till besvarandet, till exempel i vilken grad optimala svar eftersträvas.

Sammanfattande punkter:

- Det finns flera modeller för mätsituationen i en undersökning. De inkluderar till exempel respondenten, intervjuaren, frågeformuläret och de yttre ramarna (som tid och plats).
- För frågekonstruktion behövs dock en modell som mer specifikt fokuserar på hur respondenten förhåller sig till själva frågorna, dvs. modeller som fokuserar på svarsprocessen.
- Den mest spridda modellen över svarsprocessen är den så kallade fyrstegsmodellen:
1) förståelse av frågan, 2) återerinring från minnet, 3) bedömning, 4) lämnande av svar.
- En annan spridd modell – satisficingmodellen – belyser hur respondenten kan sträva efter att antingen finna ett optimalt svar eller ta genvägar och nöja sig med ett svar som är tillräckligt bra.

Svarsprocessen i företag och organisationer

Mindre forskning och teori än för individer och hushåll

I undersökningar riktade till individer och hushåll är svarsprocessen förhållandevis enkel. Med det menas att det oftast bara är en respondent inblandad och att det finns kontaktuppgifter till den personen (till exempel adressregister). I undersökningar riktade till företag, myndigheter och organisationer är situationen annorlunda. Det kan vara flera respondenter inblandade i besvarandet, till exempel olika respondenter för olika frågor i frågeformuläret. Vem på företaget eller i organisationen som är bäst lämpad att svara är ofta okänt på förhand. Kognitiva faktorer är aktuella även vid företagsundersökningar eftersom det också där är en människa som besvarar formuläret. Men frågorna besvaras generellt sett i lägre grad utifrån respondentens minne och i högre grad utifrån uppgifter hämtade från datasystem och register. Allt detta bidrar till att svarsprocessen i undersökningar riktade till företag och organisationer avviker från den till individer och hushåll. Ovanstående modeller är därför otillräckliga.

Forskningen om frågekonstruktion har länge fokuserat på undersökningar riktade till individer och hushåll. Kanske för att den har letts av psykologer med större intresse för individer än för organisationer. Detta har bidragit till att teoriutvecklingen och även mer tillämplig forskning är mindre för företagsundersökningar än den till individer och hushåll.

Hybridmodellen

De senaste decennierna har dock företagsundersökningar fått ett allt större fokus och modeller över svarsprocessen har lanserats. En sådan modell är hybridmodellen (Willimack & Nichols, 2001; 2010; Sudman, Willimack, Nichols, & Mesenbourg, 2000). Den kallas så eftersom den innehåller både kognitiva och organisatoriska faktorer. Den består av följande steg:

- inkodning
- urval och identifiering av respondenter
- uppskattning av resurser
- förståelse av dataförfrågan
- återerinring av relevant information från minne eller datasystem
- bedömning av informationen och svaret

- kommunikation av svaret
- lämning av data.

Modellen beskriver de steg som måste gås igenom från det att ett frågeformulär anländer på till exempel ett företag till att svaren är lämnade. "Inkodning" betyder att företaget registrerat den efterfrågade informationen i sina system (motsvarigheten för en individ är när information kodas in i långtidsminnet). Sedan måste lämpliga respondenter för frågeformuläret identifieras och väljas ut. Företaget måste också ta ställning till vilka resurser som krävs och hur stora resurser som de ämnar lägga på undersökningen. Detta kan säkert variera beroende på om undersökningen är frivillig eller har uppgiftslämnarskyldighet. De kommande fyra stegen i modellen är snarlika dem i den kognitiva fyrstegsmodellen (se ovan), förutom att "återerinringen" i företagsundersökningar oftare sker från datasystem än från minnet. Det sista steget innebär att framförallt större företag ibland har en kontrollfunktion innan data lämnas ut (Willimack & Nichols, 2010).

Andra modeller för svarsprocessen i företagsundersökningar

Lorenc (2006a) använder en annan modell som bygger på idén om socialt distribuerad kognition (Hutchinson, 1995). Applicerat på företagsundersökningar ses då företaget i sig som ett kognitivt system. De kognitiva processerna i systemet består då av att information distribueras mellan de anställda (från individ till individ) inom företaget, mellan individ och företagets informationssystem (datasystem, register, tidigare lämnade uppgifter etc.) och mellan individer över tid genom till exempel etablerade arbetsrutiner inom företaget. Enligt detta sker de kognitiva processerna öppet i den fysiska världen och är därmed observerbara. Till exempel sker återerinring av data från ett system snarare än inne i någons huvud. Detta antagande möjliggör analys av källor till mätfel i högre grad än vid studiet av inre kognitiva processer. Lorenc har även applicerat teorin på en skolundersökning (Lorenc, 2006b).

Andra modeller fokuserar direkt på mätsituationen, snarare än de organisatoriska faktorerna. Dessa inkluderar de olika strategier som respondenter i företagsundersökningar kan ha (Bavdaz, 2007). De visar till exempel vilken informationskälla respondenten har tillgång till (data eller minnet), vilken typ av återerinring som kan göras och hur

informationen kan omvandlas utifrån vad som efterfrågas. Vilka möjligheter och beslut som respondenten har och tar påverkar om det lämnade svaret till exempel blir en exakt uppgift, en uppskattning eller en ren gissning (läs mer i Bavdaz, 2007).

Sammanfattande punkter:

- Svarsprocessen i företag och organisationer skiljer sig från den för individer och hushåll, bland annat för att flera respondenter kan vara inblandade och för att svaren baseras på information från datasystem snarare än individers minne.
- Det finns mindre forskning om svarsprocessen i företag och organisationer.
- Hybridmodellen innehåller både kognitiva och organisatoriska steg (därav namnet).
- Företaget i sig kan ses som ett kognitivt system (socialt distribuerad kognition).
- Ett annat perspektiv är att fokusera på hur kompatibel företagens data är med frågan (finns till exempel exakta siffror eller måste de uppskattas?).

Felkällor och mätfel

Mätfelet och dess orsaker

Mätfel brukar inom surveyområdet definieras som när det observerade värdet avviker från det sanna värdet (Hansen, Hurwitz, Marks & Mauldin, 1951). I vissa fall, till exempel attitydmätningar, finns inget sant värde i strikt betydelse. Mätfelet brukar då definieras som en avvikelse från det värde vi skulle ha fått under optimala förutsättningar (Groves, 1989; Bergman & Wärneryd, 1982; läs även mer i kapitel 8). Det krävs resurser för att skatta mätfelet. Ett sätt är att jämföra de observerade värdena med en annan datakälla till vilken man har hög tilltro, till exempel register. Detta förutsätter dock andra datakällor, vilka tyvärr sällan finns tillgängliga. Ett annat perspektiv, som inte står i konflikt med eller helt ska ersätta det förra, är att arbeta med att förebygga mätfel genom att bland annat designa sina frågor väl, vilket är denna boks fokus.

Mätfel kan grundas i (Groves, 1989)

- intervjuaren
- respondenten
- frågeformuläret
- datainsamlingsmetoden.

Betydelsen av intervjuaren och datainsamlingsmetoden diskuteras främst i kapitlet om datainsamlingsmetoder. De två andra faktorerna – respondenten och frågeformuläret – är utgångspunkterna för boken som helhet.

Nedan diskuteras först källor till fel utifrån fyrstegsmodellen (CASM) och hybridmodellen, uppdelade på undersökningar till individer och hushåll respektive till företag och organisationer. Sedan följer en redovisning av mer specifika typer av fel som kan förekomma i undersökningar.

Sammanfattande punkter:

- Mätfel är när det observerade värdet avviker från det sanna värdet.
- Ibland finns inget sant värde i strikt betydelse. Då är utgångspunkten det värde vi skulle ha fått under optimala förutsättningar.
- Mätfel kan grunda sig i intervjuaren, respondenten, frågeformuläret eller datainsamlingsmetoden.

Mätfel och modellerna för svarsprocessen

Individer och hushåll

Modellerna i avsnitten ovan beskriver hur svarsprocessen går till. De tar upp vad som händer eller vad som kan hända. Därmed ger de också information om vad som kan fallera – vad som kan utgöra källor till fel. Detta är bra att känna till och försöka undvika vid frågekonstruktion.

Vid undersökningar till individer och hushåll är svarsprocessen främst kognitiv och Tourangeaus (1984) fyrstegsmodell är utgångspunkten för resonemangen nedan.

Förståelse av frågan

Först måste respondenterna förstå frågan, instruktioner och eventuellt annat material enligt frågekonstruktörens intention. De måste förstå orden i frågan och vad de betyder. De måste förstå frågan bokstavligt men även vad frågekonstruktören faktiskt menar. Det finns en del saker som kan gå fel i denna process. Frågan kan ha ord som respondenten inte känner till betydelsen av. De kan tolka odefinierade ord och termer annorlunda än avsikten. De kan missförstå vad som efterfrågas. De kan missförstå en fråga på grund av dess sammanhang (till exempel en lärare som i en enkät svarade 23 på frågan hur många barn hon hade). De kan vara okoncentrerade och missa en instruktion eller annan viktig information. Frågan kan dessutom vara felaktigt konstruerad eller så språkligt avancerad att inte alla respondenter förstår den. Allt detta är exempel på saker som kan leda till att frågan inte förstås på rätt sätt. Om frågan missförstås är naturligtvis riskerna stora att vi får ett inkorrekt svar (Tourangeau, Rips & Rasinski, 2000; Willis, 2005).

Återerinring från minnet

Alla frågor kräver att minnet aktiveras i någon grad men vissa frågor ställer högre krav på minnet än andra (läs mer i kapitel 11). En källa till fel kan vara att den sökta informationen kanske inte finns i respondentens minne. Den kanske aldrig har funnits där, har funnits där men vara svårtillgänglig eller skrivits över av annan information i minnet. Minnet kan ha rekonstruerats över tid till något som avviker från hur det faktiskt var. Respondenterna kanske inte minns exakt det som efterfrågas utan utgår från något närliggande. De kanske minns att något inträffat men placerar det fel i tiden. Sådana saker bidrar till att respondenten återhämtar bristfällig eller direkt felaktig information som underlag till sitt svar. Naturligtvis är detta steg beroende av steget innan. Om respondenten missförstått frågans fokus är riskerna stora att också fel information återerinras (Tourangeau, Rips & Rasinski, 2000; Willis, 2005).

Bedömning

I nästa steg ska respondenten bedöma och eventuellt bearbeta den återhämtade informationen till ett svar på frågan i frågeformuläret. Även i detta steg finns risker för fel. Det kan finnas luckor i minnet som respondenten fyller i med inkorrekt information. Respondenterna kanske inte bemödar sig med att söka efter fullständig information utan gör ett

överslag baserat på enskilda händelser, vilket kan bli missvisande. Ett annat exempel är att de kan göra beräkningsfel när de försöker få fram ett snitt (Tourangeau, Rips & Rasinski, 2000; Willis, 2005).

Lämnande av svar

I det avslutande steget ska respondenten välja det svarsalternativ som bäst sammanfaller med det individuellt formulerade svaret. Men respondenten kanske inte hittar något lämpligt svarsalternativ och väljer något på andra grunder. Respondenten vet kanske inte svaret men väljer ändå ett alternativ eftersom "Vet inte" saknas som svarsalternativ. På grund av bristande motivation kan icke-svar (som "Vet inte" och "Inte aktuellt") också väljas för att snabbt komma vidare i frågeformuläret. Respondenten kanske missförstår vad olika svarsalternativ står för eller justerar sitt svar till något mer socialt accepterat eller för att framstå som konsekvent (Tourangeau, Rips & Rasinski, 2000; Willis, 2005). Allt detta kan leda till fel.

Olika frågor ställer olika krav på svarsprocessen

Nedan följer några exempel på frågor med svårigheter i de olika stegen i svarsprocessen.

EXEMPEL:

Hur stor är den spatiala ytan?

Ovanstående fråga innehåller ett svårt ord. Detta kan leda till problem vad gäller förståelsen av frågan (steg 1 i fyrstegsmodellen). Som frågekonstruktör bör man helst testa om respondenterna förstår frågan. Kanske behöver den skrivas om, till exempel genom att använda enklare synonymer.

EXEMPEL:

Hur många gånger åt du kött förra månaden (dela upp dina svar på "till lunch", "till middag" och "övrigt")?

Utmaningen i frågan ovan ligger förmodligen inte i språket i sig utan i minnesprocesserna (steg 2 i fyrstegsmodellen). Minns respondenterna detta överhuvudtaget? Minns de det på den efterfrågade nivån? Vad får det för konsekvenser om respondenterna inte minns lika detaljerat som frågan gör anspråk på? Allt detta måste frågekonstruktören ta ställning till.

EXEMPEL:

Har du någon gång rökt marijuana?

Det svåra för respondenten i det tredje exemplet är förmodligen inte att förstå frågan eller att minnas om det hänt eller inte. Det svåraste steget är kanske snarare att avgöra om man ska svara "Ja" och därmed erkänna något känsligt (steg 3 i fyrstegsmodellen). Det beslutet kan påverkas av hur respondenten uppfattat intervjuaren, undersökningssituationen och vilken organisation som står bakom undersökningen.

I alla steg i den kognitiva modellen över svarsprocessen finns det saker som kan gå fel, vilket i slutänden kan leda till att respondenten lämnar ett felaktigt svar. Olika frågor kan ha sina främsta svårigheter vid olika steg i modellen.

Sammanfattande punkter:

- Respondenter kan missa någon information eller tolka centrala begrepp fel när de försöker förstå frågan.
- Respondenten kanske inte minns det som efterfrågas eller återhämtar felaktig eller bristfällig information.
- Respondenten kan göra felaktiga bedömningar eller uppskattningar.
- Respondentens svar kanske inte finns representerat bland svarsalternativen och respondenten väljer därför något annat.
- Kanske väljer respondenten "Vet inte" för att fort komma vidare.
- Olika frågor kan ha sina respektive svårigheter i olika steg i svarsprocessen. Vissa kan vara svåra att förstå, andra kan vara lätta att förstå men svåra att minnas svaren på och så vidare.

Företag och organisationer**En "ingång" på företaget - vem ska svara?**

I undersökningar till företag och organisationer finns samma felkällor som för undersökningar till individer och hushåll men dessutom några till som är organisatoriska till naturen. Det första steget är förstås vem på företaget som undersökningen ska skickas till. De flesta återkommande undersökningar med uppgiftslämnarskyldighet har kontaktpersoner på företagen. Men det är inte nödvändigtvis den personen som är bäst

lämpad att besvara frågeformuläret. Dessutom kan det ske rotation och nyrekrytering på företaget. Det är inte säkert att en tidigare kontaktperson fortfarande är aktuell. Därför kan det vara problem med att hitta en "ingång" på företaget.

Prioritering, samarbete och motivation

Kontaktpersonen måste sedan bedöma informationskravet i frågeformuläret, vad som krävs för att besvara det och vem eller vilka personer som är lämpliga att besvara det. Här finns det en risk att de mest optimala respondenterna inte väljs ut. Om det är flera respondenter inblandade krävs också någon form av kommunikation och samarbete. Det samarbetet kan fungera mer eller mindre bra, vilket kan påverka datakvaliteten. Företagets inställning till att besvara frågeformulär påverkar också prioritering och genomförande. Här är det troligtvis en skillnad mellan frivilliga undersökningar och undersökningar med uppgiftslämnarskyldighet. Den enskilde respondentens motivation påverkas säkert av företagets allmänna inställning, men sannolikt också av andra faktorer som rådande arbetssituation och personligt intresse. Om respondenten är mindre motiverad ökar risken för sämre svar (Willimack & Nichols, 2001; 2010).

Uppgifternas kompatibilitet

Förutom de organisatoriska faktorerna beskrivna ovan och de kognitiva faktorerna beskrivna i förra avsnittet, påverkas även datakvaliteten av informationssystemen och hur kompatibla de är med undersökningens frågor. Finns den efterfrågade informationen överhuvudtaget? Finns den i rätt enheter? Kan den räknas fram från andra uppgifter eller måste den uppskattas? Alla dessa faktorer kan ha inverkan på kvaliteten på svaren och kan variera mellan olika företag.

Sammanfattande punkter:

- I undersökningar till företag och organisationer finns ytterligare faktorer som kan leda till fel.
- Det är inte alltid givet vem på företaget som ska besvara frågeformuläret och det är inte säkert att den person som svarar är den mest lämpliga.
- Det kan vara flera personer inblandade i besvarandet som måste samarbeta.

- Informationssystemen och hur kompatibla de är med undersökningens frågor påverkar datakvaliteten.

Specifika typer av mätfel

Detta kapitel har handlat om svarsprocessen och felkällor. Många olika felkällor har nämnts men hur kan de ta sig uttryck i en specifik situation och hur vanliga är de?

En studie identifierade ett antal orsaker till mätfel. Studien klassificerade mätfel efter orsak hos sammanlagt 206 utvalda variabler, fördelade på 135 produkter vid SCB. Nedan följer de vanligaste feltyperna i rangordning. Listan är hämtad från SCB:s publikation *Fråga rätt med vissa justeringar* (Christianson & Polfeldt, 1996; *Fråga rätt*, 2001):

- A. Definitionsproblem. Statistikens definitioner var obekanta eller användes inte hos uppgiftslämnaren.
- B. Minnesfel. Var vanligast i individ- och hushållsundersökningar men förekom också i företagsundersökningar.
- C. Uppgiftslämnaren måste göra bedömningar. Exakta uppgifter fanns inte alls eller inte för rätt period.
- D. Periodiseringsproblem. Uppgifterna avsåg en annan referensperiod eller -tidpunkt än den efterfrågade, t.ex. då ett företag har brutet räkenskapsår.
- E. Krav på beräkning. Variabeln fanns inte, utan svaret måste kalkyleras från andra uppgifter.
- F. Korrelerad variabel användes. T.ex. svarade man med uppgift om leveranser, då produktionen efterfrågades men inte var tillgänglig.
- G. Redovisningsproblem. Företaget hade inte sina uppgifter uppdelade på de kategorier eller objektstyper som statistiken efterfrågade.
- H. Klassificeringsfel i bakgrundsvariabler p.g.a. att de var svåra att mäta.
- I. Sant värde saknades.
- J. Felaktig inkludering av komponenter i en summavariabel. Uppgiftslämnaren räknar felaktigt in komponenter som inte efterfrågas.
- K. Fel vid registreringen i blanketten som inte upptäcks.
- L. Felaktig exkludering av komponenter i en summavariabel. Uppgiftslämnaren glömmer att räkna in en eller flera komponenter som efterfrågades.

Speciellt i ekonomiska undersökningar förekom också följande fel:

- Sortfel då respondenten inte observerar i vilken enhet svaret ska ges, till exempel rapporteras i kronor istället för i tusentals kronor.
- Respondenten kan inte få fram den efterfrågade kvantitativa uppgiften och väljer att lämna svarsrutan blank, markera med ett streck eller kanske skriva 0.
- Felaktig exkludering av en komponent i en variabel medför ibland felaktig inkludering av komponenten i en annan – ofta en övrig- eller restpost.

Listan ovan (Christianson & Polfeldt, 1996; Fråga rätt, 2001) visar olika aspekter av frågan, svarsalternativen, respondenten och respondentens informationssystem som frågekonstruktören bör ha i åtanke när frågorna konstrueras. Feltyperna kan placeras in i modellerna som redovisades tidigare i detta kapitel. De diskuteras även djupare och mer utförligt i olika kapitel i denna bok.

Sammanfattning

Svarsprocessen och felkällor – sammanfattning

Forskning om frågekonstruktion – historik och teoriutveckling:

- Det har länge varit känt att frågornas utformning påverkar svaren. Forskning genomfördes redan på 1930- och 1940-talen.
- På 1970-talet utvidgades det kognitiva perspektivet med tankeprocesser inom psykologin.
- Kognitiva psykologer och surveyforskare började samverka. Idag är kognitionspsykologin given inom surveyområdet.
- Samverkan resulterade bland annat i ett antal modeller som syftar till att beskriva mätsituationen eller svarsprocessen (från det att respondenten får en fråga tills ett svar har lämnats) i en undersökning.

Kognition och kommunikation i svarsprocessen:

- Det finns flera modeller för mätsituationen i en undersökning. De inkluderar till exempel respondenten, intervjuaren, frågeformuläret och de yttre ramarna (som tid och plats).

- För frågekonstruktion behövs dock en modell som mer specifikt fokuserar på hur respondenten förhåller sig till själva frågorna, dvs. modeller som fokuserar på svarsprocessen.
- Den mest spridda modellen över svarsprocessen är den så kallade fyrstegsmodellen:
1) förståelse av frågan, 2) återerinring från minnet, 3) bedömning, 4) lämnande av svar.
- En annan spridd modell – satisficingmodellen – belyser hur respondenten kan sträva efter att antingen finna ett optimalt svar eller ta genvägar och nöja sig med ett svar som är tillräckligt bra.

Svarsprocessen i företag och organisationer:

- Svarsprocessen i företag och organisationer skiljer sig från den för individer och hushåll, bland annat för att flera respondenter kan vara inblandade och för att svaren baseras på information från datasystem snarare än individers minne.
- Det finns mindre forskning om svarsprocessen i företag och organisationer.
- Hybridmodellen innehåller både kognitiva och organisatoriska steg (därav namnet).
- Företaget i sig kan ses som ett kognitivt system (socialt distribuerad kognition).
- Ett annat perspektiv är att fokusera på hur kompatibel företagens data är med frågan (finns till exempel exakta siffror eller måste de uppskattas?).

Felkällor och mätfel:

- Mätfel är när det observerade värdet avviker från det sanna värdet.
- Ibland finns inget sant värde i strikt betydelse. Då är utgångspunkten det värde vi skulle ha fått under optimala förutsättningar.
- Mätfel kan grunda sig i intervjuaren, respondenten, frågeformuläret eller datainsamlingsmetoden.
- Respondenter kan missa någon information eller tolka centrala begrepp fel när de försöker förstå frågan.
- Respondenten kanske inte minns det som efterfrågas eller återhämtar felaktig eller bristfällig information.

- Respondenten kan göra felaktiga bedömningar eller uppskattningar.
- Respondentens svar kanske inte finns representerat bland svarsalternativen och respondenten väljer därför något annat.
- Kanske väljer respondenten "Vet inte" för att fort komma vidare.
- Olika frågor kan ha sina respektive svårigheter i olika steg i svarsprocessen. Vissa kan vara svåra att förstå, andra kan vara lätta att förstå men svåra att minnas svaren på och så vidare.
- I undersökningar till företag och organisationer finns ytterligare faktorer som kan leda till fel.
- Det är inte alltid givet vem på företaget som ska besvara frågeformuläret och det är inte säkert att den person som svarar är den mest lämpliga.
- Det kan vara flera personer inblandade i besvarandet som måste samarbeta.
- Informationssystemen och hur kompatibla de är med undersökningens frågor påverkar datakvaliteten.

10 Att fråga efter kunskap

Författare: Anna Sundström (Umeå universitet)

Följande avsnitt handlar om konstruktion av kunskapsfrågor i samband med undersökningar. Kapitel 2 handlade om frågeteknik och tog upp vad man ska tänka på i samband med frågekonstruktion. Kapitel 3 behandlar vad man ska tänka på när man formulerar svarsalternativ. I detta kapitel ligger fokus specifikt på konstruktion av frågor som syftar till att mäta kunskaper. Ett vanligt frågeformat för kunskapsfrågor är flervalsfrågor med fasta svarsalternativ, det vill säga en frågeställning som följs av ett antal svarsalternativ, därför inriktas detta avsnitt på konstruktion av denna typ av frågor. I avsnittet presenteras tips för konstruktion av kunskapsfrågor och exempel på fallgropar.

Kunskapsfrågor i enkätundersökningar

Även om det inte är så vanligt förekommande, händer det ibland att man är intresserad av att mäta respondenternas kunskaper inom ett visst ämnesområde i en enkätundersökning. Exempelvis förekommer frågor i enkätundersökningar som syftar till att mäta politisk kunskap. Den svenska valundersökningen, som genomförs regelbundet i samband med riksdagsvalet, är ett exempel på en undersökning som bland annat mäter svenska folkets politiska kunskaper (Oscarsson & Holmberg, 2008). Andra exempel är undersökningar som syftar till att mäta svenska folkets kunskaper, attityder och värderingar om demokrati och mänskliga rättigheter (Ekman & Wohlgemuth, 2008). Ytterligare exempel är undersökningar om demokratikunskaper hos svenska gymnasieelever (Skolverket, 2003). Kunskapsfrågor förekommer också i attitydundersökningar för att hjälpa oss att förklara attityder och åsikter. De används även för att undersöka om personerna som svarar har tillräckligt mycket kunskap i ämnet för att det ska vara meningsfullt att fråga om deras attityder (Bradburn, Sudman, & Wansink, 2004).

Kunskapsfrågor kan också användas i enkäter för att utforma och genomföra informationskampanjer, till exempel om trafiksäkerhet eller hälsa. Innan en effektiv informationskampanj kan genomföras behövs information om populationens kunskapsnivå. Ett sådant exempel är en enkätundersökning som genomfördes för att studera allmänhetens

kunskaper om trafiksäkerhet. Tanken var att en sådan enkät skulle kunna användas för att studera kunskaper i olika trafikantgrupper över tid samt undersöka om det förelåg kunskapsbrister inom något särskilt område (Sundström, Lyrén & Alger, 2010). Kunskapsfrågor används också i enkäter för att mäta vissa gruppers kunskaper eller förmåga inom ett särskilt område eller för att utvärdera effekterna av en viss utbildning. Inom vården finns flera exempel på enkätundersökningar som genomförts för att studera vårdpersonalens kunskaper inom ett särskilt område, exempelvis kunskaper om diabetes hos vårdpersonal inom hemtjänsten (Danielsson & Levin, 2009).

Även om det kan vara värdefullt att mäta kunskaper inom ramen för en enkätundersökning finns det också flera potentiella svårigheter och problem som bör nämnas. Kunskapsfrågor som en del i en attitydundersökning kan till exempel leda till att respondenterna känner sig avslöjade om de har bristfällig kunskap på området. Beroende på vilken typ av undersökning det rör sig om kan kunskapsfrågor också uppfattas som känsliga och därigenom påverka svarsfrekvensen negativt. När man ska genomföra en undersökning bör man av denna anledning ställa sig frågan hur värdefull informationen om respondenternas kunskaper är, och väga fördelarna med att samla in sådan information mot de potentiella nackdelarna och vilka risker det kan innebära att inkludera sådana frågor i en undersökning. När man beslutat sig för att inkludera kunskapsfrågor bör man ta ställning till vilket frågeformat som är lämpligast. Ett vanligt förekommande frågeformat är flervalfrågor med svarsalternativ. Ett annat tänkbart frågeformat inriktas på huruvida respondenten känner till de kunskaper som efterfrågas, till exempel genom att fråga: "Känner du till om det finns ...?" eller "Har du hört talas om ...?". I valet av frågeformat är risken för "fusk" något som bör beaktas (Bergman & Wärneryd, 1982).

Att mäta olika typer av kunskap

Kunskapsfrågor kan vara inriktade på att mäta olika typer av kunskap. En modell som kan användas för att kategorisera typer av kunskap är Blooms reviderade taxonomi. Taxonomin är vanligt förekommande i utbildningssammanhang och beskriver kognitiva nivåer och kunskapsnivåer (Anderson et al., 2001). Den kognitiva dimensionen innehåller sex olika kategorier: minnas, förstå, tillämpa, analysera,

värdera och skapa. Dessa kategorier kan specificeras ytterligare i olika aspekter (se Figur 10.1).

Kunskapsdimensionen beskriver fyra kunskapskategorier. Den första kategorin, *faktakunskap*, innehåller kunskap om termer och specifika detaljer och delar. *Begreppskunskap*, den andra kategorin, är kunskap om samband mellan de grundläggande delarna i en större struktur som möjliggör att de fungerar tillsammans. Den tredje kategorin är *procedurkunskap* och där ingår kunskap om hur man gör något. Hit räknas undersökningsmetoder och kriterier för att använda färdigheter, tekniker och metoder. Den fjärde kategorin, *metakognitiv kunskap*, innefattar kunskaper om generell uppfattningsförmåga, men även medvetenhet och kunskap om sin egen uppfattningsförmåga. Alla kategorier har undergrupper som förtydligar och preciserar vad som ingår i respektive kategori (se Figur 1).

Taxonomi ordnar de sex kognitiva processerna hierarkiskt efter grad av komplexitet. *Minnas* anses vara den minst komplexa processen medan *skapa* är den mest komplexa. I kunskapsdimensionen sorteras kategorierna från konkret till abstrakt kunskap där *faktakunskap* räknas som konkret kunskap och *metakognitiv kunskap* som abstrakt kunskap (se Figur 10.1).

Figur 10.1 Blooms reviderade taxonomi

BLOOMS REVIDERADE TAXONOMI		KOGNITIVA PROCESSER					
		Minnas	Förstå	Tillämpa	Analysera	Värdera	Skapa
		Känna igen Komma ihåg	Tolka Exemplifiera Klassificera Sammanfatta Dra slutsatser Jämföra Förklara	Verkställa Applicera	Särskilja Organisera Tillskriva	Kontrollera Kritisera	Generera Planera Producera
KUNSKAPSDIMENSIONEN							
Fakta- kunskap	Terminologi						
	Specifika detaljer och delar						
Begrepps- kunskap	Klassificeringar och kategorier						
	Principer och generaliseringar						
	Teorier, modeller och strukturer						
Procedur- kunskap	Ämnesspecifika färdigheter och algoritmer						
	Ämnesspecifika tekniker och metoder						
	Kriterier för att avgöra när man ska använda lämpliga metoder						
Meta- kognitiv kunskap	Strategisk kunskap						
	Kunskap om inlärningsfrågor inkluderat lämplig kontextuell och villkorlig kunskap						
	Kunskap om sig själv						

Konstruktion av kunskapsfrågor

Det finns olika typer av frågeformat. Som nämnts i föregående kapitel brukar man skilja på öppna frågor och frågor med fasta svarsalternativ. Frågor med fasta svarsalternativ är vanliga och mycket användbara vid kunskapsmätningar, därför ligger fokus i detta avsnitt på frågor med det formatet.

En flervalsfråga består av en frågestam och ett antal svarsalternativ. Frågestammen anger problemet. Ibland presenteras också information som är nödvändig för att sätta in frågan i sitt sammanhang. Listan med svarsalternativ innehåller ett korrekt alternativ och ett antal felaktiga alternativ (så kallade distraktorer). Distraktorerna ska vara möjliga svar för den som inte har den kunskap som frågan mäter. För den som har kunskapen ska distraktorerna uppfattas som felaktiga alternativ, det vill säga bara det korrekta svaret ska vara rätt.

Flervalsfrågor kan användas för att mäta kunskaper på olika kognitiva nivåer. Frågor kan exempelvis utformas för att analysera fenomen, tillämpa principer på nya situationer, förstå idéer och koncept, tolka diagram och tabeller, dra slutsatser från data och lösa problem. En frågas svårighetsgrad kan justeras med hjälp av de svarsalternativ som erbjuds. Mer homogena alternativ ger en svårare fråga, medan mer heterogena alternativ ger en lättare fråga.

Det finns både för- och nackdelar med användningen av flervalsfrågor. En fördel är att flervalsfrågor är effektiva och genererar mer tillförlitlig information än öppna frågor (Downing, 2006). Generellt kan man säga att svarstiden för flervalsfrågor är kortare jämfört med öppna frågor. Detta innebär att man på samma tid kan ställa fler frågor och därigenom täcka ett större ämnesområde om man använder flervalsfrågor jämfört med om man använder öppna frågor. Flervalsfrågor kan med lätthet administreras till en stor grupp respondenter, eftersom de kan rättas effektivt och objektivt. Den objektiva poängsättningen innebär en fördel eftersom poängsättningen inte påverkas av vem som rättar frågan. Ännu en fördel är möjligheten att använda tekniker med automatisk rättning då instrumentet administreras via dator. En ytterligare fördel med flervalsfrågor är att man genom statistiska analyser kan undersöka frågornas kvalitet och förbättra frågorna grundat på detta.

Flervalsfrågor bygger på att respondenten väljer bland ett antal svarsalternativ. Av den anledningen är denna typ av frågor inte lämpliga när det krävs att respondenten producerar ett svar, till exempel uttrycker förklaringar, idéer eller ger exempel på något. En av de största begränsningarna med flervalsuppgifter är att de fungerar dåligt om de konstruerats fel. Det är svårare och mer tidskrävande att utforma bra flervalsfrågor än öppna frågor. Det beror på att man både behöver konstruera en välfungerande fråga samt lämpliga svarsalternativ.

För att utveckla så bra och välfungerande frågor som möjligt finns etablerade principer för frågekonstruktion som bör följas när man utvecklar frågor (Downing, 2006; Frary, 1995; Haladyna, Downing & Rodriguez, 2002; Moreno, Martinez & Muniz, 2006). Med utgångspunkt i dessa principer presenteras här riktlinjer för konstruktion av frågor och svarsalternativ. Riktlinjerna kan delas in i fyra områden: riktlinjer angående hela mätinstrumentets (enkätens) syfte och innehåll, frågornas innehåll i relation till det område som ska mätas, frågornas utformning samt svarsalternativens utformning.

Mätningens syfte och innehåll

Den första fråga man måste ställa sig är vad som är syftet med mätningen och vilket område som ska mätas. För att öka validiteten i mätningen är det viktigt att området som ska mätas definieras så tydligt som möjligt. Om området som ska mätas är otydligt beskrivet finns risken att man konstruerar frågor som inte är relevanta för syftet med mätningen. Litteraturoversikter och expertbedömningar är vanliga metoder för att beskriva och definiera det område som ska mätas (läs även mer i kapitel 2).

Det är också viktigt att specificera i vilket sammanhang frågorna ska användas och vilken population som ska besvara frågorna. När man formulerar frågorna är det viktigt att man tar hänsyn till populationens egenskaper så att frågorna blir begripliga. Exempel på viktiga aspekter att ta hänsyn till är gruppens ålder, utbildningsnivå, modersmål och eventuella fysiska eller psykiska funktionshinder. Andra viktiga faktorer att ta hänsyn till är hur och var frågorna kommer att administreras. Om man inte tar hänsyn till dessa faktorer är risken stor att innehållet, språket eller formatet i frågorna blir olämpligt vilken kan leda till mätfel.

Frågornas innehåll

Den andra stora frågan som är avgörande för mätningens kvalitet är att frågorna är relevanta och mäter det innehåll man har tänkt mäta. Syftet med mätningen, innehållsområdet och sammanhanget där mätinstrumentet ska användas ska styra frågekonstruktionen. Varje fråga ska täcka en del av innehållsområdet och bilda en helhet tillsammans med de andra frågorna. Vilken kognitiv nivå och vilken kunskapsnivå som är lämplig för en fråga beror på innehållsområdet som ska mätas. Ibland kan det vara en specifik och enkel fråga, som att minnas datumet för en historisk händelse, i andra fall kan det vara en fråga som kräver en lösning på ett komplext problem.

Formulering av frågor och svarsalternativ

När det gäller själva formuleringen av frågor och svarsalternativ finns ett antal riktlinjer att följa för att frågorna ska fungera så bra som möjligt. Dessa riktlinjer beskrivs i avsnittet nedan tillsammans med ett antal exempel på hur det kan gå om dessa riktlinjer inte följs.

Frågornas utformning

A. *Frågorna ska vara klart och tydligt formulerade. Man ska inte behöva läsa svarsalternativen för att förstå frågan.*

I exemplet nedan saknas en tydlig frågeformulering.

EXEMPEL:**I ett hierarkiskt system...**

- a) ... fattar ledaren beslut ensam
- b) ... fattas gemensamma beslut
- c) ... finns inga formella beslut
- d) ... görs olika beslut på olika hierarkiska nivåer

B. *Man ska undvika frågor som är förvirrande, tvetydiga, för långa eller för kortfattade. Det finns inga universella regler för vad som är en klart och tydligt formulerad fråga. Det beror på vilka som ska besvara frågan och i vilket sammanhang frågan ska användas.*

C. *Frågorna ska vara entydiga – bara fråga efter en sak.*

D. *Man ska ge klara instruktioner om hur frågan ska besvaras.*

E. *Frågeställningen ska vara så kort som möjligt. Den ska inte innehålla överflödigt information.*

I följande fråga presenteras mycket information i frågestammen som är överflödigt.

EXEMPEL:

Depression är en psykisk störning som ingår i de affektiva störningarna. Behandling av depression kan vara inriktad dels på terapi och dels på medicinering. De nyaste depressionsmedicinerna (så kallade SSRI) påverkar:

- a) serotoninets återupptagning
- b) adrenalinutsöndring
- c) acetylkolinets återupptagning
- d) serotoninutsöndringen

F. Undvik triviala frågor.

I en undersökning till sjuksköterskestudenter och tandhygienister ställdes följande fråga, vilken kan tyckas ganska trivial för denna grupp.

EXEMPEL:

Vilken färg har ett friskt tandkött?

- a) Mörkröd
- b) Röd
- c) Ljusrosa

G. Frågestammen ska inte innehålla information i form av ledtrådar, så att respondenten direkt kan svara på uppgiften.

I nedanstående exempel ger ordet sömnig i frågestammen en ledtråd till det korrekta svaret.

EXEMPEL:

Vilket har visat sig vara effektivast om man blir sömnig under bilkörning?

- a) Stanna och sova en stund.
- b) Stanna och dricka kaffe.
- c) Stanna och ta en kort promenad.

H. Språket ska vara lätt att förstå för de grupper som ska besvara frågan. Man ska inte använda slang, fackuttryck eller svåra ord.

I en undersökning till vårdpersonal användes följande fråga. Det korrekta svarsalternativet innehåller begreppet *skjuvkraft*, en fackterm som kan vara onödigt svår för de svarande.

EXEMPEL:

Vad händer när en patient som sitter i sängen i en halvt upprätt position (60°) glider ner?

- a) Trycket ökar när huden fastnar vid underlaget.
- b) Friktionen ökar när huden fastnar vid underlaget.
- c) Skjuvkrafterna ökar när huden fastnar vid underlaget.

I. Frågorna ska vara oberoende av varandra. En fråga ska inte ge en ledtråd till svaret på en annan fråga.

I samma undersökning som beskrivits ovan i exempel H användes följande fråga, där definitionen av begreppet "skjuv" förekommer som ett av alternativen. Detta är olämpligt eftersom det innebär att frågorna är beroende av varandra.

EXEMPEL:

Vilket påstående är korrekt?

- a) Tvål kan torka ut huden och därför öka risken för trycksår.
- b) Fukt från urin, feaces eller sårsekretion orsakar trycksår.
- c) Skjuv är den kraft som inträffar när kroppen glider ner och huden fastnar vid underlaget.

Svarsalternativens utformning

För svarsalternativens utformning kan följande riktlinjer användas:

A. Utveckla så många bra svarsförslag som möjligt

Mellan tre och fem svarsalternativ är ofta tillräckligt. Ju fler svarsalternativ frågan har, desto mindre är chansen att gissa rätt. Problemet ligger dock i att lyckas konstruera flera välfungerande svarsalternativ. För att frågan ska fungera bra ska alla felaktiga svarsalternativ (distraktorer) vara rimliga och attraktiva för den som saknar kunskap, har felaktig kunskap eller bara gör en ytlig analys av frågan.

En fråga som bara har två svarsalternativ ger 50 procenters chans att gissa rätt.

EXEMPEL:

Vilken tobaksprodukt (ej light) innehåller mest nikotin?

- a) En cigarett
- b) En portionssnus

B. Endast ett alternativ ska vara korrekt.

I exemplet nedan kan både alternativ C och D vara korrekta.

EXEMPEL:**Vilket är det huvudsakliga motivet till att vi får allt fler cirkulationsplatser (rondeller) i den svenska vägmiljön?**

- a) Att underlätta så att trafiken flyter bättre.
- b) Att skapa en vackrare trafikmiljö.
- c) Att minska antalet olyckor.
- d) Att minska antalet olyckor med personskada.
- e) Att minska koldioxidutsläpp.

C. Svartalternativen ska presenteras i en logisk ordning.

Respondenten ska inte behöva ordna svartalternativen för att kunna besvara frågan. Ordna svartalternativen logiskt, till exempel fallande/stigande eller i alfabetisk ordning. Årtal ordnas kronologiskt.

D. Placeringen av rätt svartalternativ ska varieras mellan frågorna. Sträva efter en jämn fördelning av korrekta svar.

E. Svartalternativen ska vara oberoende. Alternativen ska inte överlappa eller hänvisa till varandra.

I exemplet nedan är svartalternativen beroende av varandra, eftersom intervallerna överlappar varandra.

EXEMPEL:**Hur många giftiga ämnen innehåller tobaksrök?**

- a) 900–3 000
- b) 2 800–4 000
- c) 4 000–6 000

I utformningen av svartalternativ ska man undvika svartalternativ som har samma betydelse, det vill säga alla svartalternativ ska vara unika. Undvik också formuleringar som "Alla av de ovanstående" eller "Ingen av de ovanstående".

F. Alla svarsförslag ska vara grammatiskt överensstämmande med frågan. De ska vara enkla och entydiga.

EXEMPEL:

Vad av följande är det som avgör när på året en bil ska besiktigas?

- a) Slutsiffran i registreringsnumret
- b) Tillverkningsdatum
- c) Inköpsdatum

G. Det korrekta svarsalternativet ska inte skilja ut sig från övriga alternativ så att man kan ana att detta alternativ är rätt.

Svarsförslagen ska vara ungefär lika långa. Det korrekta alternativet ska inte skilja ut sig från övriga genom att vara mycket längre eller kortare, eller på andra sätt avvika från de andra alternativen. I exemplet nedan är det korrekta svarsalternativet betydligt längre än distraktorerna.

EXEMPEL:

Patienter med nedsatt näringstillstånd löper större risk för att utveckla trycksår eftersom:

- a) Undernäring ofta uppträder tillsammans med andra faktorer, såsom begränsad aktivitet och spontana rörelser.
- b) Proteinbrist stimulerar utveckling av trycksår.
- c) Vitamin- och zinkbrist stimulerar utveckling av trycksår.

H. Formulera frågan och svarsalternativen positivt, undvik negationer.

EXEMPEL:

Vad av följande bör man inte göra om man vill släcka en eld i en panna på spisen:

- a) Lämna pannan där den står.
- b) Hälla vatten i pannan.
- c) Lägga på ett lock.
- d) Stänga av spisen.

I. Varje svarsalternativ ska vara det kortast möjliga svaret.

Information från frågestammen ska inte upprepas i svarsförslaget, då blir det onödigt långt.

EXEMPEL:**Varför finns det mer än ett politiskt parti i demokratiska länder?**

- a) Partierna begränsar politisk korruption.
- b) Partierna förhindrar politiska demonstrationer.
- c) Partierna främjar ekonomisk konkurrens.

J. Alla distraktorer ska vara möjliga för den som inte känner till det rätta svaret.

Distraktorer kan konstrueras utifrån vanliga missförstånd och vanliga felsvar. Distraktorer kan också vara konceptuella, det vill säga svar som ligger nära det korrekta svarsalternativet och som är trovärdigt för dem som inte kan det rätta svaret.

I exemplet nedan kan några av distraktorerna uteslutas utan särskild kunskap om ämnesområdet.

EXEMPEL:**Miljöterapi...**

- a) ... är en form av inredningsplanering.
- b) ... är en form för behandling av stora grupper och institutioner.
- c) ... bedrivs av ekologiskt medvetna smågrupper.
- d) ... bedrivs i särskilda miljöer.

K. Frågorna ska inte vara kuggfrågor.

Frågorna ska inte formuleras så att respondenter med kunskap kan "luras" att välja ett felaktigt svar, som i exemplet nedan.

EXEMPEL:**Vad är bäst för miljön?**

- a) Att använda farthållare så ofta du kan.
- b) Att alltid växla upp i uppförsbacke
- c) Att alltid växla ner i uppförsbacke.
- d) Att alltid växla ner i nedförsbacke.

L. Undvik ledtrådar till det korrekta svaret

Använd inte oklara frekvenser som *ibland/ofta* eller *alltid/aldrig*; eftersom de kan ge ledtrådar till det rätta svaret. Ord som *ibland*, *kanske*, *vanligtvis*,

generellt brukar förknippas med korrekta svar, medan ord som *aldrig*, *alltid*, *alla*, *endast* brukar förknippas med felaktiga svar. Undvik också att använda samma ordval i frågan som i det korrekta svarsförslaget. Man ska också se upp med skillnader i grammatik mellan svarsalternativen som ger ledtrådar till det rätta svaret, till exempel att bara ha ett svarsalternativ som passar med frågestammen. Andra frågor ska inte heller ge ledtrådar till det korrekta svaret i frågan.

M. Layouten ska vara tydlig.

Man ska undvika för litet typsnitt eller att lämna för lite luft mellan svarsalternativen. Man brukar rekommendera att de presenteras vertikalt.

N. "Vet inte"-alternativ i kunskapsfrågor

I undersökningar som innehåller kunskapsfrågor finns ibland svarsalternativet "Vet inte" för att motverka chansning. Det kan även finnas med för att respondenten inte ska uppleva frågorna som pressande om man inte kan det rätta svaret. Huruvida detta är en bra strategi för frågekonstruktion har dock diskuterats.

I en enkätstudie om tobak använde man sig av följande fråga:

EXEMPEL:

Vilken är den beroendeframkallande substansen i tobak?

- a) Kolmonoxid
- b) Tjära
- c) Nitrosaminer
- d) Nikotin
- e) Vet inte

Fördelen med ett "Vet inte"-alternativ är att det är möjligt att få information om andelen respondenter som har den kunskap som efterfrågas, hur många som har felaktig kunskap och hur många som inte har kunskap. Nackdelen med denna typ av svarsformat är dock att validiteten kan hotas då respondenterna utnyttjar "Vet inte"-alternativet i olika grad. En risk är att vissa respondenter är omotiverade och väljer detta svarsalternativ istället för att försöka besvara frågan (Fink, 2003). En annan nackdel är att den uppskattade kunskapsnivån blir lägre, eftersom

”Vet inte”-svar kan inrymma partiell kunskap. Denna kategori av svar varierar som en funktion av personlighetsegenskaper som självförtroende, eftersom vissa personer är mer benägna att chansa än andra (Mondak, 2000, 2001; Sturgis, Allum & Smith, 2007). Vissa hävdar dock att det är tveksamt om ”Vet inte”-svaren faktiskt innehåller partiell kunskap. En studie visar att andelen korrekta svar bland dem som tidigare svarat ”Vet inte” inte var bättre än slumpen när de ombads chansa på rätt svar. Det tyder på att ”Vet inte”-alternativen inte verkar inrymma partiell kunskap (Sturgis et al., 2007).

Eftersom det finns argument som talar både för och emot användandet av ”Vet inte”-svar får man ta hänsyn till syftet med undersökningen och vad resultaten ska användas till när man bestämmer sig för om ”Vet inte”-alternativ ska inkluderas eller ej, och välja det man tror ger de mest rättvisande resultaten.

Sammansättning av enkäten

När frågorna har konstruerats ska man dels försäkra sig om att de mäter det område som man avser att mäta. Man ska dels även försäkra sig om att frågorna representerar olika delområden i lämplig utsträckning med avseende på antalet frågor. Frågorna i enkäten kan ordnas efter innehåll eller blandas, beroende på vad som passar bäst för mätningens syfte.

Administrering

För kunskapsfrågor jämfört med andra typer av enkätfrågor finns det några problematiska aspekter när det gäller självadministrering. Tillförlitligheten i svaren påverkas negativt av att respondenterna har möjlighet att kolla upp svaren eller diskutera frågorna med andra. Användningen av kunskapsfrågor i självadministrerade enkäter ställer därför krav på att administreringen sker under kontrollerade förhållanden. Den kan vara datoriserad med övervakning eller att frågorna formuleras så att det är svårt att kolla upp det rätta svaret. En annan möjlig lösning är att ha en datorbaserad enkät med tidsbegräsning för varje fråga.

Sammanfattning

Ibland förekommer det att man vill inkludera kunskapsfrågor i enkätundersökningar, exempelvis för att mäta politisk kunskap eller för att genomföra och mäta effekterna av informationskampanjer. Innan man

inkluderar kunskapsfrågor i en enkätundersökning bör man väga fördelarna mot nackdelarna. Som stöd i utvecklandet av kunskapsfrågor kan Blooms reviderade taxonomi användas för att täcka olika typer av kunskaper och olika kognitiva processer. Ett vanligt frågeformat när man använder kunskapsfrågor är så kallade flervalfrågor som består av en frågeställning och ett antal svarsalternativ. Vid konstruktion av flervalfrågor finns riktlinjer för formuleringen av såväl frågor som svarsalternativen. Dessa riktlinjer behandlas i kapitlet, och exempel på vanliga fallgropar tas upp.

Att fråga efter kunskap – sammanfattning

Kunskapsfrågor i enkäter:

- Ibland förekommer det att man i enkäter vill fråga efter respondenternas kunskaper.
- Kunskapsfrågor kan vara inriktade på att mäta olika typer av kunskap.
- Blooms reviderade taxonomi – modell för att beskriva kognitiva nivåer och kunskapsnivåer.
- Frågor med fasta svarsalternativ är vanliga och mycket användbara vid kunskapsmätningar.

Frågor med fasta svarsalternativ:

Frågor med fasta svarsalternativ har många fördelar

- effektiva och genererar tillförlitlig information
- lätta att administrera, rättas effektivt och objektivt
- automatisk poängsättning möjligt vid webbenkäter
- möjliggör statistiska analyser av frågornas kvalitet.

Frågorna fungerar dåligt om de konstruerats fel – viktigt att följa etablerade principer för frågekonstruktion:

1. Vad är syftet med mätningen och vilken grupp riktar sig mätningen till?
2. Att frågorna täcker av det innehållet som ska mätas.
3. Att följa riktlinjer för formulering av frågor och svarsalternativ.

11 Att fråga om det förflutna

Författare: Andreas Persson (SCB)

Alla surveyfrågor är beroende av minnet i någon grad. Även en så pass enkel fråga som "När är du född?" kräver att minnet aktiveras i någon grad för att kunna besvaras. Men vissa frågor kräver mer av minnet än andra. Mest krävande är frågor om det förflutna som begär att respondenten söker djupare i minnet. Det kan till exempel vara *när* eller *hur många gånger* något inträffade under en period.

Detta kapitel inleds med en kort genomgång av hur centrala delar av minnet fungerar, sedan följer ett allmänt avsnitt om på vilka sätt minnet kan leda respondenten fel när frågor ska besvaras. Det allmänna avsnittet följs av två mer specifika avsnitt om hur det går till när respondenter ska besvara *när i tiden* respektive *hur många gånger* något hände. Kapitlet avslutas med ett mer praktiskt inriktad avsnitt om hur minnet kan främjas i undersökningar.

Hur minnet fungerar

Vi tar ofta minnet för givet men om man reflekterar över det blir det tydligt vilken fantastisk egenskap minnet är. Tänk hur det skulle vara om vi inte kom ihåg någonting. Samtidigt vet vi att vi inte minns allt utan glömmer en del. Så hur fungerar minnet egentligen?

Långtidsminne och arbetsminne

Det finns mycket litteratur om hur minnet fungerar. Det är ett komplext ämne. Nedanstående är en förenklad beskrivning av minnet men syftar till att ge tillräcklig insikt för att kunna förstå hur det påverkar formulärfrågor om det förflutna. Det finns flera perspektiv men det är vanligt att böcker delar upp minnet i två centrala funktioner: långtidsminne och arbetsminne (kallas även korttidsminne) (Passer & Smith, 2007; Ashcraft 1994; Langlet, 1990).

Arbetsminnet kallas det minne där vi har information som vi för tillfället uppmärksammar eller arbetar med. Vissa forskare liknar det vid ett "mentalt skrivbord". Det är ett begränsat minnessystem både i tid och i antal enheter. Information kan behållas i arbetsminnet genom

upprepning (att vi till exempel upprepar ett telefonnummer tills vi har möjlighet att ringa) eller genom att vi arbetar med information (elaborerar eller associerar). Arbetsminnet och dess begränsningar är relevant för frågekonstruktion och då framförallt för frågor i intervjuundersökningar där respondenten inte kan läsa frågorna själv och läsa om vid behov. För frågor om det förflutna är dock långtidsminnet det relevanta minnessystemet (Passer & Smith, 2007; Ashcraft 1994; Langlet, 1990).

Information från arbetsminnet kan föras över till långtidsminnet genom en process som kallas inkodning. Långtidsminnet har inte arbetsminnets begränsningar vad gäller tid och enhet utan kan mer liknas vid ett arkiv eller nätverk av lagrad information. Det som händer vid inkodning är att informationen får en plats i nätverket och associationsvägar till annan information. Vid frågor om det förflutna är det den delen av långtidsminnet som kallas det episodiska minnet (våra personliga erfarenheter) som är relevant. Inkodningen kan ske automatiskt eller genom ansträngning (att vi försöker lära oss något). Det sistnämnda ger en djupare inkodning eftersom informationen då i högre grad ges mening, organisation och länkas till annan relevant information i minnet (läs mer på andra platser i detta kapitel) (Passer & Smith, 2007; Ashcraft 1994; Langlet, 1990).

Sammanfattande punkter:

- Det finns olika teorier om hur minnet fungerar. En vanlig distinktion är att skilja på långtidsminne och arbetsminne.
- Långtidsminnet är ett minnessystem som innehåller lagrade minnen. Det kan ses som ett slags arkiv.
- Arbetsminnet är däremot ett betydligt mer begränsat minnessystem, både vad gäller tid och enheter, där vi temporärt kan behålla information som vi behöver för tillfället.
- Informationen behöver koda in i långtidsminnet (episodiskt minne kallas den del som är relevant vid frågor om det förflutna).

Om glömska

När vi frågar om det förflutna krävs det att respondenten söker i det episodiska långtidsminnet. Ett bekymmer kan vara att respondenten

glömt det vi frågar efter. Hur går det till när vi glömmer något? Enligt äldre teori beror detta på att minnen försvagas med tiden och till slut försvinner. Denna teori har dock fått mycket kritik, bland annat för att vissa minnen är väldigt långlivade och inte verkar försvagas. Dessutom är det problematiskt att se själva tidsåtgången i sig som orsaken. Snarare är det nog så att tiden ger utrymme till andra processer att påverka minnena. Detta är i linje med den mer moderna interferensteorins förklaring till "glömska". Enligt den teorin blir minnen mer svåråtkomliga för att senare information stör återerinnringen, till exempel genom att strukturen kan förändras så tidigare sökvägar inte längre är framkomliga (Passer & Smith, 2007; Ashcraft 1994; Langlet, 1990).

Inkodning och interferens är två viktiga utgångspunkter i nedanstående avsnitt, som fortsätter diskussionen om minnesfel avgränsat till surveyfrågor.

Sammanfattande punkter:

- Tidigare var försvagningsteorin populär (att tiden i sig försvagar minnen) men det är den inte längre.
- En mer modern teori är interferensteorin. Enligt detta synsätt försvinner inte minnen. De kan istället bli mer svåråtkomliga i nätverket, till exempel för att de påverkas av nyinkodad information eller för att minnet omorganiseras.

Allmänt om källor till minnesfel

Förra avsnittet beskrev hur vårt förflutna finns lagrat i den del av långtidsminnet som kallas det episodiska minnet. När vi ställer frågor till respondenterna om deras förflutna ber vi dem alltså söka efter svaret i det episodiska minnet. Hur bra eller dåligt den återerinnringen går avgör kvaliteten på den data som samlas in. Detta avsnitt går djupare in på problem som kan uppstå vid återerinnringen och därför kan leda till minnesfel för respondenterna och mätfel i data.

Det kan finnas fyra källor till minnesfel i surveyfrågor (Tourangeau, 2000):

1. **Inkodning:** Att informationen bara delvis eller inte alls inkodats i långtidsminnet och därför är svår eller omöjlig att återhämta.

2. **Förvrängning av lagrade minnen:** Informationen kan ha blivit inkodad korrekt men senare händelser kan förvränga den till något felaktigt.
3. **Misslyckad återerinring:** Även utan brister i inkodningen kan vi misslyckas med återerinringen, speciellt om vi är okoncentrerade eller omotiverade.
4. **Rekonstruktion:** När vi fyller i bristfälliga minnen med generisk information eller med andra minnen.

Angående *inkodning* krävs det först att vi måste ha uppmärksammat en händelse för att kunna koda in den i minnet. Om vi inte uppmärksammar en händelse kodas den inte heller in i långtidsminnet. Då kan den förstås inte heller återerinnras vid besvarandet av formulärfrågor. I en studie med relativt detaljerade frågor om föräldrars kunskap om sina småbarns vaccinering innehöll svaren lika mycket felaktigheter tio veckor efter vaccinering jämfört med direkt i anslutning till vaccineringen. Bristen på kunskap tio veckor efter vaccineringen verkade alltså *inte* bero på glömska. Författarna drar slutsatsen att orsakerna till bristande rapportering var att informationen aldrig hade inkodats i långtidsminnet snarare än problem med återerinring. Detta kan förekomma inom många områden i undersökningar och är viktigt att tänka på. Om inte informationen inkodats kan den inte heller återerinnras, oavsett hur mycket resurser som läggs på att underlätta och främja återerinring (Lee, Brittingham, Tourangeau, Willis, Ching, Jobe & Black, 1999; Tourangeau, 2000).

Men inkodning är också en process i sig och ju grundligare den sker desto lättare är informationen att återhämta. Om inkodningen är yttlig och svag kan informationen vara för bristfällig för ett korrekt svar. Inte heller då har respondenten tillräcklig information för att egentligen kunna besvara frågan korrekt. Enligt nätverksteorin består grundlig inkodning av att minnet placeras på lämpligt ställe i nätverket och länkas till många relevanta noder. Länkarna är viktiga eftersom det är via dem associationer sprids i minnessystemet. Att vi inte minns något kan bero på att vi söker på fel platser i nätverket, som kanske saknar länkar till den sökta informationen. Att stimulera återerinring hos respondenter innebär därför ofta att aktivera relevanta noder i minnesnätverket som kan tänkas vara länkade till den eftersökta informationen (läs mer i senare avsnitt i detta kapitel) (Tourangeau, 2000).

En studie undersökte respondenters minne av betalning av underhåll till barn och hur de angivna summorna förhöll sig till dokumenterade uppgifter. Författarna undersökte tre faktorer om företeelsen: 1) komplexitet (hur ofta, regelbundna och lika betalningarna var), 2) klarhet (hur distinkt betalningen var från andra inbetalningar, till exempel försäkringar) och 3) känslointensitet (hur nöjd eller missnöjd respondenten var med beloppet). De fann att alla tre predicerade svarsfel jämfört med dokumentationen (Dykema & Schaeffer, 2000). Dessa är alltså faktorer som kan göra att minnen blir mer tydliga och därför lättare att komma åt.

Förutom inkodningens grundlighet verkar också förfluten tid ha stor påverkan på minnens tillgänglighet. Detta kanske kan bero på att information som inte aktiveras med tiden blir svårtillgängligare i sig (se försvagningsteorin i förra avsnittet). Men troligare är alltså att det beror på att informationen i nätverket utökas och formuleras om när fler händelser fylls på över tid. Det är lätt att minnas en unik händelse som ett läkarbesök eller en idrottsmatch. Men när vi upplever fler händelser av samma sort är inte den första längre unik och kan sammanblandas med eller försvinna i flertalet. Enligt teorier om minnet bildas ett generiskt minne när det finns många händelser av samma art och som inkluderar det händelserna har gemensamt och inte det unika. Det generiska minnet är alltså baserat på gemensamma nämnare för de olika händelserna. Detta kan vara en förklaring till varför det är enklare att minnas "hur det oftast är" än hur det var vid ett specifikt men ordinärt tillfälle. Detta är också en risk i frågeformulär – att respondenterna svarar utifrån hur det oftast är snarare än hur det faktiskt var (Tourangeau, 2000).

En ytterligare felkälla är därför när det finns information i respondentens minne som är bristfällig och därför fylls i, *rekonstrueras*, med generell information. Den här typen av minnesfel kan alltså ske i en systematisk riktning – likriktning snarare än avvikelser. Det beror på att den generiska informationen baseras på det som är gemensamt för flera händelser av samma art, snarare än det som är unikt för de enskilda händelserna (Tourangeau, 2000).

När informationen kodats in finns den i minnessystemet. Hur ofta den aktiveras och upprepas (till exempel genom att vi tänker på det) påverkar hur tillgänglig den är vid återerinring. Tydliga och lättillgängliga minnen är dock inte detsamma som sanna minnen. Minnen kan *förvrängas* över

tid. Ett exempel är så kallade fotoblixtminnen – skarpa minnen som fotografier, vid specifika och starka situationer. Hit hör till exempel situationen när man fick beskedet att Olof Palme blivit mördad eller om terroristattentaten den elfte september. Många har tydliga minnesbilder av sådana extraordinära situationer. Forskning om sådana minnen visar att även om de minnesbilderna är tydliga och starka för individen så är de inte nödvändigtvis korrekta. Det beror på att minnen inte är fasta i formen utan kan påverkas av efterhandsinformation (detta kan vara ett problem vid till exempel vittnesmål), våra egna fantasier (till exempel att vi tänker igenom ett alternativt scenario som sedan kan sammanblandas med det verkliga) och att vi själva utvecklas och ser på vårt förflutna på ett annorlunda sätt (Myers, 2002). Minnen är alltså inte huggna i sten utan kan utvecklas och förvrängas från hur det faktiskt var.

Situationella faktorer kan också leda till att återerinjering misslyckas. Störningsmoment, respondentens motivation eller intervjuens tempo kan leda till att respondentens inte söker tillräckligt djupt i minnet för att hitta informationen (Tourangeau, 2000).

Alla dessa fyra källor (inkodning, rekonstruktion, förvrängning och misslyckad återerinjering) kan alltså leda till fel. En del av detta kan vara svårt att göra så mycket åt som frågekonstruktör men man bör vara medveten om det när man arbetar med retrospektiva data. Nedanstående avsnitt fokuserar på minnesprocessen och felkällor vid två specifika typer av återerinjering – när något hände (datering) och hur många gånger det hände (frekvensbestämning).

Sammanfattande punkter:

Det kan finnas fyra källor till minnesfel i surveyfrågor:

- Inkodning – Att informationen bara delvis eller inte alls inkodats i lång-tidsminnet och därför är svår eller omöjlig att återhämta.
- Förvrängning av lagrade minnen – Informationen kan ha blivit inkodad korrekt men senare händelser kan förvränga den till något felaktigt.
- Misslyckad återerinjering – Även utan brister i inkodningen kan vi misslyckas med återerinjeringen, speciellt om vi är okoncentrerade eller omotiverade.
- Rekonstruktion – När vi fyller i bristfälliga minnen med generisk information eller med andra minnen.

Att minnas *när* något hände (datering)

I undersökningar är det ganska vanligt att fråga om när något inträffade.

EXEMPEL

När flyttade du in i din nuvarande bostad?

Hur länge var du sjuk?

Frågor om datum och tid kan variera på många olika sätt, till exempel frågor om specifika datum eller förfluten tid mellan eller efter händelser. Tourangeau, Rips och Rasinski (2000) delar upp det i tre huvudtyper:

- **Ett specifikt tillfälle** – när en händelse inträffade.

EXEMPEL: *Vilket datum träffade du läkaren?*

- **Hur lång tid något tog** – löptid.

EXEMPEL: *Hur länge var du borta från arbetet?*

- **Förfluten tid efter en händelse.**

EXEMPEL: *För hur länge sedan hände detta?*

Det finns många varianter av frågor om *specifika tillfällen*. De kan naturligtvis variera angående vilken typ av händelser de fokuserar på, men även i precision, om svaret till exempel ska anges i dagar eller månader och om det finns en referensperiod eller inte (Tourangeau, Rips & Rasinski, 2000).

Hur lång tid något tog – löptiden – kan egentligen räknas ut med hjälp av start- och slutpunkt men är trots det ganska vanlig i undersökningar. Kanske för att det ibland är just löptiden som är intressant, inte start- eller slutpunkt. Det gäller exempelvis om man vill mäta hur länge någon varit arbetslös. Frågorna kan variera i precision, till exempel efterfråga olika stora tidsenheter. Ett sätt att signalera hög precision är att lägga in en markör i texten (*exakt* hur många dagar ...). På motsvarande sätt kan en osäkerhetsmarkör läggas in för att signalera låg precision (*ungefär* hur lång tid ...) (Tourangeau, Rips & Rasinski, 2000).

Frågor om *förfluten tid efter en händelse* är ganska ovanliga i undersökningar. Det beror bland annat på att den oftast kan räknas ut ändå om man känner till när en händelse inträffade (Tourangeau, Rips & Rasinski, 2000).

Hur går återerinringen till?

Hur besvaras frågor om *när* något hände? Det finns flera strategier som respondenterna kan använda. En möjlighet är att återhämta det exakta datumet. Minnesforskningen antyder dock att vi inte brukar koda in händelsers datum. I en studie (Friedman, 1987) var det endast 10 procent som direkt återhämtade datumet. När så sker brukar dock informationen vara ganska korrekt. En viktig insikt är att kalendertid inte nödvändigtvis är någon viktig faktor för minnet. Om inte månaden i sig har en viktig funktion för minnet är det tveksamt varför den ska sorteras in på det sättet. Händelser är sällan heller inkodade efter "senaste månaden" eller "senaste året", vilket är vanliga tidsmässiga avgränsningar i surveyfrågor. Dessa är inga bra markörer för minnet eftersom de är föränderliga – "senaste månaden" syftar på olika månader allteftersom tiden går och månader skiftar. Därför överensstämmer de tidbestämningar som ofta används i frågor inte nödvändigtvis särskilt väl mot minnets struktur (Tourangeau, Rips & Rasinski, 2000; Tourangeau, 2000).

En annan möjlighet för individen, om datumet är svårt att återhämta, är att anknyta till någon närliggande händelse som enklare kan fastställas i tid (till exempel ens födelsedag) eller att återhämta andra detaljer som kan hjälpa till att fastställa tidpunkten (på hösten brukar jag till exempel ...). Studier som undersökt sådana riktmärkens roll varierar i hur många testpersoner som använt strategin (mellan 10–30 %). Förutom att direkt återhämta datumet verkar det vara den bästa strategin (Tourangeau, Rips & Rasinski, 2000; Tourangeau, 2000).

Återerinring i praktiken

Hur går då återerinringen till i praktiken? Vilket år och vilken månad införskaffade du din första mobiltelefon? Om du försökte besvara frågan, hur gick du tillväga? Vilka vägar tog tankarna? Kanske var den första mobilen en viktig händelse som du automatiskt vet när den inträffade? Troligtvis behöver dock de flesta personer söka djupare i minnet och använda riktmärken för att kunna besvara frågan. En del kanske kan utgå från något tidsmässigt ankare, som att man fick sin första mobil när man började sitt första jobb. När man började sitt första jobb vet kanske de flesta men kan annars till exempel fastställas utifrån vilket år man tog examen. Någon vet kanske vilken modell den första mobilen var och vet också under vilka år den modellen var aktuell. Någon annan kanske minns att han eller hon blev uteläst i snön men räddad av den då nya

mobilen. Personen kan då fastställa att mobilen köptes på vintern. Möjligheterna är många och blir ännu fler ju fler minnen som nystas upp.

Återerinring kan alltså ses som en cyklisk process. Hämtad information interagerar med tidigare och introducerar nya begränsningar och möjligheter som påverkar vilken information som hämtas från minnet i nästa vända. Ju bättre en person minns en händelse desto större chans att de också kan placera den rätt i tiden.

Att utgå från en tidsperiod som ankare är dock inte en lika skarp strategi som att minnas datumet i sig, eftersom det finns en risk att placera händelsen fel i tiden även om man minns den. Detta kallas för teleskopering (Tourangeau, 2000).

Teleskopering

Det händer att respondenter minns en händelse men placerar den fel i tiden. Det kan innebära att de minns en händelse som att den inträffade längre tillbaka i tiden än vad den faktiskt gjorde. Detta kallas för teleskopering bakåt. Det motsatta kan också förekomma – att respondenten minns en händelse som att den inträffade mer nyligen än den faktiskt gjorde. Det kallas för teleskopering framåt. När människor ska placera händelser i tiden kan de alltså flytta dem framåt eller bakåt i tiden, jämfört med när de faktiskt inträffade (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Vad gäller mätningen i undersökningar är teleskopering framförallt relevant i anslutning till referensperiodens gränser. Teleskopering kan förstås ske även inom en referensperiod men har då ingen betydelse för mätningen. Eftersom referensperioders främre avgränsning oftast utgörs av svarstillfället (nutid, när respondenten deltar i intervjun eller fyller i enkäten) är det framförallt referensperiodens borte gräns som kan drabbas av teleskopering. Ett exempel är om en fråga har referensperioden "den senaste månaden". Då är den främre gränsen svarstillfället och den borte gränsen en månad tillbaka från svarstillfället (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Vid den borte gränsen av referensperioden innebär teleskopering framåt en överskattning – händelser som egentligen inträffade tidigare än referensperioden flyttas fram så de blir inkluderade. Teleskopering bakåt innebär tvärtom – att händelser som egentligen inträffade inom

referensperioden flyttas bakåt så att de blir exkluderade – en underskattning (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Vid fristående referensperioder som inte är knutna till svarstillfället, till exempel en referensvecka som kan ligga någon vecka tidigare än svarstillfället, kan även den främre gränsen av referensperioden drabbas av teleskopering. Konsekvenserna blir då tvärtom jämfört med de beskrivna i förra stycket. Teleskopering framåt resulterar då i en underskattning (händelser flyttas fram, ut ur referensperioden) och teleskopering bakåt resulterar i en överskattning (händelser flyttas bak, in i referensperioden). Men det är som sagt vanligare att den bortre gränsen av referensperioden drabbas av teleskopering eftersom den främre gränsen ofta utgörs av svarstillfället (det går inte att flytta händelser framåt från nutiden eller bakåt från framtiden in i nutiden) (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Fenomenet teleskopering kan alltså leda till både under- och överrapportering av händelser. Vilket är vanligast? Enligt vissa forskare är överrapportering vanligast (Wärneryd, 1990). Det kan bero på att minnen blir mer bristfälliga med tiden och därför är tydligare för mer nutida än för senare händelser. Då borde risken att placera en händelse fel i tiden vara större för händelser längre tillbaka i tiden (utanför referensperioden) än för händelser kortare tillbaka i tiden (innanför referensperioden). Det är därför troligare att en respondent flyttar en tidigare händelse framåt i tiden än en mer nutida händelse bakåt i tiden. Resultatet blir då överrapportering (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Vad beror då fenomenet på? Det kan även finnas andra förklaringar till resultaten i de studier som undersökt saken. Respondenterna kanske är mer benägna att inkludera en händelse om de väl minns den, oavsett referensperioden, för att de hellre vill ge för mycket än för lite information. Sådana respondentbeteenden kan ge samma typ av effekt som teleskopering men brukar inte betraktas som teleskopering i sig.

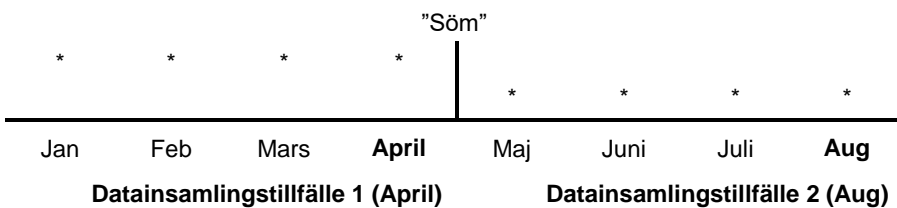
Teleskopering har med begränsningar i minnet att göra – respondenten minns en händelse men placerar den fel i tiden. Det kan i praktiken finnas flera typer av minnesbrister som kan ge samma resultat. Underrapporteringar kan ju till exempel bero på att 1) en händelse inte finns alls i respondentens minne, 2) händelsen finns där men

respondenten misslyckas med återerinringen eller 3) respondenten lyckas återerinnra händelsen men placerar den fel i tiden (teleskopering). Dessa tre olika typer av problem kräver olika typer av lösningar. Problem 1 går inte att lösa. Vid problem 2 bör fokus ligga på att främja återerinringen av händelsen i sig. Vid problem 3 är det snarare den tidsmässiga återerinringen som bör främjas (Sudman, Bradburn & Schwarz, 1996; Langlet, 1990; Tourangeau, Rips & Rasinski, 2000).

Sömeffekter

I longitudinella undersökningar som har flera mätpunkter mellan insamlingstillfällena kan det uppstå så kallade sömeffekter ("seam effects" på engelska) (Tourangeau, Rips & Rasinski, 2000). Ett exempel är en undersökning där respondenterna intervjuas var fjärde månad (april och augusti i figuren nedan) men svarar för alla månader mellan intervjutillfällena – den aktuella månaden vid intervjutillfället och månadsvis tre månader bakåt. Mellan den sista mätpunkten vid det förra intervjutillfället (april i figuren nedan) och den första mätpunkten vid det senare intervjutillfället (maj i figuren nedan) uppstår en söm. En sömeffekt är när skillnaderna mellan mätpunkterna i sömmen (april/maj) avviker för starkt från dem inom samma insamlingstillfälle (Tourangeau, Rips & Rasinski, 2000). Effekten kan vara så pass stor att den utgör ett problem. Detta är inte en säsongseffekt, så vad är det?

Figur 11.1



Det kan finnas flera skäl till sömeffekter. Här fokuseras det främst på de kognitiva. Orsaker till sömeffekter kan vara att respondenten svarar med så lite ansträngning som möjligt (satisficing) eller att de inte minns ordentligt. Respondenterna kan till exempel utgå från den senaste månaden eller göra en summering och låta det svaret gälla för alla månader inom intervallet (så kallad "constant wave response"). Om respondenten använder denna strategi blir förändringar inom intervallet

underrapporterade och hamnar i sömnen till förra intervallet. Liknande effekter kan dock kanske också uppstå för att minnet kan vara sämre för mer avlägsna än närliggande händelser. Men då borde det finnas sådana tendenser även inom intervallet, inte bara i sömmarna (Tourangeau, Rips & Rasinski, 2000).

I en studie besvarade respondenterna en enkät varje vecka i åtta veckor. De intervjuades vid två tillfällen (efter vecka 4 och 8). Data från enkäten användes som jämförelse för att utreda eventuell sömmeffekt i intervjuerna. Studien visade sömmeffekter och att storleken på effekten var relaterad till minneskraven i frågan. Effekten var också starkare om ämnesrelaterade frågor kom i anslutning till varandra. För att minska sömmeffekten kan det därför vara en idé att gruppera frågorna kronologiskt snarare än tematiskt. Ett annat alternativ är att använda beroende intervju där respondentens svar från förra referensperioden upprepas. Denna teknik har dock sina egna problem (till exempel att eventuella fel består snarare än upptäcks och justeras) (Rips, Conrad & Fricker, 2003). Att ha fler mättillfällen och kortare referensperioder kan vara ett sätt att mildra effekten. Andra strategier för att hjälpa återerinring kan också minska risken för sömmeffekter (Callegaro, 2008).

Sammanfattande punkter:

- Tre typer av frågeansats för när något hände:
 1. Ett specifikt tillfälle – när en händelse inträffade. Exempelvis: Vilket datum träffade du läkaren?
 2. Hur lång tid något tog – löptid. Exempel: Hur länge var du frånvarande från arbetet?
 3. Förfluten tid efter en händelse. Exempel: För hur länge sedan hände detta?
- Hur besvaras frågor om när något hände? Det finns flera strategier som respondenterna kan använda. En möjlighet är att återhämta det exakta datumet. En annan möjlighet, om datumet är svårt att återhämta, är att anknyta till någon närliggande händelse som enklare kan fastställas i tid.
- Återerinring kan alltså ses som en cyklisk process. Hämtad information interagerar med tidigare och introducerar nya begränsningar och möjligheter som påverkar vilken information som hämtas från minnet i nästa vända.

- Teleskopering är när respondenter minns en händelse men placerar den fel i tiden.
- Det kan innebära att de minns en händelse som att den inträffade längre tillbaka i tiden än vad den faktiskt gjorde. Detta kallas för teleskopering bakåt.
- Det motsatta kan också förekomma – att de minns en händelse som mer nyligen än den faktiskt inträffade. Det kallas för teleskopering framåt.
- Vad gäller mätningen i undersökningar är teleskopering framförallt relevant i anslutning till referensperiodens gränser.
- Fenomenet teleskopering kan leda till både under- och överrapportering av händelser. Vilket är vanligast? Enligt vissa är överrapportering vanligast. Det kan bero på att minnen blir mer bristfälliga med tiden och är tydligare för mer nutida än för senare händelser.
- I longitudinella undersökningar som har flera mätpunkter mellan insamlingstillfällena kan det uppstå så kallade sömeffekter ("seam effects" på engelska): orimliga skillnader i "sömmen" mellan insamlingstillfällena.
- En förklaring kan vara att respondenterna utgår från den senaste månaden eller gör en summering och låter det svaret gälla för alla månader inom intervallet (så kallad "constant-wave response"). Om respondenten använder denna strategi blir förändringar inom intervallet underrapporterade och hamnar i sömmen till förra intervallet.
- För att minska sömeffekten kan det vara en god idé att gruppera frågorna kronologiskt snarare än tematiskt.
- Att ha fler mättillfällen och kortare referensperioder kan vara ett sätt att mildra effekten.

Att minnas *antalet gånger* något hände (frekvensbestämning)

I frågor ombeds ofta respondenterna att inte bara minnas någonting, utan även att besvara antalet gånger något inträffade. Den typen av frågor kan kräva bedömningar eller beräkningar. Ibland behöver information sammanfattas eller beräknas för att det ska kunna gå att svara. Detta är en väldigt vanlig typ av fråga i undersökningar.

EXEMPEL

Hur ofta motionerade du förra månaden?

Hur många gånger har du besökt en läkare under det senaste året?

Den här typen av fråga innebär att respondenten måste använda både minnet och göra bedömningar.

Personer har ofta svårt för att uppskatta frekvenser korrekt. Felen kan vara stora men hur stora de blir beror förstås på frågan och vilken svarsstrategi respondenten använder vid besvarandet (Tourangeau, Rips & Rasinski, 2000).

Om svarsstrategier vid denna frågetyp

Kognitiva tester, till exempel att tänka högt eller probing (läs mer i kapitel 14), har visat att respondenter kan använda många olika strategier för att komma fram till ett svar när antal eller frekvenser efterfrågas. Tourangeau, Rips och Rasinski (2000) ger några exempel:

1. Respondenten kanske har frekvensen *lagrad som en egen post* i minnet. Svarsprocessen blir då ganska simpel. Det är dock sannolikt inte särskilt många frekvenser som lagras på detta sätt utan endast lågfrekventa händelser av stor betydelse för individen, till exempel hur många platser man bott på.
2. En annan strategi är att återerindra *informationen om alla aktuella tillfällen och summera*. Det kan emellertid vara väldigt krävande för respondenterna om det handlar om många tillfällen. Det är nog ingen särskilt vanlig strategi förutom inom ämnesområden där det är rimligt och frågorna "tvingar" respondenten att begrunda varje tillfälle.
3. En tredje strategi är *uppskatta med hjälp av en frekvens* som redan finns i minnet.
4. Respondenterna kan också räkna ut en frekvens för en kortare period eller ett intervall som sedan *omräknas för en längre period eller ett intervall*.
5. Respondenterna kan också göra en *uppskattning utifrån svarsalternativen*. Om de till exempel upplever att de gjort en aktivitet väldigt ofta kanske de väljer det högsta alternativ oavsett vilka faktiska frekvenser som ingår i det svarsalternativet.
6. De kan också svara utifrån *"första intrycket"* utan vidare eftertanke (Sudman, Bradburn & Schwarz, 1991).

Vad styr val av strategi?

Vad är det då för faktorer som påverkar vilken strategi respondenten väljer? Tourangeau, Rips och Rasinski (2000) beskriver följande faktorer:

- Frågans krav. Med det menas hur många tillfällen eller händelser som ingår i svaret. Högre frekvenser ökar risken för att tumregler och uppskattningar används. På motsvarande sätt ökar längre referensperioder risken för att tumregler och uppskattningar används.
- Respondentens motivation. En studie visade att testpersoner ansträngde sig mer när de fick veta att deras svar var viktiga (Burton & Blair, 1991).
- Minnets tillgänglighet. Ju mer distinkta och viktiga händelser är, desto mer tillgängliga är de i minnet. Triviala och vardagliga händelser är svårare att minnas (Dykema & Schaeffer, 2000).
- Andra förhållanden kring uppgiften som till exempel frågeformuleringen eller vilken tid det finns för genomsökning av minnet.

Antalet händelser verkar alltså ha betydelse. Om de är för många tenderar respondenter att övergå från räkning till uppskattning. I en studie räknade de flesta om det var tre eller färre händelser. Från fyra och uppåt använde fler och fler uppskattningar. Hur regelbundna och lika händelser är påverkar också val av strategi. Om händelser är regelbundna och likartade ökar användningen av uppskattningar. Om händelser är oregelbundna och olikartade ökar tvärtom användningen av räkning. Referensperiodens längd kan ha också betydelse (se nedan). Detta sammanfaller dock med antalet händelser eftersom det senare ökar i takt med det förra, och det finns tecken på att om man kontrollerar antalet händelser så har inte referensperiodens längd någon större effekt i sig (Tourangeau, Rips & Rasinski, 2000).

Svarstrategier och datakvalitet

Att räkna alla aktuella tillfällen har ansetts som en bättre strategi för mätningen än uppskattning. Det är nödvändigtvis inte alltid sant. Vad är det som kan påverka hur korrekt en uppskattning är? Det verkar finnas en viss stabilitetsbias, det vill säga att personer utgår från något i minnet och sedan missar förändringar över tid eller temporära skillnader beroende på till exempel semester eller vädret. Vissa respondenter kan försöka justera för sådant men kan misslyckas så att de antingen överkorrigerar eller underkorrigerar (det sistnämnda är vanligast). Det

kan också vara så att utgångspunkten, tagen ur minnet, är bristfällig. Det leder antingen till en underskattning (händelser missas) eller överskattning (till exempel teleskopering) (Sudman, Bradburn & Schwarz, 1991).

Det finns ganska lite forskning som jämför räkning och uppskattning. Det finns indikationer på att uppskattning leder till större fel än räkning, men att felen ofta är osystematiska (de kan vara både överskattningar och underskattningar) medan felen från räkning kan vara systematiska (Sudman, Bradburn & Schwarz, 1991).

Vad gäller svar baserade på "första intrycket" är respondenterna ofta inte själva medvetna om hur de kommit fram till sitt svar. Dessa svar borde vara förknippade med mest osäkerhet och fel. "Första intrycket" verkar vara en vanligare strategi om frågan är vag och generell (Sudman, Bradburn & Schwarz, 1991).

Referensperiodens påverkan på förståelsen av frågan

När frågekonstruktören väljer referensperiod för frågorna utgår de ofta från en förväntad frekvens av det som efterfrågas. Om händelsen eller beteendet är ovanligt väljs en längre referensperiod för att undvika att missa förekomsten. Vid vanliga händelser och beteenden väljs en kortare referensperiod för att undvika minnesfel.

Referensperioden är dock inte bara en frågeteknisk avgränsning utan kan också påverka vad respondenten uppfattar som frågans utgångspunkt (Schwarz, Strack, Muller & Chassein, 1988). Vissa frågeområden är otydliga till naturen. Känslor är ett sådant exempel. En känsla som till exempel ilska har en stor vidd av uttryck som kan inkluderas. Referensperioden kan påverka vilken intensitet av känslan som respondenten uppfattar att frågeställaren menar. Om det efterfrågas hur många gånger respondenten var arg "igår", uppfattas nog "arg" som något mildare än om referensperioden är "senaste året". Det är en ganska naturlig tolkning. Om det var mild ilska som avsågs vore en referensperiod på ett år orimligt lång med tanke på minnets begränsningar. På motsvarande sätt, om det var extrem ilska som avsågs vore "igår" en onaturligt kort referensperiod eftersom de flesta personer inte blir vansinniga dagligen. Ovanstående bekräftas av forskning av Schwarz och kollegor. När referensperioden utgår från "förra veckan" rapporteras 2,5 gånger så stor förekomst som om referensperioden utgår från ett år.

En konkurrerande hypotes till ovanstående är dock glömska – att mindre extrema och intensiva känslor glöms bort med tiden. Denna hypotes ger i praktiken samma utfall som den om att referensperioden påverkar hur begreppet (till exempel ilska) uppfattas. En studies (Winkielman, Knäuper & Schwarz, 1998) resultat förkastar dock glömska-hypotesen, eftersom effekten försvann om begreppet (ilska) definierades i frågan. Om effekten berodde på glömska skulle det inte ha någon betydelse om begreppet definierades eller ej. Det är viktigt att kunna särskilja vilken av flera konkurrerande mekanismer som är giltig, eftersom det kan påverka hur man ska hantera problemet. Om minnet var förklaringen vore en lämplig åtgärd att stödja återerinring. Om det istället är vagheten i begreppet, som i studien ovan, ska istället resurser läggas på att förtydliga frågan.

Referensperioden bör vara anpassad utifrån det som efterfrågas och samtidigt måste frågeformuleringen vara tydlig. Annars kan förståelsen styras av annan information i frågan, till exempel referensperioden.

Sammanfattande punkter:

- Svarsstrategier för frågor om *antalet gånger*:
 - Respondenten kan ha frekvensen lagrad som en egen post i minnet.
 - En annan strategi är att återerinnra informationen om alla aktuella tillfällen och summera.
 - En tredje strategi är att uppskatta med hjälp av en frekvens som redan finns i minnet.
- Respondenterna kan också räkna ut en frekvens för en kortare period eller ett intervall som sedan generaliseras till en längre period.
- Respondenterna kan också göra en uppskattning utifrån svarsalternativen. Om de till exempel upplever att de gjort en aktivitet väldigt ofta kanske de väljer det högsta alternativ oavsett vilka faktiska frekvenser som ingår i det svarsalternativet.
- De kan också svara utifrån "första intrycket" utan vidare eftertanke.

- Valet av strategi styrs av frågans krav på ansträngning, respondentens motivation, informationens tillgänglighet i minnet och andra förhållanden kring uppgiften som till exempel frågeformuleringen eller vilken tid det finns för genomsökning av minnet.
- Referensperioden är inte bara en frågeteknisk avgränsning utan kan också påverka vad respondenten uppfattar som frågans utgångspunkt.

Att främja återerinring och minska minnesfel

Många frågor i surveyundersökningar frågar om det förflutna och är därför beroende av respondenternas minne. Några problem kan vara att respondenten inte minns alls, bara minns en del eller minns fel angående det som efterfrågas. Hur kan man motverka fel och underlätta för respondenterna?

När vi frågar efter det förflutna måste respondenterna söka efter information i den del av långtidsminnet som kallas episodiska minnet (se tidigare avsnitt i detta kapitel). Denna process, när information hämtas från långtidsminnet, kallas återerinring och att underlätta den processen minskar både risken för fel och respondentbördan.

Inkodning

Det första man behöver tänka på som frågekonstruktör är personers inkodning av en händelse – processen där information kodas in i minnet. Människor exponeras för mycket information men långt ifrån allt uppmärksammas och ännu mindre kodas in i minnet. Man kan till exempel fråga några arbetskamrater hur många gånger de var tvungen att stanna för rödljus på resan till jobbet och få komma nog känna sig säkra på det, trots att rödljus är en händelse som uppmärksammas eftersom man måste stanna. Som frågekonstruktör innebär det här att man inte ska fråga om saker som personer inte kan förväntats ha kodat in i minnet. Naturligtvis bör man inte heller lägga resurser på att främja återerinring av sådana saker eftersom det är en lönlös strategi. Det finns ingen information att hämta.

Inkodning är dessutom en process som kan ske med olika djup. Om individen begrundar och arbetar med informationen (t.ex. associerar) är det troligare att den kodas in grundligare i minnet, till exempel knyts till

fler relevanta noder i minnesnätverket. För att ta ett exempel från skolvärlden är det här nog en av anledningarna till att vissa lärare vill att eleverna antecknar med egna ord snarare än kopierar lärarens anteckningar. Det här medför också att inkodningen blir personlig eftersom olika personer till exempel kan associera på olika sätt. Konsekvensen blir att information som kan främja återerinring för en person inte nödvändigtvis passar en annan person lika bra. Ett exempel på detta är att det är lättare att minnas innehållet i en föreläsning om man läser sina egna föreläsningssanteckningar jämfört med om man läser någon annans. Detta beror på att ens anteckningar är ett resultat av den egna förståelsen och därför matchar minnet av föreläsningens innehåll bättre än någon annans anteckningar (Passer & Smith, 2007).

Men inkodning är inte bara en intellektuell eller kognitiv process. Forskning har visat att känslor, kontext och miljöer kan påverka. Exempelvis kan man minnas bättre om man befinner sig i samma miljö som händelsen inträffade i (Passer & Smith, 2007).

Att främja återerinring – allmänt

Att främja återerinring innebär att presentera s.k. ledtrådar (retrieval cues på engelska) som matchar hur informationen kodats in i minnet. Allra bäst är det om ledtrådarna är både korrekta och specifika för det avseende de har. Om en ledtråd är för generell är det inte säkert att den är behjälplig oavsett om den korrekt. På samma sätt leder en specifik ledtråd inte rätt om den inte är korrekt. Det borde vara svårast att erbjuda specifika ledtrådar till händelser som i liten grad är unika. Det betyder till exempel att om en person bara har haft en sjukhusvistelse borde det inte vara svårt med ledtrådar som specifikt knyter an till den vistelsen. Men om en person har haft många sjukhusvistelser är det svårare att hitta ledtrådar som är specifika för just en, och inte en annan, vistelse (Shum & Rips, 1999). Den personen borde kunna minnas sjukhusvistelser i allmänhet men kanske inte en specifik vistelse, om den inte var unik och frågeformuleringen anknöt till detta. Detta kan begränsa vad man kan fråga och få information om.

Tekniker att främja återerinring i undersökningar

Eftersom minnesfel kan vara en allvarlig felkälla i undersökningar med frågor om det förflutna, har det utvecklats ett antal tekniker för att främja respondenternas återerinring. Återerinring är dock en komplex och i viss

mån personlig process (i betydelsen att ledtrådarna kan vara personliga, se ovan). Det kan vara svårt att hitta en standardlösning som fungerar för alla respondenter. Kanske är det skälet till att de flesta teknikerna nedan är kopplade till intervjuundersökningar som är en mer flexibel metod än pappersenkäter. Ett annat skäl kan vara att besöks- och telefonintervjuer har varit de dominerande metoderna för många individ- och hushållsundersökningar.

Teknikerna nedan beskrivs utförligt i Langlet (1990) ("Att ta hänsyn till minnets begränsningar", i "Att fråga" av Wärneryd m.fl.) och är endast kort sammanfattade här.

- **Underlättad erinring:** Tänkbara ledtrådar presenteras för att hjälpa respondentens minne på traven. Om man ska fråga om kost kan man till exempel presentera en lista eller en bildkarta med matvaror för respondenten. En risk är dock att exemplen som finns med på listan eller på bilderna blir överrepresenterade jämfört med sådant som inte finns på listan.
- **Dokumentation:** Här används till exempel kvitton, räkningar eller kontrakt som underlag för frågorna. Tekniken minskar minnesfel men kan vara svår att tillämpa i andra datainsamlingsmetoder än intervju.
- **Kalenderintervju:** En flexibel intervjuform som innebär att man använder en kalender som utgångspunkt för frågorna, för att till exempel hjälpa respondenten att placera saker rätt i tiden. Tekniken kan minska minnesfel men fungerar främst i besöksintervjuer, är tidskrävande och kan öka intervjuarvariansen.
- **Uppvärmning:** Ett exempel är att inleda med några frågor som har som primärt syfte att stimulera, "värma upp", minnet om det man egentligen vill fråga om. Om man till exempel ska fråga om matinköp kan uppvärmningsfrågorna innehålla ämnen som i vilken butik respondenten brukar handla, om de har inköpslista och vilka veckodagar de brukar handla. Tekniken kan minska minnesfel men ökar respondentbördan och administrationstiden.
- **Nedbrytning:** Ett sätt att strukturera frågeinnehållet. Istället för att ställa en övergripande fråga som respondenten spontant på sin höjd kan uppskatta svaret på, kan flera mer specifika frågor ställas. Om de specifika frågorna är täckande kan man sedan räkna ut den övergripande nivån. Tekniken har visat sig fungera men innebär att frågeformuläret blir längre.

- **Anknyta till rätt begrepp:** Det innebär att i frågeformuleringen sträva efter använda de ord som bäst hjälper respondenterna att minnas, det vill säga ord som fungerar som bra ledtrådar för minnet. Denna teknik kan i princip användas för alla frågor oavsett datainsamlingsmetod. Men det är inte helt enkelt att känna till eller hitta rätt ord. Det kan krävas kvalitativa förtester så man får insikt i respondenterna språkbruk och hur de tänker kring ämnet man vill undersöka.
- **Långa frågor:** Detta har visat sig förbättra minnet i vissa fall i telefonintervjuer, förmodligen för att de ger respondenten mer betänketid. Dock får inte frågor förlängas så att innebörden förändras. Dessutom ökar det intervjutiden så tekniken bör bara användas i specifika fall.
- **Svarsalternativens utformning:** De har stor påverkan på respondentens svarsprocess. De kan fungera som ledtrådar för minnet men de kan också bli styrande (läs mer i kapitel 3).
- **Datainsamling nära inpå referensperioden:** Eftersom minnen generellt sett försämras med tiden underlättar det att ha datainsamlingen nära inpå den referensperiod som används i undersökningen.

Det finns alltså ett antal tekniker som kan hjälpa respondenten att minnas. De flesta teknikerna har dock en baksida, vilket oftast är att de förlänger svarstiden. Därför bör de generellt endast användas i specifika fall snarare än som generell lösning.

Sammanfattande punkter:

- Som frågekonstruktör behöver man tänka på personers inkodning av information i minnet. Om informationen inte har inkodats, kan den inte återerinnras.
- Att främja återerinnring innebär att presentera s.k. ledtrådar (retrieval cues) som matchar hur minnena är kodade och organiserade.
- Ledtrådarna bör vara både korrekta och tillräckligt specifika för det avseende de har. En svårighet är dock att det kan skilja sig mellan olika personer vad som utgör bra ledtrådar.

- Det finns flera tekniker för att främja återerinring: underlättad erinring, dokumentation, kalenderintervju, uppvärmning, nedbrytning, att anknyta till rätt begrepp, långa frågor, svarsalternativens utformning och ha datainsamlingen nära in på referensperioden.
- De flesta teknikerna har dock en baksida, vilket oftast är att de förlänger svarstiden. Därför bör de generellt endast användas i specifika fall snarare än som generell lösning.

Sammanfattning

Att fråga om det förflutna – sammanfattning

Långtidsminne och arbetsminne:

- Det finns olika teorier om hur minnet fungerar. En vanlig distinktion är att skilja på långtidsminne och arbetsminne.
- Långtidsminnet är ett minnessystem som innehåller lagrade minnen. Det kan ses som ett slags arkiv.
- Arbetsminnet är däremot ett betydligt mer begränsat minnessystem, både vad gäller tid och enheter, där vi temporärt kan behålla information som vi behöver för tillfället.
- Informationen behöver kodas in i långtidsminnet (episodiskt minne kallas den del som är relevant vid frågor om det förflutna).

Om glömska:

- Tidigare var försvagningsteorin populär (att tiden i sig försvagar minnen) men det är den inte längre.
- En mer modern teori är interferensteorin. Enligt detta synsätt försvinner inte minnen. De kan istället bli mer svåråtkomliga i nätverket, till exempel för att de påverkas av nyinkodad information eller för att minnet omorganiserar.

Allmänt om källor till minnesfel:

Det kan finnas fyra källor till minnesfel i surveyfrågor:

- Inkodning - Att informationen bara delvis eller inte alls inkodats i lång-tidsminnet och därför är svår eller omöjlig att återhämta.

- Förvrängning av lagrade minnen – Informationen kan ha blivit inkodad korrekt men senare händelser kan förvränga den till något felaktigt.
- Misslyckad återerinring – Även utan brister i inkodningen kan vi miss-lyckas med återerinringen, speciellt om vi är okoncentrerade eller omotiverade.
- Rekonstruktion – När vi fyller i bristfälliga minnen med generisk information eller med andra minnen.

Att minnas när något hände (datering):

- Tre typer av frågeansats för när något hände:
 1. Ett specifikt tillfälle – när en händelse inträffade. Exempelvis: Vilket datum träffade du läkaren?
 2. Hur lång tid något tog – löptid. Exempel: Hur länge var du frånvarande från arbetet?
 3. Förfluten tid efter en händelse. Exempel: För hur länge sedan hände detta?
- Hur besvaras frågor om när något hände? Det finns flera strategier som respondenterna kan använda. En möjlighet är att återhämta det exakta datumet. En annan möjlighet, om datumet är svårt att återhämta, är att anknyta till någon närliggande händelse som enklare kan fastställas i tid.
- Återerinring kan alltså ses som en cyklisk process. Hämtad information interagerar med tidigare och introducerar nya begränsningar och möjligheter som påverkar vilken information som hämtas från minnet i nästa vända.
- Teleskopering är när respondenter minns en händelse men placerar den fel i tiden.
- Det kan innebära att de minns en händelse som att den inträffade längre tillbaka i tiden än vad den faktiskt gjorde. Detta kallas för teleskopering bakåt.
- Det motsatta kan också förekomma – att de minns en händelse som mer nyligen än den faktiskt inträffade. Det kallas för teleskopering framåt.
- Vad gäller mätningen i undersökningar är teleskopering framförallt relevant i anslutning till referensperiodens gränser.

- Fenomenet teleskopering kan leda till både under- och överrapportering av händelser. Vilket är vanligast? Enligt vissa är överrapportering vanligast. Det kan bero på att minnen blir mer bristfälliga med tiden och är tydligare för mer nutida än för senare händelser.
- I longitudinella undersökningar som har flera mätpunkter mellan insamlingstillfällena kan det uppstå så kallade sömeffekter ("seam effects" på engelska): orimliga skillnader i "sömmen" mellan insamlingstillfällena.
- En förklaring kan vara att respondenterna utgår från den senaste månaden eller gör en summering och låter det svaret gälla för alla månader inom intervallet (så kallad "constant-wave response"). Om respondenten använder denna strategi blir förändringar inom intervallet underrapporterade och hamnar i sömmen till förra intervallet.
- För att minska sömeffekten kan det vara en god idé att gruppera frågorna kronologiskt snarare än tematiskt.
- Att ha fler mättillfällen och kortare referensperioder kan vara ett sätt att mildra effekten.

Att minnas antalet gånger något hände (frekvensbestämning):

- Svarsstrategier för frågor om *antalet gånger*:
 - Respondenten kan ha frekvensen lagrad som en egen post i minnet.
 - En annan strategi är att återerindra informationen om alla aktuella tillfällen och summera.
 - En tredje strategi är att uppskatta med hjälp av en frekvens som redan finns i minnet.
- Respondenterna kan också räkna ut en frekvens för en kortare period eller ett intervall som sedan generaliseras till en längre period.
- Respondenterna kan också göra en uppskattning utifrån svarsalternativen. Om de till exempel upplever att de gjort en aktivitet väldigt ofta kanske de väljer det högsta alternativ oavsett vilka faktiska frekvenser som ingår i det svarsalternativet.
- De kan också svara utifrån "första intrycket" utan vidare eftertanke.

- Valet av strategi styrs av frågans krav på ansträngning, respondentens motivation, informationens tillgänglighet i minnet och andra förhållanden kring uppgiften som till exempel frågeformuleringen eller vilken tid det finns för genomsökning av minnet.
- Referensperioden är inte bara en frågeteknisk avgränsning utan kan också påverka vad respondenten uppfattar som frågans utgångspunkt.

Att främja återerinring och minska minnesfel:

- Som frågekonstruktör behöver man tänka på personers inkodning av information i minnet. Om informationen inte har inkodats, kan den inte återerinnas.
- Att främja återerinring innebär att presentera s.k. ledtrådar (retrieval cues) som matchar hur minnena är kodade och organiserade.
- Ledtrådarna bör vara både korrekta och tillräckligt specifika för det avseende de har. En svårighet är dock att det kan skilja sig mellan olika personer vad som utgör bra ledtrådar.
- Det finns flera tekniker för att främja återerinring: underlättad erinring, dokumentation, kalenderintervju, uppvärmning, nedbrytning, att anknyta till rätt begrepp, långa frågor, svarsalternativens utformning och ha datainsamlingen nära inpå referensperioden.
- De flesta teknikerna har dock en baksida, vilket oftast är att de förlänger svarstiden. Därför bör de generellt endast användas i specifika fall snarare än som generell lösning.

12 Att fråga om attityder

Författare: Andreas Persson (SCB)

Attityder kan vara svåra att mäta

Attityder och åsikter tillhör de kanske mest problematiska objekten att mäta. Vi kan inte *se* dem. Till skillnad från i serietidningarnas värld kan vi inte utläsa personers tankar och åsikter i tankebubblor ovanför deras huvuden. Attityder och åsikter är, som det kallas, inte observerbara. De existerar endast inom individen. Ingen utom individen själv kan heller bekräfta svarens giltighet. Detta medför att det kan vara svårt att utvärdera den här typen av frågor. Det finns inget "sant" värde att jämföra med som det kan göra för mer objektiva frågor. Det finns också sällan datakällor att pröva svaren mot (till exempel administrativa register). Ett undantag är väljarbarometrar. Där kan valresultatet på sätt och vis ses som en validering av undersökningsresultatet (även om det finns en risk för t.ex. förändringar mellan mättillfället och valdagen). För att utvärdera den här typen av frågor får man därför använda andra tekniker. Man kan till exempel studera om flera frågor som ska mäta samma sak hänger ihop statistiskt och om begrepp förhåller sig till varandra på det sättet som teorin på området förespråkar (läs mer i kapitel 8).

På grund av dessa svårigheter har det gjorts rätt mycket forskning om just attitydfrågor (Schuman & Presser, 1981). Det finns också en bred teoribildning kring själva attitydbegreppet. Detta kapitel syftar till att redovisa både de teoretiska utgångspunkterna och de frågetekniska.

Vad är en attityd?

Vardagligt synsätt

Attityder fyller en viktig funktion i dagens samhälle. Vanligt folk pratar om dem, vissa försöker påverka dem genom opinionsbildning och opinionsinstitut försöker mäta dem. Vad är då en attityd? Enligt Svenska Akademiens ordlista över svenska språket betyder "attityd" en "inställning till något" (SAOL, 2013). Men attityder har också en teoretisk dimension, ett arbete som går tillbaka ända till 1930-talet. Vad är då det teoretiska synsättet på en attityd?

Teoretiskt synsätt

En kort definition är att en attityd är "en positiv eller negativ utvärderande reaktion till någon eller något, uttryckt i övertygelser, känslor och beteende" (Myers, 2002). Det är alltså en slags beredskap för att agera eller reagera på ett visst sätt eller i en viss riktning. En attityd är knuten till ett attitydobjekt som kan vara platser, idéer, situationer, grupper av individer eller fysiska saker. Enligt vissa forskare är människors attityder representerade i människors minnesstrukturer (Myers, 2002; Oskamp & Schultz, 2005).

Enligt en äldre syn bestod en attityd av tre komponenter ("tri-componential view" på engelska):

- känslomässig komponent – vilka känslor individen har förknippade med attitydobjektet
- beteendemässig komponent – hur personen agerar och handlar vad gäller attitydobjektet
- kognitiv komponent – övertygelser och tankar kring attitydobjektet.

Ovanstående kallas för attityders ABC eftersom det är begynnelsebokstäverna för de tre komponenterna på engelska (Affections, Behaviors och Cognitions). Modellen har stött på en hel del kritik, bland annat om det verkligen går att särskilja komponenterna från varandra och om alla attityder måste ha alla tre. Kanske kan vissa attityder vara rent kognitiva medan andra är fullständigt känslomässiga (Oskamp & Schultz, 2005)?

Enligt en modernare syn är dessa komponenter fristående, separata delar ("separate entities" på engelska) som kan vara aktiva och relaterade, beroende på vilken typ av attityd det handlar om (Oskamp & Schultz, 2005).

I en annan gren av forskningen har attitydbegreppet "avteoretiserats" över tid till den ungefärliga betydelsen "något man tycker bra eller illa om" (Schwarz & Bohner, 2001).

Sammanfattande punkter:

- "Attityd" har en vardaglig betydelse i stil med en "inställning" till något.

- Det finns också ett teoretiskt synsätt på attityder. En attityd brukar definieras som "en positiv eller negativ utvärderande reaktion till någon eller något, uttryckt i övertygelser, känslor och beteende".
- En attityd kan bestå av olika komponenter, till exempel en känslomässig, en beteendemässig och/eller en kognitiv komponent.
- Det finns även grenar som "avteoretiserat" attitydbegreppet mot dess vardagliga betydelse: något man tycker om eller illa om.

Svårigheten med attitydfrågor

Att mäta attityder är ett av de svårare områdena inom mättekniken.

Svårigheterna består bland annat av:

- Begreppsförvirring för att termen attityd används med olika betydelse i olika sammanhang.
- Studier visar att kopplingen mellan attityd och beteende inte alltid är så stark som den kan förväntas vara.
- Det kan finnas olika nivåer av en attityd, till exempel en konkret och en abstrakt nivå, vilka kan ge motsägelsefulla resultat.
- En del attitydmätningar är osäkra, varierar orimligt mycket över tid och mellan olika sammanhang.

Avsnitten nedan behandlar bland annat dessa svårigheter och hur man bör förhålla sig till dem när man formulerar attitydfrågor.

Begreppsförvirring

En svårighet med attitydfrågor har med själva terminologin att göra. Beteckningen "attitydfrågor" används både om relativt enkla, direkta mätningar där till exempel endast en rak fråga används för att mäta en "attityd" och för mer komplicerade mätningar där flera indikatorer används för att uppskatta en underliggande attityd. Denna variation kan bidra till förvirring om vad som egentligen åsyftas med "attitydfrågor". Det är viktigt att vara medveten om dessa varianter. Egentligen är "åsiktsfrågor" kanske en mer lämplig beteckning för de enklare direkta mätningarna men dessa kallas som sagt ofta för "attitydfrågor".

Kopplingen mellan attityd och beteende

En annan svårighet inom forskningen har varit att utreda relationen mellan attityd och beteende. Ibland är det ju hur folk *agerar* som är det viktigaste. Det finns dock många exempel på att relationen mellan attityd och beteende är förvånansvärt svag. Det skulle i sin tur betyda att folk inte följer sina övertygelser. Detta är naturligtvis inte helt sant. Den svaga relationen kan bero på många saker, bland annat om vi gjort en god mätning av attityden. Om vi inte lyckats mäta folks faktiska inställning utan till exempel får tillrättalagda svar är det klart att kopplingen till beteende kan vara svag. En annan sak som påverkar är hur specifikt kopplad attityden är till beteendet. En allmän attityd till hälsa, kondition eller fysisk träning är inte nödvändigtvis en bra prediktor för jogging. Den mer specifika attityden till just jogging är sannolikt betydligt bättre (Myers, 2002). Andra saker som påverkar kopplingen mellan attityd och beteende är hur viktig attityden är för oss och om vi handlar medvetet eller per automatik. Det sociala sammanhanget kan främja eller sätta bojor på vissa beteenden, oavsett vad individen svarat i en undersökning.

En svag koppling mellan attityd och beteende kan alltså bero på att attitydobjektet är vagt, att mätningen är svag, att de sociala sammanhangen skiljer sig åt eller att mätningarna av attityd och beteende ligger på olika nivåer.

Sammanfattande punkter:

- Det finns en begreppsförvirring om vad som avses med "attitydfrågor". Det kan användas både om relativt enkla direkta mätningar och om mer komplicerade mätningar där flera indikatorer används för att uppskatta en underliggande attityd.
- Kopplingen mellan attityd och beteende är ofta svag. Det kan bero på att attitydbegreppet är teoretiskt svagt, att mätningen är svag eller att mätning och beteende ligger på olika nivåer.

Abstrakt och konkret nivå av en attityd

En svårighet med attitydfrågor har alltså med begreppshierarkier att göra, som i exemplet ovan med fysisk träning och jogging. Schuman (1972) (Wikman 1990) menar att den svaga kopplingen och osäkerheten kan bero på just det – skillnader mellan abstrakt och konkret nivå av attityden, där verkligheten alltid är konkret men frågenivån i

frågeformuläret ofta är abstrakt. Han använder frågor om diskriminering på arbetsmarknaden som exempel. I Schumans (1972) studie fanns en generell fråga om diskriminering. Denna var alltså på en abstrakt nivå och 85 procent svarade att de var emot diskriminering. I en senare fråga beskrevs ett scenario där en personalchef agerade diskriminerande för att undvika motsättningar. På den frågan tyckte 39 procent att personalchefen agerade rätt. I den situationen var det alltså acceptabelt att diskriminera enligt 39 procent av respondenterna. Vad beror dessa skillnader på?

Svaren från de frågorna (på en abstrakt respektive konkret nivå) sammanfaller inte som förväntat. Tvärtom skiljer de sig så pass mycket åt att det kan ifrågasättas om de mäter samma sak. Sannolikt finns de största skillnaderna mellan frågorna just i skillnaden mellan abstrakt och konkret. Det abstrakta är inte alltid giltigt för individen i det konkreta fallet där andra relevanta faktorer kan vara inblandade. I konfliktsituationen kan individen alltså göra en kompromiss på den abstrakta attitydens bekostnad. Schumans frågeexempel är föråldrat men konflikten mellan abstrakt och konkret nivå är fortfarande lika aktuell (Wikman, 1990). Personer kan alltså ha en värdering på en abstrakt nivå, men när den exemplifieras med en konkret situation kan den hamna i konflikt med andra värderingar. Då måste individen förhandla om och kompromissa med sina värderingar. Resultatet kan tyckas motsägelsefullt men det är så verkligheten ofta ser ut – konflikt och kompromisser. Att det finns skillnader mellan undersökningar (attityder) och verkligheten (beteende) kan därför bero på att frågorna i undersökningen ofta är på en abstrakt nivå medan verkligheten är konkret och ibland kräver kompromisser. Utmaningen för frågekonstruktörer blir då hur verklighetens konfliktsituationer ska kunna matchas i en undersökning och hur vi samtidigt kan kontrollera vad det faktiskt är vi mäter (Schuman, 1972; Wikman, 1990).

Sammanfattande punkter:

- Attitydfrågors "osäkerhet" kan bero på nivån i frågorna – är den abstrakt eller konkret?
- Individer kan ha en attityd på en abstrakt nivå men den är inte alltid giltig i det konkreta fallet där andra faktorer och attityder också kan vara relevanta. I det konkreta fallet kan det finnas konflikter som kräver kompromisser.

- Verkligheten är konkret. Konflikten och kompromisser mellan abstrakt och konkret nivå kan därför vara en förklaring till exempelvis skillnader mellan attityd och beteende.

Abstrakt och konkret nivå – konsekvenser för frågekonstruktion

Respondenters svar kan alltså skilja sig åt beroende på om attityden i frågan är på en abstrakt eller konkret nivå. Ibland varnas det generellt för att ställa frågor på en för abstrakt nivå, till exempel "Bör FN i första hand kämpa för frihet i alla länder?". Det är då svårt att veta vad respondenten fokuserar på och inkluderar i sina svar – det finns så mycket möjligt innehåll i de breda begreppen (Wikman, 1990). Denna brist är särskilt bekymmersam om attitydobjektet är något som respondenterna kan antas ha liten kunskap om. Om man inte vet vad svaren uttrycker begränsas förstås användbarheten av resultatet.

En lösning på att frågorna är för breda är förstås att smalna av dem för att göra dem mer konkreta (läs även mer i kapitel 8). Det finns dock en del saker man bör ta hänsyn till om man ska konkretisera frågeinnehållet.

Frågans täckningsgrad

Ett sätt att göra frågor mer konkreta är att ge frågan mer specifika utgångspunkter och tillföra beskrivningar. I samma skede uppkommer en annan utmaning. Vilka utgångspunkter ska vi ha i frågan och vilka beskrivningar ska tillföras? Det kan ha väldigt stor påverkan.

Alla tre frågor nedan syftar till att mäta inställningen till privata sjukhus, där andelen positiva svar står i parentes efter frågan (det bör påpekas att resultaten nedan inte är aktuella utan från mätningar på 1990-talet, men som principiellt exempel är de fortfarande giltiga):

1. Är det bra eller dåligt om du får möjlighet att välja mellan vård på ett privat sjukhus och vård på ett sjukhus som drivs av landstinget, förutsatt att det inte kostar mer på det privata sjukhuset? (Bra, 81 %)
2. Tycker du det vore bra eller inte bra om också privata företag kunde äga och driva sjukhus om kostnaden för patienten vore densamma? (Bra, 56 %)
3. Tycker du att privatägda företag ska få driva sjukhus med vinst, som kan tas ut ur verksamheten? (Ja, 32 %)

Alla frågor ovan innehåller specificeringar, om än i olika grad. De innehåller dock ganska olika typer av specificeringar med olika utgångspunkter och innehåll. Den första frågeformuleringen innehåller "möjligheten att välja" och "att det inte kostar mer på det privata sjukhuset". Båda specificeringarna upplevs förmodligen som positiva av många respondenter. Den andra frågeformuleringen påminner om den första men säger inget om "möjligheten att välja". Den tredje frågeformuleringen innehåller ingen av de två positiva specificeringarna utan fokuserar bara på "... vinst, som kan tas direkt ur verksamheten". Slutsatsen är att alla tre frågor har specificeringar men väldigt olika sådana.

Föga förvånande skilde sig svaren åt avsevärt på dessa tre frågor, i linje med om specificeringarna var positiva eller negativa. Ovanstående frågor är vinklade men illustrerar likväl svårigheten med vilka specificeringar frågan ska innehålla för att bli mer konkret. Verkligheten är komplex. Många sakfrågor innehåller flera aspekter, både fördelar och nackdelar. Alla specificeringarna i frågorna ovan kan mycket väl vara giltiga i sig för att mäta inställningen till privata sjukhus, men om frågan bara inkluderar vissa av dem kan den ge helt andra svar än om andra specificeringar hade ingått.

Ett annat exempel på samma svårighet kommer från mätningar om tv-reklam i slutet på 1980-talet (Bohm, 1988; Wikman, 1990).

EXEMPEL:

Inledning: Vad tycker Ni om att ha reklam i svensk TV – mellan programmen, sammanlagt 15 minuter per kväll i varje kanal. TV skulle då få reklaminkomster som gör att de får mer pengar till sin verksamhet.

Fråga 1: Är Ni för eller emot att man betalar en del av TVs verksamhet genom att på det här sättet införa reklam i svensk TV? (För: 80 %, Emot: 14 %)

Fråga 2: Men det här förväntas också leda till att företagen lägger ner mer pengar på reklam. Då skulle priserna på varorna höjas ute i handeln. Det skulle alltså bli mer pengar till TV-verksamheten men dyrare varor. Är Ni för eller emot att på det här sättet införa reklam i svensk TV? (För: 47 %, Emot: 43 %)

Den andra frågan tillför en specificering. I detta fall är det en nackdel med attitydobjektet och det får konsekvenser på resultatet. När vi tillför

specificeringar ger vi samtidigt frågan en *inriktning* – vi väljer att inkludera ett visst innehåll snarare än ett annat. Därmed har vi påverkat frågans täckningsgrad. Det är viktigt att inriktningen är i linje med det vi planerat att mäta. I praktiken är detta svårt, speciellt om man ska mäta något komplicerat med många olika aspekter, till exempel för- och nackdelar. Det är ofta omöjligt att mäta alla inriktningar av en sakfråga i endast en fråga, utan det kan krävas flera frågor för att täcka in allt. Detta, att använda flera frågor (eller items som det också brukar kallas), är ett vanligt tillvägagångssätt i praktiken för att mäta attityder mer på djupet.

Sammanfattande punkter:

- Ett sätt att konkretisera en fråga är att tillföra specificeringar.
- Ett dilemma är att välja vilka specificeringar som ska ingå i frågan eftersom det påverkar frågans täckningsgrad och därmed också svaren på frågan.
- För komplicerade frågor kan inte alla aspekter täckas av en enskild fråga. Om frågan bara tar upp vissa aspekter har den låg täckningsgrad. Det kan styra svaren i och med att specificeringarna ger frågan *en inriktning*.

Specificeringar och förståelse

Specificeringarna hittills i avsnittet har utgjorts av tydliga fördelar och nackdelar. Men även till synes neutrala specificeringar kan påverka resultatet. Bradburn & Danis (1984) diskuterar skillnaderna mellan två frågor som båda syftar till att mäta opinionen till Koreakriget.

1. Tycker du att USA gjorde ett misstag eller ej, genom att besluta sig att försvara Korea?
2. Tycker du att det var rätt eller fel av USA att sända amerikanska trupper för att stoppa kommunistinvasionen i Sydkorea?

Den andra frågan mätte en betydligt mer positiv opinion. Det kan finnas flera skäl till det. Det som brukar anges i litteraturen är att den andra frågan nämner "kommunismen" och då aktiverar respondenternas tankar kring detta. Frågorna ställdes under en tidsperiod när det fanns en stark och utbredd rädsla för kommunismen. Att nämna detta i frågeformuleringen kan därför ge mer positiva svar för den frågeformuleringen än för den andra. Om detta stämmer kan det ifrågasättas vad som

egentligen mäts i fråga 2. Frågan syftade till att mäta inställningen till inblandningen i Korea, *inte* till kommunismen. Dessa två saker – Koreakriget och kommunismen – är förmodligen relaterade till varandra, men de är definitivt inte samma sak. Detta exempel visar hur olika specificeringar och beskrivningar kan aktivera olika tankar hos respondenterna, vilket kan leda till att frågan mäter olika saker (Bradburn & Danis, 1984; Wikman, 1990).

Specificeringarna tillför alltså en inriktning och en avgränsning i respondentens tänkande kring frågan. Detta är inte särskilt förvånande eftersom det är specificeringens egentliga syfte – att avgränsa eller inringa något abstrakt. Men för att få en bra mätning krävs det förstås att specificeringarna leder till en korrekt avgränsning för respondenten, givet det som ska mätas.

Specificeringar – sammanfattande slutsatser

Specificeringar kan leda till framförallt två problem:

- Specificeringarna är "felaktiga" i den betydelsen att de leder respondenten till att tänka på och svara för ett annat attitydobjekt än det vi tänkt oss (till exempel att respondenten svarar för kommunismen istället för Koreakriget).
- Specificeringarna är giltiga i sig men styr frågans inriktning så att endast vissa delar, snarare än andra, av attitydobjektet täcks (till exempel frågorna om privat vård i förra avsnittet).

Ett skydd mot båda typerna av problem är att ställa flera frågor, med olika specificeringar, för samma attitydobjekt. Då blir eventuella fel i enskilda frågor inte heller lika utslagsgivande.

Sammanfattande punkter:

- Även till synes neutrala specificeringar kan inrikta respondentens förståelse på oönskade sätt och därmed påverka svaren.
- Flera frågor med olika specificeringar ökar täckningsgraden och ger en mer rättvis mätning.

Hur går det till när attitydfrågor besvaras?

Det finns många exempel på att svaren på attitydfrågor kan variera beroende på sammanhang. De tydligaste exemplen på detta är så kallade kontexteffekter (läs mer i kapitel 13). Det betyder att svaren på en fråga beror på kontexten, till exempel den föregående frågan. Vissa studier har även funnit att resultat på attitydfrågor kan variera mycket över tid och det är främst de som är osäkra som ändrar sig (Wikman, 1990).

Tourangeau, Rips och Rasinski (2000) beskriver hur tänkande i sig, att respondenterna uppmanas att begrunda sitt ställningstagande, kan påverka deras svar. Detta kan ofta ske i en polariserande riktning, vilket då leder till en slags självgenererad attitydförändring. Det finns även studier som visar att en tredjedel av respondenterna anger åsikter om påhittade ämnen eller ämnen som de rimligtvis inte kan ha en åsikt om (Schuman & Presser, 1981). Den här typen av resultat väcker frågan om hur det egentligen går till när attitydfrågor besvaras. Hur ser svarsprocessen ut?

En traditionell modell

Enligt en traditionell syn finns attityder färdigutvecklade och lagrade i minnet. Enligt denna syn är attityderna arkiverade och redo att plockas fram vid behov. När respondenten får en formulärfråga söker han eller hon alltså i minnet tills attityden påträffas, hämtar svaret och rapporterar det sedan i frågeformuläret. Modellen förutsätter att attityder existerar på förhand, innan frågan ställs, och är stabila (Tourangeau, Rips & Rasinski, 2000). Som redovisas ovan verkar dock dessa förutsättningar ofta inte stämma. Många attitydmätningar ger osäkra resultat. Ibland kan modellen säkert vara korrekt på ett generellt plan, även om attitydfrågor alltid är kontextberoende i någon grad. Vissa attityder existerar på förhand och är relativt stabila över tid. Om vissa sakfrågor har vi tydliga och fasta åsikter och attityden finns hyfsat stabil inom oss. Ofta innefattar sådana "säkra åsikter" sådant som är relevant för oss och ligger oss nära, till exempel frågor om familjeliv, intresseområden och vårt arbete. I många andra fall verkar dock svaren på attitydfrågor inte nödvändigtvis finnas på förhand utan snarare konstrueras vid svarstillfället (Tourangeau, Rips & Rasinski, 2000).

Sammanfattande punkter:

- Enligt en traditionell syn finns attityder färdiga att hämtas i minnet.
- Forskning som visar attitydfrågors "osäkerhet" ifrågasätter den traditionella synen.
- Ibland finns säkert åsikter färdiga i minnet, till exempel om ämnen vi är intresserade av. Men för många ämnen har vi nog inte färdiga åsikter utan åsikten konstrueras vid svarstillfället.

En annan modell

Enligt en annan modell konstrueras ibland attityder vid svarstillfället. I denna process kan respondenten använda olika tillvägagångssätt. Tourangeau, Rips och Rasinski (2000) skiljer till exempel på två olika tillvägagångssätt när attityder skapas vid svarstillfället: att utgå från 1) generella attityder eller 2) specifika övertygelser.

Dessa påminner om top-down- eller bottom-up-strategier (eller induktion och deduktion), alltså att man antingen applicerar en generell regel på ett enskilt fall (top-down) eller skapar en generell regel utifrån ett specifikt fall (bottom-up). Översatt till konstruktion av attityder betyder detta att i det ena fallet (1) använder respondenten en attityd på en mer generell nivå och applicerar den på den specifika frågan. I det andra fallet (2) utgår respondenten från något specifikt, till exempel en enskild händelse, för att besvara något på en mer generell nivå.

Sannolikt är det så att attityder som konstrueras vid svarstillfället är mer känsliga för faktorer som gäller just då men som inte nödvändigtvis gäller för ett annat svarstillfälle. Sådana faktorer kan vara frågeformuleringens ordalydelse, frågans placering i frågeformuläret eller undersökningens ämne. Det kan också vara rena tillfälligheter som ligger bortom undersökningen (till exempel en just då aktuell nyhets-händelse eller om respondenten är på gott humör). Det går inte att kontrollera för allt detta men man bör ha övriga faktorer i åtanke när man skriver attitydfrågor och framförallt allt när man återanvänder frågor från andra frågeformulär. När man frågar om saker som respondenterna kanske inte har en given åsikt om, det vill säga där attityden kan komma att konstrueras vid svarstillfället, bör man därför också vara extra noggrann med frågeformuleringarna och testa frågorna innan de tas i bruk (Tourangeau, Rips & Rasinski, 2000).

Sammanfattande punkter:

- Enligt ett synsätt kan attityder skapas vid svarstillfället genom att 1) antingen utgå från en generell attityd och applicera den på den specifika frågan (top-down eller deduktion) eller tvärtom, 2) genom att forma en attityd baserat på en specifik händelse (bottom-up eller induktion).
- Attityder som konstrueras vid svarstillfället borde vara känsliga för sammanhanget som råder just när frågorna besvaras. Det kan vara annorlunda vid ett annat svarstillfälle.

Att konstruera attitydfrågor

Nedan följer en del riktlinjer för konstruktion av attitydfrågor, uppdelat i frågeformulering och svarsalternativ.

Frågeformuleringen

Definiera attitydobjektet

Det första man måste göra är att definiera attitydobjektet noggrant (Bradburn, Sudman & Wansink, 2004). Vad karaktäriserar attitydobjektet? Hur särskiljs det från och överlappar med angränsande attitydobjekt (både horisontellt och vertikalt)? Alltså skapa en slags begreppskarta. Har man detta klart definierat är det mycket enklare att skriva frågor och riskerna för att blanda in andra attitydobjekt i frågeformuleringen än det man är intresserad av blir betydligt mindre. Detta är ett slags konceptualisering och beskrivs mer utförligt i kapitel 8.

Fundera på vilka uttryck attityden kan ta sig både känslomässigt, beteendemässigt och kognitivt (Bradburn, Sudman & Wansink, 2004), eftersom det också kan vara underlag till frågekonstruktionen. Tänk dock på att beteende och attityd inte alltid sammanfaller.

Välj konkretionsnivå

Var uppmärksam på konkretionsnivån, hur generella eller specifika frågorna är, och hur det förhåller sig till syftet med mätningen. Personer kan ha en abstrakt attityd men som inte ger utslag i det konkreta fallet. Frågor som är alltför abstrakta i ansatsen kan däremot ge svårtolkade svar eftersom tolkningsrymden kan vara för stor. När frågor konkretiseras är det viktigt att bedöma att det definierade attitydobjektet

fortfarande är i fokus i frågan och att annat helst inte inkluderas. Konkretisering av frågan kan också leda till obalans – att bara vissa delar av ett attitydobjekt innefattas. Därför är ofta den mest lämpliga lösningen att ställa flera frågor som var och en för sig belyser olika delar av attitydobjektet och som slås ihop vid analysen. Det ger en mer rättvis bild och man undviker riskerna med att dra slutsatser baserat på endast en fråga (Wikman, 1990).

Var uppmärksam på respondenternas kunskap

Var försiktig med att ställa frågor om sådant som respondenterna inte känner till. Som tidigare nämnts kan vissa respondenter svara på frågor även om påhittade ämnen. Det ska inte tolkas som att respondenterna svarar "hur som helst" utan kan vara ett uttryck för att de gör sitt bästa och till exempel bygger sitt svar på någon mer generell attityd än det påhittade ämnet (Schwarz, 1996). Att ställa frågor om ämnen som respondenterna inte känner till påverkar förstås mätningen negativt. Det bör undvikas men det kan vara svårt att på förhand veta hur mycket respondenterna känner till och det kan förstås också skilja sig åt mellan olika respondenter. För att få information om det sistnämnda kan man till exempel erbjuda "Ingen åsikt" som svarsalternativ (läs mer i kapitel 3) eller lägga till frågor om hur säkra eller osäkra respondenterna är på sitt svar eller hur insatta de är i ämnet. Sådana kunskapsfrågor ökar förstås omfånget på frågeformuläret och kan därför oftast inte användas som generell lösning.

Sammanfattande punkter:

- Definiera attitydobjektet. Detta är viktigt och styr vilka enskilda frågor som ska ställas.
- Tänk på konkretionsnivån och att flera frågor, med olika specificeringar, ofta behövs för att täcka attitydobjektet.
- Var försiktig med att fråga om sådant som respondenterna inte känner till. Det går att mäta deras kunskap, säkerhet eller intresse för attitydobjektet med separata frågor.

Svarsalternativ

Svarsalternativ har sitt eget kapitel i denna bok. Attitydfrågor är dock troligen den frågetyp som uppvisar flest typer av svarsalternativ, alltifrån VAS-skolor till glada och sura ansikten (där respondenten ska markera vilket ansikte som bäst beskriver deras inställning till något). Därför tas några av de kändaste exemplen upp även i det här kapitlet. De har med tiden använts även för att mäta andra subjektiva egenskaper och tillstånd än just attityder. Nedan ges en kort beskrivning av några attitydtypiska svarsskolor (läs mer i kapitel 3). För mer detaljerade beskrivningar av svarsskalorna nedan hänvisas till annan litteratur (till exempel Oskamp & Schultz, 2005).

Likert: Likertskalan, som fått sitt namn efter Rensis Likert, är en summationsskala där resultatet från flera frågor (eller så kallade items) slås ihop. När Likertskalan lanserades var en nyhet att respondenterna inte skulle svara "Ja" eller "Nej" utan till vilken grad de instämde i eller höll med om varje delfråga. Frågorna ställs som påståenden (Trost, 1991; Oskamp & Schultz, 2005).

EXEMPEL:

	Stämmer mycket bra	Stämmer ganska bra	Stämmer ganska dåligt	Stämmer mycket dåligt	Vet inte
På min arbetsplats respekterar vi varandra	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Enligt Likertmetoden ska frågorna för-testas så att de verkligen ger utslag mellan dem som har olika nivåer av attityden. De frågor som inte ger tillräckligt stor spridning ska tas bort. Likertskalan har fått stort genomslag och används inom många fler områden än attitydmätningar. Den har dock fått en del kritik, bland annat för risken för hålla-med-tendenser (läs mer i kapitel 3).

Guttman: En nackdel med Likertskalan är att summan i summeringen inte har någon unik mening. En respondent kan få samma poängsumma på många sätt (till exempel genom att svara mittenvärdet 3 på alla frågor eller extremvärdena 1 och 5 på varannan fråga). Guttmans skala råder bot på detta eftersom den är kumulativ. Den innehåller ja-och-nej-frågor som är utprovade så att de mäter samma underliggande objekt och växer

i styrka, fråga för fråga. Summan ger därför information om attitydens styrka (Oskamp & Schultz, 2005).

Semantisk-differential: Osgoods semantiska differentialskala är speciell på så sätt att den, på sätt och vis, har många olika typer av svarsalternativ och ingen riktig frågeformulering. Den innehåller en "huvudfråga" vilken utgörs av attitydobjektet. "Huvudfrågan" är dock inte formulerad som en fråga utan är en beskrivning av attitydobjektet. Svarsalternativen består av motsattpar av adjektiv och respondenten ska markera sitt svar någonstans mellan dessa motsatser.

Adjektiven innehåller olika dimensioner. En utvärderande dimension (till exempel bra-dåligt), en intensitetsdimension (till exempel stark-svag) och en aktivitetsdimension (till exempel aktiv-passiv). Den utvärderande dimensionen anses vara viktigast för att mäta attityden (Oskamp & Schultz, 2005; Trost, 1991).

Sammanfattande punkter:

- Det finns ett antal svarsskalor som är tätt förknippade och utvecklade för attitydmätningar, till exempel Likerts, Guttmans och Semantisk-differential skala.

Exempel på hur innehållet i detta kapitel kan tillämpas i praktiken finns i Wikman (1990) ("Att studera attityder och opinioner", kapitel i "Att fråga" av Wärneryd m.fl.). Exemplet är nästan fyrtio år gamla men de frågetekniska principerna är fortfarande aktuella.

Sammanfattning

Att fråga om attityder – sammanfattning

Vad är en attityd?

- "Attityd" har en vardaglig betydelse i stil med en "inställning" till något.
- Det finns också ett teoretiskt synsätt på attityder. En attityd brukar definieras som "en positiv eller negativ utvärderande reaktion till någon eller något, uttryckt i övertygelser, känslor och beteende".

- En attityd kan bestå av olika komponenter, till exempel en känslomässig, en beteendemässig och/eller en kognitiv komponent.
- Det finns även grenar som "avteoretiserat" attitydbegreppet mot dess vardagliga betydelse: något man tycker om eller illa om.

Svårigheten med attitydfrågor:

- Det finns en begreppsförvirring om vad som avses med "attitydfrågor". Det kan användas både om relativt enkla direkta mätningar och om mer komplicerade mätningar där flera indikatorer används för att uppskatta en underliggande attityd.
- Kopplingen mellan attityd och beteende är ofta svag. Det kan bero på att attitydbegreppet är teoretiskt svagt, att mätningen är svag eller att mätning och beteende ligger på olika nivåer.
- Attitydfrågors "osäkerhet" kan bero på nivån i frågorna – är den abstrakt eller konkret?
- Individer kan ha en attityd på en abstrakt nivå men den är inte alltid giltig i det konkreta fallet där andra faktorer och attityder också kan vara relevanta. I det konkreta fallet kan det finnas konflikter som kräver kompromisser.
- Verkligheten är konkret. Konflikten och kompromisser mellan abstrakt och konkret nivå kan därför vara en förklaring till exempelvis skillnader mellan attityd och beteende.
- Ett sätt att konkretisera en fråga är att tillföra specificeringar.
- Ett dilemma är att välja vilka specificeringar som ska ingå i frågan eftersom det påverkar frågans täckningsgrad och därmed också svaren på frågan.
- För komplicerade frågor kan inte alla aspekter täckas av en enskild fråga. Om frågan bara tar upp vissa aspekter har den låg täckningsgrad. Det kan styra svaren i och med att specificeringarna ger frågan *en inriktning*.
- Även till synes neutrala specificeringar kan inrikta respondentens förståelse på oönskade sätt och därmed påverka svaren.
- Flera frågor med olika specificeringar ökar täckningsgraden och ger en mer rättvis mätning.

Hur går det till när attitydfrågor besvaras?

- Enligt en traditionell syn finns attityder färdiga att hämtas i minnet.
- Forskning som visar attitydfrågors "osäkerhet" ifrågasätter den traditionella synen.
- Ibland finns säkert åsikter färdiga i minnet, till exempel om ämnen vi är intresserade av. Men för många ämnen har vi nog inte färdiga åsikter utan åsikten konstrueras vid svarstillfället.
- Enligt ett synsätt kan attityder skapas vid svarstillfället genom att 1) antingen utgå från en generell attityd och applicera den på den specifika frågan (top-down eller deduktion) eller tvärtom, 2) genom att forma en attityd baserat på en specifik händelse (bottom-up eller induktion).
- Attityder som konstrueras vid svarstillfället borde vara känsliga för sammanhanget som råder just när frågorna besvaras. Det kan vara annorlunda vid ett annat svarstillfälle.

Att konstruera attitydfrågor:

- Definiera attitydobjektet. Detta är viktigt och styr vilka enskilda frågor som ska ställas.
- Tänk på konkretionsnivån och att flera frågor, med olika specificeringar, ofta behövs för att täcka attitydobjektet.
- Var försiktig med att fråga om sådant som respondenterna inte känner till. Det går att mäta deras kunskap, säkerhet eller intresse för attitydobjektet med separata frågor.
- Det finns ett antal svarsskalor som är tätt förknippade och utvecklade för attitydmätningar, till exempel Likerts, Guttman och Semantisk-differential skala.

13 Ordningsföljder och kontext

Författare: Andreas Persson (SCB)

Ordningsföljder och kontext

En schamporeklam betonar att man aldrig får en andra chans att göra ett första intryck. I partiledardebatter på tv lottas i vilken ordning partiledarna får tala direkt till tittarna. I media rapporteras entusiastiskt vilket startnummer det svenska bidraget fått i den europeiska melodifestivalen. Har ordningen verkligen någon betydelse? Det sunda förnuftet kan hävda att vi människor borde vara kapabla att bedöma ett antal objekt individuellt eller att vi kan göra en rättvis jämförande bedömning oavsett i vilken ordning objekten presenteras. Forskning inom den kognitiva psykologin och psykofysiken hävdar det motsatta – att ordningsföljden kan ha betydelse.

Den första informationsbiten är unik

På ett övergripande plan kan ordningsföljdens betydelse bero på att den första informationsbiten har en unik position i en informationsmängd. När den första informationsbiten når oss är vi, på sätt och vis, som ett tomt blad. Vi är opåverkade. När nästa informationsbit kommer är situationen helt annorlunda. Den första informationsbiten kan då utgöra en referenspunkt gentemot vilken den andra bedöms. Tankar väckta och aktiverade av den första informationsbiten kan påverka hur den andra uppfattas. Det finns också mindre tid för eftertanke ju fler alternativ individen har att ta ställning till. Situationen är alltså annorlunda beroende på *när* i en ordningsföljd som en informationsbit presenteras.

Ordningsföljder och surveyundersökningar

Hur är detta relevant för surveyundersökningar? Det är lätt hänt att som frågekonstruktör se på frågor som unika och självständiga, att respondenterna behandlar alla frågor som om de vore den första informationsbiten beskriven ovan. Mycket talar dock för att så inte alltid är fallet. Ordningsföljden kan ha betydelse för svaren på frågorna – att svaren på frågorna blir annorlunda jämfört med om de skulle ha haft en

annan ordningsföljd. Detta kapitel fokuserar på två typer av sådana ordningseffekter:

- Kontexteffekter: En samlingsterm för effekter som främst har att göra med *ordningsföljden mellan eller inom frågor*. Kontexten kan dock vara bredare än frågeformuläret och till exempel påverkas av hur undersökningen presenteras som helhet.
- Positionseffekter: En samlingsterm för effekter som har att göra med *svarsalternativens ordningsföljd*.

Båda dessa effekter diskuteras i olika avsnitt nedan.

Sammanfattande punkter:

- Informationsbitars ordningsföljd kan ha betydelse.
- Den första informationsbiten är unik och fungerar som referens för senare information.
- Även i surveyundersökningar kan ordningsföljd och tidigare etablerat sammanhang påverka svaren på enskilda frågor. Kontexteffekter och positionseffekter är två exempel.

Frågors ordningsföljd

Kontexteffekter på grund av frågornas ordningsföljd

Ordningen på frågorna kan påverka svaren. Detta kallas för en kontexteffekt med en övergripande term. Kontexteffekter kan ställa till med problem i undersökningar. Det gäller kanske framförallt i longitudinella undersökningar där innehållet i frågeformuläret kan skifta mellan insamlingarna och eventuella skillnader på grund av ordningsföljd kan då tolkas som faktiska förändringar. På ett år, mellan 1988 och 1989, verkade antalet amerikaner som felaktigt trodde att bloddonation kan leda till aids öka kraftigt (från 28,9 % till 43,5 %). Vad berodde detta på? En vilseledande ryktesspridning? Eller kanske en kontexteffekt? 1988 kom frågan om bloddonation efter en fråga om blodtransfusion. 1989 roterades frågorna så hälften fick frågan om donation före frågan om transfusion. Kunde det vara skälet till ökningen? Detta undersöktes i ett experiment som varierade ordningsföljden vilket gav ett liknande resultat – tio procentenheter fler trodde att donation kan orsaka aids om frågan ställdes före frågan om transfusion jämfört med efter. Effekten av ordningsföljden beror nog i det här fallet på att "donation" i sig kan vara

otydligt och kan blandas ihop med "transfusion". Men när "transfusion" föregår "donation" så försvinner den risken (Colasanto, Singer & Rogers, 1992). Det här är ett tydligt exempel på att ordningsföljden kan påverka vissa respondenters tolkning.

Ovanstående är ett exempel på att det kan vara vanskligt att byta ordningsföljd på relaterade frågor.

Ordningsföljd: enhet-helhet

En del av forskningen om kontexteffekter har fokuserat på hur ordningsföljden av två frågor påverkar svaren på frågorna. Förhållandet mellan de två frågorna kan påverka svaren på olika sätt. Flera studier har undersökt hur svaren påverkas om man först ställer en specifik fråga och sedan en generell fråga inom samma ämnesområde. Detta förhållande kallas för enhet-helhet ("part-whole" på engelska). I en studie ställdes en specifik och en generell fråga om inställningen till abort. För den specifika frågan hade det ingen betydelse om den ställdes först eller sist men för den generella frågan var det drygt tio procentenheter fler som svarade positivt om den specifika frågan ställdes först. Författarna hävdar att kontexteffekter är särskilt sannolika att inträffa när något komplext sammanfattas i en generell fråga (Schuman, Presser & Ludwig, 1981).

Riktningen på kontexteffekten kan dock variera. Svaren på den senare frågan kan gå i linje med den tidigare frågan. Detta kallas för en assimileringseffekt. Men även det motsatta kan inträffa, nämligen att svaren på den senare frågan avviker på något sätt från svaren på den tidigare frågan. Detta kallas för en kontrasteffekt (se Schuman & Presser, 1981; Schwarz, Strack & Mai, 1991 för en översikt).

Detta kan tyckas förvirrande vid första anblicken men om man förstår orsakerna till effekterna är det inte så svårt. Båda effekterna är ett resultat av att individens bedömning av den andra frågan påverkas av den första. Assimileringseffekten är ett resultat av att respondenterna *inkluderar* information från den första frågan även på den andra frågan. Kontrasteffekten är tvärtom ett resultat av att respondenten *exkluderar* information från den första frågan när de ska svara på den andra frågan. Förklaringarna om varför respondenten inkluderar ("priming" och "carry over") eller exkluderar (subtraktion) diskuteras under de följande rubrikerna (Sudman, Bradburn & Schwarz, 1992).

Sammanfattande punkter:

- Ordningsföljden på frågor kan påverka svaren. Detta kallas för en kontexteffekt.
- En variant är enhet-helhet: en specifik fråga och en generell fråga som också omfattar det specifika.
- Svaret på den generella frågan blir annorlunda om den specifika frågan ställs först.
- Det finns olika typer av effekter (jämfört med om den specifika frågan inte föregått den generella): 1) att respondenterna svarar i linje med den specifika frågan när de svarar på den generella frågan (assimileringsseffekt) eller 2) att den specifika frågan får respondenterna att tänka och tolka om den generella frågan i avvikande riktning (kontrasteffekt).

Förklaringar – Priming och ”carry-over”

När kan vi förvänta oss en assimileringseffekt respektive en kontrasteffekt? Assimileringsseffekten kan bero på så kallad priming, att tankar och känslor väckta i den föregående frågan färgar bedömningen i den efterkommande frågan. Många frågor i surveundersökningar kan ha ganska breda betydelser. De kan tolkas på olika sätt och inkludera många olika saker. Det finns alltså mycket information som skulle kunna vara aktuell i till exempel en generell bedömning. Därför kan tankar väckta i en föregående fråga få stort utslag jämfört med om denna fråga inte ställts (Schwarz, Strack & Mai, 1991). Det kan då leda till en kontexteffekt – att svaren på enskilda frågor påverkas av föregående frågor. Effekten, att aktiverad information ”spiller över” på efterkommande frågor, brukar också kallas för en ”carry-over” mekanism. Ett exempel är att personer bedömer politiker i allmänhet som mindre pålitliga om de först fått ta ställning till en enskild politiker som anses opålitlig, som till exempel Richard Nixon (huvudperson i Watergateskandalen). Tankar aktiverade om Nixon ”spiller över” på den allmänna frågan om politiker (Schwarz & Bell, 1992).

Det finns dock en del faktorer som borde påverka effekten. Ett exempel är hur länge informationen förblir aktiverad, det vill säga hur länge primingen ”håller i sig”. Om det ligger många frågor emellan kontextfråga och målfråga (den fråga som påverkas av kontextfrågan), kanske de kan fungera som en buffert och den aktiverade informationen

hinner blekna. En annan faktor som kan påverka är i vilken grad kontextfråga och målfråga är relaterade. Eventuella effekter av priming borde bli större ju mer relaterade de är. Vid attitydfrågor borde även respondentens attityd i sig kunna påverka effekten, till exempel om den är enhetlig eller splittrad. Carry-over-mekanismen borde vara starkast för dem som har en motstridig inställning till frågeområdet. För sådana respondenter kan nämligen den aktiverade informationen vara icke-representativ för inställningen som helhet och ge annorlunda svar än om informationen inte varit aktiverad. En studie fann att kontexteffekten var starkast för just dem som hade starka men motstridiga inställningar till frågeområdet och att resultatet är att buffertfrågor var svaga (Tourangeau, Rasinski, Bradburn & D'Andrade, 1989).

Denna typ av effekter är beroende av respondenternas arbetsminne – att information aktiverad vid en fråga förblir aktiverad vid en senare fråga. Därför borde olikheter i arbetsminnets kapacitet också påverka effektens storlek. Ålder är en faktor i det sammanhanget. Generellt sett har äldre arbetsminne med lägre kapacitet än yngre och det kan alltså påverka effekten av kontext (Knäuper, Schwarz, Park & Fritsch, 2007). Detta är bra att känna till om man analyserar skillnader mellan åldersgrupper på kontextkänsliga frågor.

Sammanfattande punkter:

- Assimilerings effekter kan bero på så kallad priming – att tankar och känslor aktiverade vid en fråga "spiller över" på efterföljande frågor.
- Efterföljande frågor uppfattas då annorlunda än om inte föregående fråga hade ställts, och i linje med den föregående frågan.
- Effekten är beroende av arbetsminnet och är därför till exempel svagare hos äldre respondenter.

Subtraktionsförklaringar

Hur ska då kontrasteffekten – att kontextfrågan påverkar målfrågan i avvikande riktning – förstås? En teori utgår från konversationsregler och normer (läs mer i kapitel 7). Sådana normer hävdar bland annat att en talare inte ska fråga om onödig information som redan är känd. Konsekvensen i en survey kan då bli att respondenten uppfattar frågan

som att efterfråga ny information som inte ingått i tidigare frågor, annars borde inte frågan ha ställts. Detta kan betyda att respondenten uppfattar en generell fråga om nöjdhet med livet som att exkludera nöjdhet med äktenskap om de redan svarat för äktenskapet i en föregående fråga. I praktiken tolkar de då frågan som om den varit formulerad "Om du bortser från ditt äktenskap, hur nöjd är du med livet?". Detta skulle då kunna leda till en kontrasteffekt, beroende på vilka andra faktorer respondenten tar med i bedömningen förutom äktenskapet, och om respondenten är nöjd eller missnöjd med dessa. Denna teori har fått stöd i forskningen och kallas subtraktion (Schwarz, Strack & Mai, 1991).

En studie undersökte samma sak fast med öppna frågor. Svaren för de öppna frågorna gav stöd för att subtraktion kan vara en förklaring till kontrasteffekten (Mason, Carlson & Tourangeau, 1994).

Kontrasteffekter kan alltså bero på att respondenterna, i linje med konversationsreglerna, exkluderar saker som de redan svarat på i en tidigare fråga. Detta kallas för subtraktion.

Sammanfattande punkter:

- Kontrasteffekt kan bero på subtraktion – att respondenten inte tar med saker i den generella frågan som de redan svarat på i den specifika frågan.
- Subtraktion kan grunda sig i konversationsreglerna – man ska inte fråga efter eller ge redan känd information (sådant som redan frågats efter och besvarats på tidigare frågor).

Ordningsföljd *inom* frågor

Ordningen i frågeformuleringen

Hittills har detta kapitel diskuterat effekten av ordningsföljden *mellan* olika frågor. Ordningsföljden kan dock ha en effekt även *inom* frågor – det vill säga i själva frågeformuleringen. En del frågor inkluderar en jämförelse mellan två objekt i frågeformuleringen, till exempel är "Är Aftonbladet en bättre eller sämre tidning än Expressen?" eller "Är tennis roligare eller tråkigare att titta på än fotboll?". Har det i sådana här frågor någon betydelse vilket av jämförelseobjekten som skrivs först eller sist? Spontant kan man tycka att det inte borde ha det. Båda objekten (till exempel tennis och fotboll) är inkluderade och får samma utrymme i

frågan. Forskning i kognitiv psykologi antyder dock att det faktiskt kan ha betydelse. Även om båda objekten är inkluderade i frågan får de olika roller kognitivt (Tversky, 1977, se översikt i Wanke, Schwarz och Noelle-Neumann, 1995). Det objekt som står först kan bli en utgångspunkt och kallas då för subjekt. Det andra objektet, som subjektet jämförs med, kallas för referent. Forskningen visar att jämförelsen mellan objekten kan ske på subjektets villkor och att egenskaper som referenten men inte subjektet har då kan negligeras. Detta hänger ihop med det som stod inledningsvis i det här kapitlet – att den första informationsbiten har en unik position i en informationsmängd. Därför kan olika egenskaper bli aktuella i bedömningen beroende på vilket objekt som kommer först och i vilken mån objekten delar egenskaper (Wanke, Schwarz & Noelle-Neumann, 1995).

Wanke, Schwarz och Noelle-Neumann (1995) visar att ovanstående även kan gälla för enkätfrågor. Hur ska frågekonstruktören förhålla sig till detta? En möjlighet är att ha den ordningsföljd som är naturlig för respondenten eftersom en avvikande ordningsföljd skulle kunna leda till andra svar än vad respondenten skulle ha svarat spontant. Detta kan dock vara okänt på förhand och svårt att implementera i frågeformuleringen. En annan möjlighet är att ta bort riktningen i frågan och istället ha den i svarsalternativen. Då finns det dock en risk att problemet bara förskjuts till svarsalternativen. Ett tredje alternativ, som visserligen innebär en annan ansats, är att låta respondenterna skatta de olika objekten var och en för sig.

Sammanfattande punkter:

- En del frågor inkluderar en jämförelse mellan två objekt. Denna jämförelse blir inte nödvändigtvis helt rättvis.
- Det förstnämnda objektet kallas subjekt och det andra för referent och jämförelsen kan ske utifrån subjektets villkor (och bortse från till exempel styrkor som referenten men inte subjektet har). Denna problematik är svår att korrigera för i frågeformuleringen.

Konversationsnormer och konventioner

Ordningsföljden inom frågan kan också påverka den allmänna förståelsen av frågan. På flera ställen i denna bok nämns att konversationsnormer, regler som reglerar kommunikation, även är giltiga i kommunikation mellan respondent och frågeställare. En del av forskningsresultaten om frågekonstruktion skulle kunna förklaras av att respondenterna följer sådana normer och tar för givet att frågeställaren också gör det (läs mer i kapitel 7).

Syftet med konversationsnormer är att underlätta kommunikation av *mening* (i betydelsen "innebörd") och minska risken för missförstånd. Holbrook, Krosnick, Carson och Mitchell (2000) skiljer på konversationsnormer och konversationskonventioner där den sistnämnda är en konvention eller praxis som många följer men som inte påverkar meningen i det som kommuniceras. Det är till exempel vanligare att säga "Mamma, pappa, barn" än "Pappa, barn, mamma" men meningen är densamma. Att det kan påverka svaren att bryta konversationsnormer finns det alltså indikationer på, men vad gäller för att bryta konversationskonventioner?

Eftersom konventionen inte styr meningen som kommuniceras borde det inte leda till missförstånd att bryta den. Det kan däremot leda till ökad börda och vara förvirrande för respondenten. Även i surveysammanhang förekommer språkliga och designmässiga konventioner. Det är till exempel mycket vanligare att skriva "Är du för eller emot...?" jämfört med "Är du emot eller för...?". En annan konvention är i vilken ordning svarsalternativ skrivs (till exempel att "Ja" kommer före "Nej"). Ovan nämnda studie kom fram till att det kan förvirra respondenterna att bryta sådana normer. Det kan göra att besvarandet tar längre tid och kan också leda till fler fel. Resultaten gällde dock endast för respondenter med låg utbildning. En slutsats kan därför vara att det finns ett värde i att följa konventionen om det inte finns särskilda skäl att avvika från den (Holbrook, Krosnick, Carson & Mitchell, 2000).

Sammanfattande punkter:

- Att bryta konventioner kan förvirra respondenten och bör, generellt sett, endast göras om det finns ett syfte.

Andra typer av kontext

En bredare syn på kontext

Kontexten behöver inte endast bestå av föregående frågor utan även av andra, bredare faktorer. Det bredaste lagret utgörs nog av vår samtid och kultur. Även detta breda lager kan påverka hur frågor tolkas och förstås och kan medföra att samma fråga får olika betydelser i olika kulturer. (Den kulturella kontextens betydelse tas inte upp som enskilt kapitel i denna bok men är en utmaning för till exempel undersökningar som genomförs i flera länder eller för mångkulturella och flerspråkliga länder. Läs mer i Harkness, Braun, Edwards, Johnson, Lyberg, Mohler, Pennell & Smith, 2010).

En del andra faktorer som utgör kontexten (det sammanhang som respondenten besvarar frågeformuläret i) ligger bortom frågekonstruktörens kontroll, till exempel samhällsströmningar eller nyhetsrapportering. Man bör ändå vara medveten om dem. Andra faktorer ligger inom frågekonstruktörens och undersökningsorganisationens kontroll men kanske inte uppfattas som att utgöra en del av sammanhanget. En sådan faktor kan exempelvis vara sidvisningen i en webbenkät. Forskning har till exempel visat att svaren på frågor korrelerar starkare om de visas på samma skärmbild i webbundersökningar jämfört med om de delas upp på flera bilder (Tourangeau, Couper & Conrad, 2004). Kanske för att de uppfattas som att de hör ihop konceptuellt om de ingår i samma sammanhang, men det kan i det här fallet också bero på att risken för enkelriktade svarstilar ökar om man har flera frågor av samma typ på samma sida.

Intrycket av undersökningen – ett tolkningsramverk

En annan faktor i de bredare lagren av kontexten kan vara hur undersökningen lanseras och vilken undersökningsorganisation som står bakom undersökningen. Det kan finnas flera potentiella mekanismer som kan ge utslag på svaren. En mekanism kan vara att lanseringen av undersökningen etablerar ett tolkningsramverk i vilket respondenten förstår frågorna. Eftersom många frågor innehåller vaga ord kan tolkningsramverket påverka vilken tolkning respondenten gör (Galesic & Tourangeau, 2007). Är en tv en möbel? Respondenternas tolkningar kan säkert vara olika om det aktuella frågeformuläret handlar om inredning jämfört med elektroniska produkter.

En annan potentiell mekanism utgår från de konversationsregler som reglerar kommunikation (läs mer i kapitel 7). En sådan regel är samarbetsprincipen som hävdar att deltagare i ett samtal ska bidra med information som är värdefull och nyttig utifrån konversationens ämne. Applicerat till surveys kan denna princip innebära att respondenten uppfattar sin uppgift som att bidra med värdefull information till frågeställaren. Vad som är värdefullt eller inte kan förstås påverkas av det övergripande sammanhanget (påminner om det fenomen som kallas "demand characteristics" vilken kan vara förödande i forskningssammanhang, om testpersoner uppfattar vad forskaren är intresserad av och försöker bekräfta det). En sista indirekt mekanism som påverkar svarsfördelningen är bortfall. Lanseringen av undersökningen kan påverka vilka som väljer att delta eller inte, vilket kan leda till bortfallsfel (Galesic & Tourangeau, 2007). Den sista mekanismen diskuteras dock inte vidare här.

Kan denna typ av mekanismer ge utslag? Flera studier indikerar detta. I en enkät om situationer som kan uppstå på arbetsplatser svarade respondenterna något olika beroende på om undersökningen titulerades "Undersökning om sexuella trakasserier" och genomfördes av en feministgrupp med syfte att motarbeta sexuella trakasserier, jämfört med om undersökningen titulerades "Undersökning om atmosfären på arbetsplatser" och genomfördes av en neutral forskargrupp. Undersökningsgruppen med sexuella trakasserier som ramverk rapporterade fler situationer som tydligt inkluderade sexuella trakasserier (till exempel att någon stryker sig mot en annan persons kropp). De var även mer obekväma med otydliga situationer och rapporterade mindre förekomst av situationer som tydligt inte är sexuella trakasserier (till exempel att chefen sätter omöjliga deadlines). Effekterna är mycket små men kan kanske ändå illustrera mekanismen att lanseringen av undersökningen kan etablera ett ramverk som kan ha en viss påverkan på både hur frågorna förstås och vad frågeställaren tycker är relevant och värdefull information (Galesic & Tourangeau, 2007; Tourangeau, Groves, Kennedy & Yan, 2009). En liknande studie (Norenzayan & Schwarz, 1999) visar att respondenter betonade delvis olika typer av förklaringar till massmord, i linje med om forskarna bakom enkäten tillhörde "Institute of Personality Research" eller "Institute of Social Research".

Sammanfattande punkter:

- Kontexten kan ses som något bredare än frågeformulärets enskilda frågor.
- Vissa aspekter ligger bortom frågekonstruktörens kontroll, till exempel samhällsströmningar.
- Andra aspekter ligger inom frågekonstruktörens kontroll, till exempel hur undersökningen presenteras – detta kan etablera ett tolkningsramverk för respondenten och påverka hur denne tänker och svarar på frågorna.

Datainsamlingsmetodens roll

En annan faktor som kan påverka kontexteffekter är datainsamlingsmetoden. Både i telefonintervjuer och i självadministrerade enkäter finns en ordningsföljd mellan frågorna. I telefonintervjuer är den ordningsföljden absolut. Det finns ingen möjlighet för respondenten att läsa frågorna i förväg eller gå tillbaka och själv justera tidigare lämnade svar. I självadministrerade enkäter finns båda möjligheterna. Påverkar det eventuella kontexteffekter, alltså hur en föregående fråga påverkar en efterkommande?

En studie undersökte detta med tre frågor. En fråga handlade om att skänka pengar till välgörenhet och två frågor om inställningen till skatter och välfärd. När frågan om att skänka pengar till välgörenhet kom först i en telefonintervju svarade respondenterna i snitt att de skulle skänka 45 tyska mark. När frågan föregicks av de två frågorna om skatter och välfärd var motsvarande snitt endast 29 tyska mark. Resultatet visar en tydlig kontexteffekt där ordningsföljden har betydelse. I den självadministrerade enkäten var det emellertid ingen signifikant skillnad beroende på frågornas ordningsföljd. Intressant nog var det inte heller någon signifikant skillnad mellan de självadministrerade svaren och svaren i telefonintervjun när frågan om att skänka pengar till välgörenhet kom efter frågorna om inställningen till skatt och välfärd. Detta kan tyda på att ordningsföljden även hade betydelse i de självadministrerade varianterna men att effekten uteblev, sannolikt för att respondenterna hade möjlighet att läsa frågor i förväg eller justera sina svar (Schwarz & Hippler, 1995).

Sammanfattande punkter:

- I enkäter kan respondenten läsa frågorna i förväg och gå tillbaka och justera sina svar.
- I telefonintervjuer kan respondenten inte gå tillbaka och justera sina svar, vilket kan vara en förklaring till att den datainsamlingsmetoden visat starkare kontexteffekter.

Positionseffekter

Hittills har kapitlet handlat om kontexteffekter som främst har att göra med ordningsföljden mellan frågor eller enskilda ord i frågeformuleringen. Ordningsföljden kan också ha betydelse för svarsalternativen. Detta – att svarsalternativens ordningsföljd har en betydelse – brukar kallas för *positionseffekter* eftersom det syftar på att svarsalternativens position i en lista, snarare än deras innehåll, kan påverka fördelningen. Nominalskalor är ofta de skalor som har flest svarsalternativ, ibland långa listor, och de saknar dessutom en given inbördes ordning. Nominalskalor är därför den skaltyp som löper störst risk för positionseffekter.

Primacy- och recency-effekter

Positionseffekter kan vara av främst två typer. Den ena är en tendens att i högre grad välja svarsalternativ som kommer tidigt i en lista. Detta kallas för en primacy-effekt. Det andra är en tendens att i högre grad välja svarsalternativ som kommer sent i listan vilket kallas för en recency-effekt. Litteraturen om dessa effekter finns främst inom minnesforskning, men i viss mån också i forskningen om frågekonstruktion i surveyundersökningar.

Effekterna beror på minnet

Primacy-effekt är inom minnesforskningen en tendens att bäst minnas den första informationen i en informationsmängd. Recency-effekt är en tendens att bäst minnas den sista informationen i en informationsmängd. Forskning har visat att när testpersoner ska återge alternativ från en lista så är minneskurvan U-formad. Det betyder att de främst minns de första och sista alternativen. Detta kan tyckas motsägelsefullt, men det är det inte. För att förstå det behöver man känna till hur arbetsminnet och

långtidsminnet fungerar (läs mer om framförallt långtidsminnet i kapitel 11).

Arbetsminnet är ett begränsat minnessystem där vi temporärt kan behålla information som vi behöver eller använder för tillfället (som ett telefonnummer). Långtidsminnet är mer som ett stort bibliotek av lagrade och mer hållbara minnen. Det som händer när en testperson exponeras för en lista med alternativ är att för de första alternativen kan vi upprepa och lära in dem så pass väl att de inkodas i långtidsminnet. Allteftersom alternativen blir fler fylls emellertid det begränsade arbetsminnet. Vi hinner inte lära in dem och de förs därmed inte över till långtidsminnet. De sista alternativen i listan har däremot förmånen att inte ersättas av ny information och finns därför kvar i testpersonernas arbetsminne. En primacy-effekt är alltså ett resultat av att personer i början av en uppräknings har tid att koda in informationen i långtidsminnet. En recency-effekt sker däremot på grund av att de sista alternativen i en uppräknings inte ersätts av nya och därför finns kvar i arbetsminnet en stund efteråt för att sedan blekna eller ersätts med annan typ av information (Passer & Smith, 2007).

Att det förhåller sig på ovanstående vis har bekräftats genom olika forskningsstudier. Om man till exempel försvårar inövning av de första alternativen genom att presentera alternativen i listan väldigt fort, försvagas primacy-effekten (Passer & Smith, 2007). Om man däremot väntar en stund (det kan räcka med så kort tid som 30 sekunder) eller om man fyller arbetsminnet med annat innan man ber testpersoner återupprepa de sista alternativen i listan så försvagas eller försvinner recency-effekten (Passer & Smith, 2007).

Översatt till surveyfrågor innebär alltså en primacy-effekt att de svarsalternativ som kommer tidigt i en lista får högre stöd (fler väljer dem jämfört med om de haft lägre positioner i samma lista). På motsatt sätt innebär en recency-effekt högre stöd för svarsalternativ i slutet på en lista.

Sammanfattande punkter:

- En primacy-effekt är i minnespsykologin en tendens att bäst minnas den första informationen i en informationsmängd. I surveysammanhang betyder det förhöjt stöd för svarsalternativ som kommer tidigt i en lista.

- En recency-effekt är i minnespsykologin en tendens att bäst minnas den sista informationen i en informationsmängd. I surveysammanhang betyder det förhöjt stöd för svarsalternativ som kommer i slutet på en lista.
- Primacy-effekter beror på att individer i början av en uppräkningslista har tid att koda in informationen i långtidsminnet och därför minns den bättre än senare information som inte hinner kodas in i samma grad.
- Recency-effekter beror på att den sista informationen finns kvar i respondentens arbetsminne (innan den bleknar eller ersätts av annan information).

Datainsamlingsmetodens roll för positionseffekter

Primacy- och recencyeffekter i surveyundersökningar beror på insamlingsmetod. Vid pappersenkäter läser förmodligen respondenterna alternativen i sin egen takt och det finns tid för upprepning och inövning. Där finns det alltså risker för en primacy-effekt (Schwarz, Knäuper, Oysermann & Stich, 2008). I telefonintervjuer presenteras däremot svarsalternativen i en rask följd (och dessutom endast som ljud). Det ges därför inte så stora möjligheter att upprepa eller lära in alternativen. Detta innebär en risk för fokus på de sistnämnda alternativen som fortfarande finns kvar i arbetsminnet efter att uppläsningen avslutats, en så kallad recency-effekt (Schwarz, Knäuper, Oysermann & Stich, 2008).

En studie fann att recency-effekter framförallt förekommer vid svårare frågor, längre svarsalternativ och svarsalternativ som inte var varandra uteslutande. De var även vanligare i slutet av frågeformuläret (Holbrook, Krosnick, Moore & Tourangeau, 2007). Allt detta är förenligt med förklaringen att det är arbetsminnets begränsningar som orsakar recency-effekten.

Forskningen är inte entydig

Enligt en reviewartikel (en översikt av forskningsläget) med studier om ordningseffekter hade tio studier hittat primacy-effekter, 28 hade hittat recency-effekter och 41 fick inga signifikanta effekter. Alla recency-effekter hittades i intervjuer (ljudmässig information) och alla utom två primacy-effekter hittades i undersökningar med enkäter (visuell information) (Krosnick, 2000). Artikeln bekräftar alltså att effekterna relativt entydigt verkar hänga ihop med datainsamlingsmetoden.

Forskningen inom surveyområdet är dock *inte* entydig när primacy- och recency-effekter får genomslag och deras påverkan ska inte överdrivas (41 av studierna i reviewartikeln ovan hittade inga signifikanta effekter). Detta kan bero på att det kan finnas andra effekter som också påverkar och ibland ger utslag. Telefonundersökningar är till exempel mer känsliga för social önskvärdhet. Om det sista alternativet också är det minst socialt önskvärda kanske detta kan styra i det fallet, snarare än en recency-effekt (de Leeuw, 2008; Dillman, 2007). Det kan finnas många fler sådana konfliktsituationer och det gör forskningsresultaten lite spretiga. Som beskrivs i det förra avsnittet varierar effekten också beroende hur svår frågan är, hur långa svarsalternativen är, hur omfattande frågeformuläret är och var i frågeformuläret den aktuella frågan är placerad.

Sammanfattande punkter:

- I enkäter finns utrymme att läsa om frågorna och koda in informationen, vilket är förknippat med primacy-effekter.
- Intervjuer är beroende av respondentens arbetsminne, vilket är förknippat med recency-effekter.
- Flera studier har dock inte funnit någon effekt alls.

Minnesforskning och satisficing

Ibland kallas det för primacy-effekt när en respondent inte läser igenom alla alternativ ordentligt utan väljer något av de första för att det är "tillräckligt bra" eller för att komma vidare med så liten ansträngning som möjligt (så kallad satisficing). Resultatet blir detsamma som av det som kallas primacy-effekt inom minnesforskningen (i båda fallen väljs de första svarsalternativen på "felaktiga" grunder) men den bakomliggande mekanismen är inte densamma. Primacy-effekten har med kognitiva processer att göra medan satisficing främst beror på bristande motivation. Det är alltså två separata mekanismer, även om resultatet blir detsamma.

Krosnick (2000) skriver att benägenheten till satisficing, utöver motivation, också påverkas av uppgiftens svårighet och respondentens kognitiva förmåga. Uppgiftens svårighet beror i sin tur på hur svår frågan är att förstå, hur tillgänglig informationen är att återhämta i

minnet, om svarsalternativen är tydliga och lättförståeliga, intervjuens tempo och om det finns distraktioner i närmiljön (Krosnick, 2000).

Krosnick föreslår att satisficing kan minskas genom att öka motivationen (förklara syftet med undersökningen, instruera respondenterna att tänka igenom svaren ordentligt, ha korta enkäter) och förenkla uppgiften så mycket som möjligt (använd enkla och entydiga ord, fråga om nutida händelser, fråga om en sak i taget). Det är de traditionella rekommendationerna och inga nyheter för en erfaren frågekonstruktör. Men är det alltid bäst att förenkla? Kanske kan vissa respondenter bli mer motiverade av mer "utmanande" frågor så att tendensen till satisficing minskar när frågorna blir mer komplexa? En studie testade just detta genom att jämföra positionseffekterna i en komplex och en simpel uppgift. Den komplexa uppgiften var att ranka offentliga tjänstemän i hur pass ansvariga de var för skadorna efter orkanen Katrina. En komplex och krävande men samtidigt förmodligen en relevant och motiverande uppgift. Denna uppgift jämfördes med den simplare uppgiften att besvara några frågor av ordinalskaletyp. I den här studien hittades i stort sett inga positionseffekter för den komplexa uppgiften men signifikanta sådana för den enkla (Malholtra, 2009). Det kan alltså krävas att frågekonstruktören ställer en uppgifts komplexitet mot respondentens motivation. Att förenkla en uppgift är inte nödvändigtvis alltid den bästa lösningen - vi kan inte ta för givet att respondenterna är ointresserade.

Positionseffekter kan också "jämnas ut" genom att randomisera svarsalternativens ordningsföljd eller undvikas genom att bryta ner stora frågeblock med samma svarsskala till separata frågor (Krosnick, 2000).

Sammanfattande punkter:

- Förhöjt stöd för tidiga svarsalternativ kan bero på att respondenten väljer något av de första alternativen bara för att komma vidare (så kallad satisficing).
- Förhöjt stöd för tidiga svarsalternativ kan minimeras genom att dela upp frågor i block, att ha "Vet inte"-alternativ eller att randomisera ordningsföljden.

Studier av ögonrörelser

Det finns alltså två konkurrerande förklaringar till förhöjt stöd för tidiga svarsalternativ i en lista – minneseffekt och satisficing. I survey-forskningen kan vi oftast bara studera dessa förklaringar indirekt och då går det inte att avgöra vilken av dem som är giltig. I en studie användes emellertid en eye-tracking-teknik som registrerar ögats rörelser och fixering. Med denna teknik går det att undersöka hur länge respondenter har fokus på olika svarsalternativ och om alla svarsalternativ läses. Studien hittade stöd för båda förklaringarna. Vissa respondenter läste alla svarsalternativ men ägnade mer tid åt de första svarsalternativen. Andra läste inte alla alternativ i listan (Galesic, Tourangeau, Couper, & Conrad, 2008). Båda förklaringarna – minneseffekt och satisficing – kan alltså vara giltiga.

Sammanfattande punkter:

- Studier som undersökt både minneseffekter och satisficing direkt (genom att registrera ögonrörelser) fann stöd för båda förklaringarna.

Minneseffekter och långa frågor

Den mesta forskningen om dessa minneseffekter i surveyundersökningar har fokuserat på svarsalternativen. Samma mekanismer (primacy och recency) kan nog även vara aktuella vid långa frågeformuleringar och instruktioner. En studie provade två varianter av samma fråga (Beatty, Fowler & Cosenza, 2007). Frågan bestod av dels en explicit frågeformulering ("Var får du din läkarhjälp?") och en massa exempel (privat mottagning, sjukhus etc.). I ena fallet började frågan med frågeformuleringen och därefter exemplen. I det andra fallet var det tvärtom. Det visade sig att svarsmönstret i slutänden blev ungefär detsamma men att ställa frågeformuleringen först och sedan exemplen ledde till fler avbrott i intervjun. Författarna till denna studie diskuterar inte i termer av minneseffekter men resultatet skulle kunna vara ett utfall av arbetsminnets begränsning – på grund av de många exemplen i frågan har respondenterna hunnit glömma den ursprungliga frågeformuleringen (den har försvunnit ur arbetsminnet) och de avbryter därför intervjuaren och ber om upprepningar och förtydliganden. Dessa minneseffekter kan sålunda säkert vara aktuella även för långa frågeformuleringar och instruktioner.

Sammanfattande punkter:

- Problemet med för mycket information för arbetsminnet kan vara relevanta för långa frågeformuleringar också, inte bara för listor med svarsalternativ.

Sammanfattning

Ordningsföljder och kontext – sammanfattning**Ordningsföljder och kontext:**

- Informationsbitars ordningsföljd kan ha betydelse.
- Den första informationsbiten är unik och fungerar som referens för senare information.
- Även i surveyundersökningar kan ordningsföljd och tidigare etablerat sammanhang påverka svaren på enskilda frågor. Kontexteffekter och positionseffekter är två exempel.

Ordningsföljd: enhet-helhet:

- Ordningsföljden på frågor kan påverka svaren. Detta kallas för en kontexteffekt.
- En variant är enhet-helhet: en specifik fråga och en generell fråga som också omfattar det specifika.
- Svaret på den generella frågan blir annorlunda om den specifika frågan ställs först.
- Det finns olika typer av effekter (jämfört med om den specifika frågan inte föregått den generella): 1) att respondenterna svarar i linje med den specifika frågan när de svarar på den generella frågan (assimileringsseffekt) eller 2) att den specifika frågan får respondenterna att tänka och tolka om den generella frågan i avvikande riktning (kontrasteffekt).

Förklaringar – Priming och "carry-over":

- Assimileringsseffekter kan bero på så kallad priming – att tankar och känslor aktiverade vid en fråga "spiller över" på efterföljande frågor.
- Efterföljande frågor uppfattas då annorlunda än om inte föregående fråga hade ställts, och i linje med den föregående frågan.

- Effekten är beroende av arbetsminnet och är därför till exempel svagare hos äldre respondenter.

Subtraktionsförklaringar:

- Kontrasteffekt kan bero på subtraktion – att respondenten inte tar med saker i den generella frågan som de redan svarat på i den specifika frågan.
- Subtraktion kan grunda sig i konversationsreglerna – man ska inte fråga efter eller ge redan känd information (sådant som redan frågats efter och besvarats på tidigare frågor).

Ordningsföljd inom frågor:

- En del frågor inkluderar en jämförelse mellan två objekt. Denna jämförelse blir inte nödvändigtvis helt rättvis.
- Det förstnämnda objektet kallas subjekt och det andra för referent och jämförelsen kan ske utifrån subjektets villkor (och bortse från till exempel styrkor som referenten men inte subjektet har). Denna problematik är svår att korrigera för i frågeformuleringen.
- Att bryta konventioner kan förvirra respondenten och bör, generellt sett, endast göras om det finns ett syfte.

Andra typer av kontext:

- Kontexten kan ses som något bredare än frågeformulärens enskilda frågor.
- Vissa aspekter ligger bortom frågekonstruktörens kontroll, till exempel samhällsströmningar.
- Andra aspekter ligger inom frågekonstruktörens kontroll, till exempel hur undersökningen presenteras – detta kan etablera ett tolkningsramverk för respondenten och påverka hur denne tänker och svarar på frågorna.
- I enkäter kan respondenten läsa frågorna i förväg och gå tillbaka och justera sina svar.
- I telefonintervjuer kan respondenten inte gå tillbaka och justera sina svar, vilket kan vara en förklaring till att den datainsamlingsmetoden visat starkare kontexteffekter.

Positionseffekter – effekterna beror på minnet:

- En primacy-effekt är i minnespsykologin en tendens att bäst minnas den första informationen i en informationsmängd. I surveysammanhang betyder det förhöjt stöd för svarsalternativ som kommer tidigt i en lista.
- En recency-effekt är i minnespsykologin en tendens att bäst minnas den sista informationen i en informationsmängd. I surveysammanhang betyder det förhöjt stöd för svarsalternativ som kommer i slutet på en lista.
- Primacy-effekter beror på att individer i början av en uppräkningslista har tid att koda in informationen i långtidsminnet och därför minns den bättre än senare information som inte hinner kodas in i samma grad.
- Recency-effekter beror på att den sista informationen finns kvar i respondentens arbetsminne (innan den bleknar eller ersätts av annan information).

Datainsamlingsmetodens roll för positionseffekter:

- I enkäter finns utrymme att läsa om frågorna och koda in informationen, vilket är förknippat med primacy-effekter.
- Intervjuer är beroende av respondentens arbetsminne, vilket är förknippat med recency-effekter.
- Flera studier har dock inte funnit någon effekt alls.

Minnesforskning och satisficing:

- Förhöjt stöd för tidiga svarsalternativ kan bero på att respondenten väljer något av de första alternativen bara för att komma vidare (så kallad satisficing).
- Förhöjt stöd för tidiga svarsalternativ kan minimeras genom att dela upp frågor i block, att ha "Vet inte"-alternativ eller att randomisera ordningsföljden.
- Studier som undersökt både minneseffekter och satisficing direkt (genom att registrera ögonrörelser) fann stöd för båda förklaringarna.

Minneseffekter och långa frågor:

- Problemet med för mycket information för arbetsminnet kan vara relevanta för långa frågeformuleringar också, inte bara för listor med svarsalternativ.

14 Att utvärdera frågor

Författare: Andreas Persson (SCB)

Varför utvärdera?

Att det kan vara svårt att konstruera bra frågor har nog framgått vid det här laget. Det finns flera saker som kan bli fel eller mindre bra vilket kan leda till att man delvis eller till och med helt missar målet. Det är ett skäl till att frågorna bör utvärderas innan de ska användas men det finns fler. Här är en lista över de främsta skälen till att utvärdera frågorna:

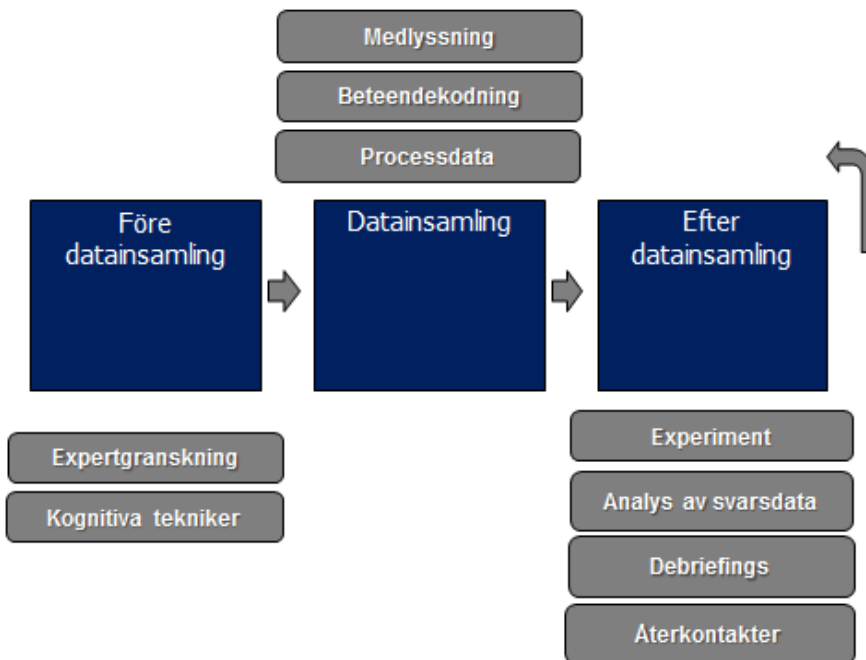
- **Identifiera problem:** Detta görs för att bli varse om vilka frågor som kan vara problematiska i något avseende och som därför behöver omarbetas.
- **Problemdiagnos:** Det ger en kännedom om eventuella problems natur så att man kan reducera eller neutralisera dem.
- **Underlätta administrationen:** Ett problematiskt frågeformulär tar mer resurser att administrera. Det gäller speciellt i telefonintervjuer där problematiska frågor kan ta mycket tid i anspråk och leda till en rörig intervjusituation.
- **Minska respondentbördan:** Genom utvärdering kan frågeformuläret förbättras så att det blir enklare för respondenten att besvara.
- **Minska efterarbetet (till exempel korrigerings av fel):** Ett frågeformulär som resulterar i mycket fel i insamlad data kräver mer efterarbete för att kunna användas. Mängden efterarbete kan minskas genom att utvärdera frågorna i förväg.
- **För att få kvalitetsmått på centrala variabler:** För att publiceras i vetenskapliga tidskrifter krävs ibland mått på till exempel centrala variabelers tillförlitlighet (som Cronbach's alpha) och ibland även validitetsmått.
- **För att få indikationer på hur ett oväntat resultat ska förstås:** Ibland kan ett oväntat resultat bero på att respondenterna tolkat frågorna annorlunda än vad som varit avsikten. Genom att utvärdera frågorna kan förvånande resultat bli begripliga.

Om en utvärdering visar att frågorna fungerar bra kan man också ha större tilltro till datamaterialet och de slutsatser som dras. Även tredje part, de som tar del av undersökningsresultatet, kan känna mer tillit till siffrorna om frågeformuläret testats.

Olika metoder för att utvärdera frågor

Det finns många metoder att utvärdera frågor med och frågetestning är ett eget ämnesområde i sig med en egen litteratur (till exempel Presser, Rothgeb, Couper, Lessler, Martin, Martin & Singer, 2004 eller Madans, Miller, Maitland & Willis, 2011). Detta avsnitt har som syfte att ge en övergripande orientering om vilka metoder som kan användas och något om hur de kan tillämpas i praktiken. För djupare studier i ämnet hänvisas till mer specialiserad litteratur, till exempel ovan nämnda böcker.

Figur 14.1



Figuren ovan visar de vanligaste metoderna som används för att utvärdera frågor. Precis som figuren visar är de naturligt knutna till olika faser i en undersökning. Vissa förekommer främst före datainsamlingen, så kallade förtest ("pre-test" på engelska). De är ofta kvalitativa och småskaliga till naturen, till exempel expertgranskning och kognitiva intervjuer (se beskrivningar nedan) och använder inga kvantitativa data. Att endast kvalitativa metoder används i detta steg beror på att det före

datainsamling inte finns någon kvantitativ data tillgänglig såtillvida man inte genomfört en pilotundersökning (en provundersökning eller generalrepetition av den riktiga undersökningen).

Andra metoder är främst knutna till själva datainsamlingen och använder data som skapas i densamma, till exempel processdata och medlyssning. Under boxen längst till höger i figuren ovan ligger de metoder som används efter att undersökningen genomförts. Det kan innebära att analysera resultatet från ett inbäddat experiment eller att analysera om olika variabler hänger ihop som förväntat i svarsdata. Det kan även gälla att genomföra separata utredningar, till exempel debriefingar (strukturerade gruppdiskussioner) med produktionspersonal (till exempel intervjuare), respondenter eller återkontaktstudier där respondenterna får beskriva hur de tänkte när de besvarade undersökningen.

Vad som är "före" respektive "efter" datainsamlingen är inte lika cementerat som figuren ovan kan ge intryck av. Vid återkommande eller longitudinella undersökningar är ju utvärderingar efter en produktionsomgång samtidigt före nästa produktionsomgång, även om de ligger längst till höger i figuren ovan. Då handlar det snarare om ett cykliskt och kontinuerligt förbättringsarbete. Om man genomför en pilotundersökning kan de metoder som ligger under boxen "Efter datainsamlingen" dessutom användas före huvudundersökningens datainsamling.

Metoderna skiljer sig åt på många sätt. Exempelvis är vissa metoder kvalitativa, förhållandevis resurssnåla och kan ge god förståelse för bakomliggande faktorer, men inte så mycket om utbreddheten eller de statistiska konsekvenserna av eventuella problem i frågeformuläret. Andra är kvantitativa, ibland mer resurskrävande och kan ge ett precist värde på en frågas prestation men inte alltid ge någon närmare inblick i vad som orsakat ett visst problem med en fråga. Avsnitten nedan har som mål att ge en introduktion till de olika metoderna vad gäller syfte och utförande samt styrkor och svagheter.

Sammanfattande punkter:

- Det finns många skäl att utvärdera sina formulärfrågor (se punktlista ovan).
- Det finns många olika utvärderingsmetoder och de är naturligt knutna till olika faser i en undersökning (se figur ovan).

- Metoderna varierar också bland annat i resurskrav, om de är kvalitativa eller kvantitativa, om de ger ett statistiskt säkerställt värde eller en djupare förståelse.

Expertgranskning

Syftet med expertgranskning

Expertgranskning⁴ ("expert review" på engelska) är en metod som praktiseras på de flesta nationella statistikbyråer i västvärlden. Den innebär att ett frågeformulär granskas på ett systematiskt sätt av tredje part som har gedigna erfarenheter och kunskap om frågekonstruktion – så kallade experter på frågekonstruktion (därav namnet expertgranskning). Att det är en tredje part som genomför granskningen är viktigt för att minska konsekvenser av "hemmablindhet" – att man har svårt att se problem med frågor man själv skrivit. Detta beror främst inte på egocentrism utan är grundat i hur språket fungerar – att vi människor är mer intresserade av innebörd än faktiska ord. Om vi skrivit frågorna själva känner vi naturligtvis redan till innebörden eftersom vi vet vad vi menar. Då kan det vara svårt att granska på bokstavlig nivå (läs mer i kapitel 7).

Ibland innefattas även ämnesexperter och statistiker (ESS, 2006) i expertgranskningen eftersom de kompetenserna också är viktiga för frågornas utformning och för att frågekonstruktören inte nödvändigtvis har dessa.

Ett syfte med expertgranskningen är att försöka säkerställa att frågorna samlar in den önskade informationen. Ett annat syfte är att identifiera andra potentiella problem samt justera dem enligt de principer som gäller för fråge- och formulärkonstruktion (till exempel den forskning och de erfarenheter som redovisas i denna bok). Till skillnad från till exempel kognitiva intervjuer inbegriper inte expertgranskningen några testpersoner som företräder respondenterna och den bör därför helst kompletteras med någon sådan metod (Yan, Kreuter & Tourangeau, 2012; ESS, 2006).

Utvärderingar av expertgranskningen som metod visar att vad de olika experterna hittar för problem i sina granskningar varierar, men att

⁴ Metoden kan även ha andra benämningar. På SCB kallas den även "Mätteknisk granskning".

experterna som helhet ändå är framgångsrika i att förutsäga vilka frågor som leder till bortfall eller felaktig rapportering (Olson, 2010).

Expertgranskning och kognitiva intervjuer

Expertgranskningen identifierar i regel många problem, speciellt om den genomförs med ett kodningsschema eller checklista, till exempel QAS (Question appraisal system) (Rothgeb, Willis & Forsyth, 2001). Tidigare forskning visar att expertgranskningar ibland identifierar fler problem än till exempel kognitiva intervjuer (Presser & Blair, 1994; Willis, 2005) och ibland inte (Rothgeb, Willis & Forsyth, 2001). Däremot skiljer sig metoderna en del angående vilka problem som identifieras ($r=.40$) (Willis, 2005; Yan, Kreuter & Tourangeau, 2012). Detta är ytterligare ett skäl till att de bör kombineras.

Proceduren

Expertgranskning är oftast den förtestmetod som är billigast och enklast att genomföra. Den kräver inte mer resurser än den tid det tar för experterna att granska frågeformuläret och rapportera resultatet. Det rekommenderas att expertgranskningen genomförs tidigt i förtest-fasen (ESS, 2006). Den är lämpligt som ett första steg före till exempel kognitiva tester för att eliminera eventuella "slarvfel" och liknande. Den utgör även ett underlag för vad de kognitiva testen ska fokusera på. En expertgranskning kan med fördel inkludera flera experter eftersom uppgiften ofta innehåller bedömningar snarare än självklara rätt och fel.

Granskningen kan ske:

- **ostrukturerat** – genom att frågeformuläret granskas mer informellt utan att en checklista eller kodningsmall används
- **strukturerat** – genom att varje fråga kodas enligt en standardiserad kodningsmall, till exempel QAS.

Oftast fokuserar granskningen på följande områden (Biemer & Lyberg, 2003; Lessler & Forsyth, 1996):

1. språket, termer och begrepp
2. strukturen på frågorna
3. svarsalternativen
4. ordningsföljden
5. navigeringen genom frågeformuläret
6. instruktioner

7. layout
8. typografin.

Ostrukturerade/informella expertgranskningar

I ostrukturerade expertgranskningar diskuteras frågeformuläret utan att utgå från någon checklista eller mall. Diskussionen kan ske i gruppform och ledas av en moderator som ser till att diskussionen berör relevanta teman som formuleringar, funktionalitet och förståelse. Diskussionen kan spelas in för mer djupgående analys (ESS, 2006).

Ostrukturerade expertgranskningar kan också ske mer informellt utan moderator och utan att diskussionen spelas in. Denna typ av granskning handlar mer om att få feedback från kollegor eller att några kollegor arbetar tillsammans med att granska ett frågeformulär (ESS, 2006). Det rekommenderas dock att expertgranskningarna har någon mån av struktur för att inte missa problem i frågeformuläret (ESS, 2006).

Strukturerade expertgranskningar

I strukturerade expertgranskningar följer granskningen en checklista, till exempel att varje fråga i frågeformuläret kodas utifrån ett särskilt kodningsschema. Detta underlättar en systematisk granskning av frågeformuläret och gör att granskarna kan ge och diskutera sin feedback på ett standardiserat sätt (ESS, 2006).

QAS är ett exempel på en checklista för formulärgranskning (Willis & Lessler, 1999). Den innehåller punkter som till exempel instruktioner (om de är motsägande eller för komplicerade), tydlighet i frågorna och brister vad gäller respondenternas kunskap och minne. En sådan checklista kan vara en bra utgångspunkt, speciellt om granskaren är ovan.

Tidsåtgång och resultat

Resultatet från och tidsåtgången för en expertgranskning beror givetvis på frågeformulärets längd, skick och komplexitet. Det beror också på vilken form av expertgranskning som genomförs. Om en strukturerad expertgranskning genomförs där varje fråga kodas, tar granskningen naturligtvis längre tid (ESS, 2006).

Resultatet av en expertgranskning är ett justerat och förbättrat frågeformulär som antingen kan fortsätta testas (till exempel med kognitiva test) eller tas i produktion (om inga fler test ska genomföras).

Det rekommenderas dock att expertgranskningar kompletteras med någon form av test som involverar möjliga respondenter.

Sammanfattande punkter:

- Expertgranskning innebär att en tredje part, experter på frågekonstruktion, granskar frågeformuläret på ett systematiskt sätt.
- Syftet med expertgranskning är att identifiera möjliga problem och föreslå förbättringar.
- Expertgranskningen kan ske ostrukturerat och informellt eller mer strukturerat utifrån en checklista eller kodningsmall.
- Utvärderingar av metoden visar att den är framgångsrik när det gäller att identifiera problem, även om det kan variera vilka problem olika granskare uppmärksammar.

Kognitiva tekniker

Syftet med kognitiva tekniker

Kognitiva intervjuer bygger på kognitiv psykologi som är läran om hur människan behandlar information (Lundh, Montgomery & Waern, 1992). "Kognitiva intervjuer" används ibland som ett samlingsnamn för flera tekniker som kan användas i intervjuform. Dessa tekniker bygger på antagandet att kognitiva processer kring frågorna i ett frågeformulär kan identifieras och beskrivas muntligt, genom verbala rapporter från respondenterna (Ericsson & Simon, 1993).

Eftersom kognitiva intervjuer undersöker testpersoners tankar kring frågorna kan de ge god information om förståelsen av frågorna, bearbetning av svar och anpassning av svaret till frågeformulärets svarsalternativ. Intervjuerna syftar till att upptäcka dessa fel och få idéer om lösningar, snarare än att fastslå generella principer om frågekonstruktion (Fråga rätt, 2001; ESS, 2006). Kognitiva intervjuer används alltså mest i tillämpat syfte, att förbättra frågorna i ett visst frågeformulär, snarare än att forska om allmänna principer för frågekonstruktion. Detta har också bidragit till att metoden används på lite olika sätt – det finns ingen helt etablerad och fastställd procedur. Däremot råder det inga tvivel om att kognitiva intervjuer kan ge väldigt värdefull information i förbättringsarbetet med formulärfrågor – information som är svår att få på något annat sätt.

Olika kognitiva tekniker

Att bara observera hur frågorna besvaras ger inte tillräcklig information för att avgöra om problem förekommer i svarsprocessen. Inte sällan svarar respondenterna snabbt och övertygande trots att det inte är säkert att de har förstått frågan korrekt eller vara säkra på svaret. För att upptäcka när och varför det inträffar finns en rad kognitiva tekniker som kan kombineras på olika sätt beroende på vilka svårigheter man förutser i frågeformuläret. Vilken eller vilka tekniker man väljer att använda beror alltså på vad man misstänker för problem med frågorna. Valet anpassas också till datainsamlingsmetoden och kategorin av respondenter i den planerade undersökningen (Fråga rätt, 2001). Det finns även generella rekommendationer om när en teknik är att föredra framför en annan (ESS, 2006). Teknikerna redovisas kort nedan.

Tänka högt ("think aloud" på engelska)

Tekniken innebär att testpersonen "tänker högt", beskriver sina tankar i realtid när frågorna besvaras. Intervjun inleds vanligtvis med en instruktion och en övning som hjälper testpersonen att förstå proceduren (Fråga rätt, 2001). Tekniken kan användas både samtidigt som frågorna besvaras och retrospektivt efter att frågeformuläret eller varje enskild fråga besvarats. Enligt den renläriga metodiken ska dock testpersonen tänka högt i realtid, inte retrospektivt.

Tänka högt har sitt ursprung i kognitiv forskning om problemlösning. Testpersoner fick i olika studier tänka högt medan de övervägde ett schackdrag, löste ett mekaniskt problem eller en tankenöt. Påverkar då inte talandet tänkandet? Det är lätt att föreställa sig att tänka högtmomentet i sig påverkar tänkandet, speciellt för eftertänksamma personer som funderar mycket innan de formulerar sig. Forskningen ger dock inte så mycket stöd för att talandet skulle påverka tänkandet. Personer som tänker högt löser uppgifter på i princip samma tid och på samma sätt som personer som inte tänker högt (Willis, 2005).

Idén bakom tekniken är att om personer tänker högt i realtid kan man få ett flöde direkt från arbetsminnet – den information som individen använder eller arbetar med just då. Om personerna tvärtom överväger hur de ska formulera sig, sker kopplingar till långtidsminnet och det finns då risk att man både påverkar tänkandet i sig och får en tillrättalagd utsaga. Det är alltså viktigt att försöka få till ett spontant och oredigerat verbalt flöde från testpersonen. Detta är också anledningen till att

tekniken inte fungerar, eller åtminstone blir delvis något annat, om den utförs retrospektivt.

En fördel med tänka högt-tekniken är att den är mer explorativ och mindre intervjuarstyrd än till exempel probingtekniken (se nedan). Risken för oönskad intervjuarpåverkan är därför mindre. Eftersom tekniken är mindre intervjuarstyrd kräver den mindre träning av och förarbete för intervjuaren jämfört med probingtekniken. En nackdel är att den istället ställer högre krav på testpersonen som ska tänka högt. En del personer har svårt att klä sina tankar i ord eller få igång ett flöde (Prüfer & Rexroth, 2005). En risk med tänka högt-teknikens explorativa natur är också att testpersonerna svävar ut från ämnet och fastnar i långa utläggningar kring saker som inte är relevanta för att avgöra huruvida frågan fungerar eller inte (Willis, 2005).

Tänka högt-tekniken är alltså lämplig om syftet med den kognitiva intervjun är explorativt. Om syftet däremot är att i detalj utreda redan identifierade problem kan probingtekniken vara ett skarpare verktyg.

Probingteknik

Probes är engelska och kan i det här fallet översättas ungefär med utredningsfrågor. Tekniken innebär att intervjuaren ställer utredningsfrågor eller så kallade probes efter att testpersonen besvarat en formulärfråga i syfte att få testpersonerna att berätta om hur de resonerat sig fram till sina svar, hur de förstår olika ord, termer och begrepp, och vad som innefattas i svaret. Liksom för tänka högt-metoden kan utredningsfrågor ställas *direkt efter* att enskilda frågor besvarats eller efter att hela frågeformuläret besvarats (probingen kallas då retrospektiv). Att ställa utredningsfrågor direkt kan vara en styrka eftersom det kan vara svårt för testpersoner att efteråt minnas och återge sina tankar kring specifika frågor. Om man däremot ställer utredningsfrågorna direkt är det risk att själva ifyllandet störs och att testpersonen reagerar annorlunda på senare frågor än vad som skulle ha varit fallet utan avbrotten (Willis, 2005; Fråga rätt, 2001; ESS 2006).

Utredningsfrågorna kan vara mer eller mindre *generella*. Med generell följdfråga menas att den är allmänt formulerad, till exempel "Kan du berätta mer om detta?" (Willis, 2005; Fråga rätt, 2001). En specifik följdfråga är däremot skraddarsydd för en viss fråga i frågeformuläret. Utredningsfrågorna kan vara *gemensamma* och ställas till alla testpersoner eller *villkorliga*. Det innebär att de bara ställs givet att något speciellt

inträffar under intervjun, till exempel att testpersonen tvekar eller svarar på ett särskilt sätt (Willis, 2005; Fråga rätt, 2001).

Utredningsfrågorna kan också kategoriseras efter vad de syftar till att belysa för typ av problem, som till vilket steg i den kognitiva modellen över svarsprocessen de är knutna till. De kan ha för avsikt att ge information om testpersonens *förståelse* eller *informationshämtning* avseende en viss fråga (Fråga rätt, 2001).

Exempel på probes:

EXEMPEL:

Vad inkluderar ditt svar?

Hur gick du tillväga när du besvarade frågan?

Kan du ge exempel på vad du tycker ingår respektive inte ingår i frågan?

Jämfört med tänka högt-metoder är probes ett mer precist verktyg i och med att intervjuaren kan rikta fokus på de delar som på förhand misstänks vara problematiska. Intervjuaren kan behålla kontroll över intervjusituationen så att inte onödig tid läggs på irrelevanta diskussioner. Genom probes kan man även få testpersoner att beskriva sådant som de inte skulle berätta spontant genom att tänka högt. Det är enklare att få testpersoner att försöka beskriva sina tankar när man frågar om dem än att få dem att tänka högt spontant (Willis, 2005). En nackdel är att testpersonens svar styrs av hur utredningsfrågorna är utformade och vad de fokuserar på (Willis, 2005). I värsta fall kan resultatet av probingtekniken bli en självuppfyllande profetia – att intervjuaren tycker sig hitta problem där de misstänker att de finns genom att ställa ledande utredningsfrågor (Willis, 2005).

Vinjetter

Vinjetter är fall- eller scenariobeskrivningar av till exempel en situation eller en persons beteende, som testpersonen ombeds ta ställning till. Vinjetter används bland annat när man är osäker på om den vardagliga betydelsen stämmer överens med den formella eller tvärtom. Teknikens huvudsakliga syfte är att undersöka hur konsekvent vinjetterna tolkas och klassas av testpersonerna och om detta överensstämmer med den intention som finns (Fråga rätt, 2001; ESS, 2006).

Vinjetter kan användas som frågor direkt i frågeformuläret och resultaten kan då kvantifieras (ESS, 2006). De kan dock även användas kvalitativt i kognitiva intervjuer och bidrar då med en referenssituation som testpersonen kan luta sitt tänkande mot. Om det finns mycket kunskap på förhand om vilka faktorer eller situationer som kan vara problematiska i ett frågeformulär, kan en vinjett vara att föredra framför probing. Probing kan nämligen kräva många utredningsfrågor för att nå "pudels kärna" medan vinjettens referenssituation kan bidra med detsamma på ett enklare och mindre betungande sätt. Vinjetter lämpar sig bäst om det är några få begrepp eller frågor som ska utredas. Det tar för lång tid och är för omständligt att testa ett större antal med vinjetter som verktyg.

Kortsortering

Kortsortering är precis vad benämningen antyder. Testpersonen uppmanas att sortera och kategorisera objekt (till exempel ett antal begrepp) som representeras av kort. Kortet kan innehålla en beskrivande text eller bild på det som kortet representerar. Testpersonen får sedan resonera kring korten och oftast försöka "reducera" antalet kort genom att gruppera de som hör ihop på något vis. Tekniken kan ge förståelse för hur testpersonen grupperar och associerar begrepp till varandra – vad som hör ihop, hur nära olika begrepp ligger varandra, vad som är över- och undergrupper och så vidare. Antingen kan man låta testpersonerna tänka högt under sorteringen eller också kan man intervjua dem om deras sortering efteråt. Resultatet blir ett slags begreppskartläggning (ESS, 2006).

Kortsorteringen är oftast ett kvalitativt verktyg men om kortsorteringen genomförs strukturerat kan den även analyseras kvantitativt, till exempel vilka kort som tenderar att klustra tillsammans. En styrka med kortsorteringen är att den inte är lika beroende av testpersoners verbala förmåga som många av de mer intervjuinriktade teknikerna.

Kortsortering kräver mycket förarbete, administration och även efterarbete. Tekniken används inte så ofta för att testa specifika formulärfrågor utan kanske snarare i konceptualiseringsfasen när undersökningens begrepp ska definieras.

Ett viktigt metodbeslut är vilka "regler" som gäller för sorteringen. Ska testpersonen få placera korten hur som helst eller i givna grupper? Ett annat viktigt beslut är förstås vilka kort (och därmed begrepp eller ord)

som ska ingå i sorteringen. En möjlighet är också att tillåta öppen sortering – att testpersonerna själva får lägga till kort som de anser saknas.

Sammanfattande punkter:

- Syftet med kognitiva tekniker är att få insikt i respondenternas tankeprocess när de besvarar frågor för att förbättra mätningen och minska bördan.
- Tänka högt innebär att testpersonen med ord återger sitt tankeflöde. Tekniken är explorativ och kräver inte så mycket förarbete men desto mer av respondenten.
- Probing är mer en traditionell intervjuetodik, där tankeprocessen utreds genom att ställa frågor. Tekniken är mer styrd än tänka högt.
- Vinjetter är fall- eller scenariobeskrivningar av till exempel en situation eller ett beteende som testpersonen ombeds ta ställning till. Tekniken går "rakt på sak" om ett redan identifierat problem.
- Kortsortering innebär att testpersoner får gruppera kort som representerar till exempel ord eller begrepp. Resultatet blir en begreppskartläggning.

Design av ett kognitivt test

De kognitiva teknikerna beskrivna i förra avsnittet kan kombineras med varandra och genomförs i intervjuform, ett så kallat kognitivt test. Ett typiskt sådant test består av en testperson och en testledare (intervjuaren). Testpersonerna ska representera riktiga respondenter och de ska rekryteras utifrån huvudundersökningens målgrupp, alternativt om man är intresserad av att undersöka vissa undergrupper inom målgruppen. Hur många testintervjuer som ska genomföras i en omgång går inte att ange generellt. I den bästa av världar kan man fortsätta testa tills mättnad har nåtts (att de nya intervjuerna inte bidrar med någon ny meningsfull information), justera frågeformuläret och sedan testa igen. I praktiken finns dock sällan de resurserna. Det optimala antalet intervjuer varierar givetvis också beroende på frågeformuläret som man vill testa och i vilken grad man vill täcka olika möjligheter i testet. Om frågeformuläret innehåller många olika filter krävs det fler testpersoner

för att täcka in alla dessa jämfört med ett frågeformulär utan filter. Samma logik gäller om undersökningen har många subpopulationer som kan förväntas ha olika typer av problem.

I praktiken finns det dock nästan alltid ekonomiska och tidsmässiga begränsningar som blir styrande. Ofta innehåller ett kognitivt test 5–15 testpersoner men det kan variera. Metodlitteraturen hävdar också att även ett begränsat antal testintervjuer ger ett meningsfullt test (Willis, 2005). Senare forskning har dock visat att antal upptäckta problem ökar med antalet genomförda kognitiva intervjuer (Blair & Conrad, 2011). Hur många testintervjuer som ska ingå i testet är alltså inte självklart och man bör vara medveten om att ett test med få intervjuer riskerar att missa problem, även allvarliga sådana. Hur man ser på antalet intervjuer hänger också ihop med hur man ser på det kognitiva testets syfte. Vissa använder resultatet främst kvalitativt och då spelar det mindre roll i hur många intervjuer ett visst problem förekommit. Andra vill använda resultatet kvantitativt och anser att utbreddheten av problem är viktigt eftersom det ger information om hur undersökningens skattningar kan påverkas. Utifrån ett sådant synsätt krävs fler intervjuer och också mer avancerad urvals- och rekryteringsmetodik av testpersoner. Vår erfarenhet på SCB, som även stöds av många andra (bland annat Willis, 2005), är att även ett fåtal intervjuer kan ge väldigt värdefull information.

Testledare och andra inblandade måste redan före testet utreda vad i frågeformuläret de ska fokusera på. Testtiden är begränsad och man bör inte hålla på längre än max två timmar. En del av den tiden går åt till besvarandet av frågeformuläret i sig. Det finns därför inte tid att utreda allt utan man måste välja ut vissa bitar att fokusera på i testet. Val av kognitiv teknik beror på testens fokus och andra faktorer beskrivna ovan. Oavsett vilken teknik som väljs krävs en del förarbete. Probes måste formuleras, vinjetter måste skrivas eller kort måste framställas.

Testen genomförs med fördel öga mot öga, inte via telefonen. I vissa fall kan man vara tvungen att genomföra testen på telefon men det är en svaghet. Kommunikationen är mycket mer begränsad på telefon och även koncentrationen kan vara det. Vissa av teknikerna, som kortsortering, är direkt omöjliga att genomföra på telefon. Om mer detaljerade analyser än allmänna intryck ska göras bör testen spelas in och vid ännu högre ambitioner skrivas ut.

Vid många kognitiva test sker analysen av resultatet informellt. Problemen kan ofta vara ganska självklara och uppenbara när de väl

upptäckts och då är det inte nödvändigt med en djupare kvalitativ analys. I andra fall kan det krävas att testmaterialet kategoriseras och etiketteras i meningsfulla kategorier som är definierade och särskiljda från varandra.

Sammanfattande punkter:

- De kognitiva teknikerna kan användas i intervjuform med en testledare och en testperson – ett kognitivt test.
- Hur många intervjuer man ska göra varierar. Ju fler desto bättre, men även ett fåtal kan ge värdefull information.
- Vilken teknik som används och på vilken nivå resultatet ska analyseras bestäms utifrån testets syfte.

Medlyssning och beteendekodning

Syftet med medlyssning och beteendekodning

Medlyssning ("monitoring" på engelska) innebär att en tredje part lyssnar på intervjuer som genomförs i fält. Det kan ge en unik inblick i intervjusituationen. Syftet kan vara att kontrollera att intervjun har ägt rum, för att utvärdera intervjuprocessen och ge feedback till intervjuarna som ett led i intervjuarutbildningen (Biemer & Lyberg, 2003; Fowler 1995). Intervjuerna kan medlyssnas i realtid men oftast sker det utifrån inspelade intervjuer eftersom det innebär många fördelar beträffande bearbetning och analys.

Proceduren

Medlyssning kräver i praktiken ett datorstöd som automatiskt spelar in intervjuerna enligt programmerade urval (det är sällan alla intervjuer spelas in eftersom det vid större undersökningsorganisationer fort blir en ohanterlig mängd). För separata medlyssningsstudier går det att spela in intervjuerna manuellt (att intervjuerna till exempel startar en inspelning vid intervjuns start) men det är svårt att bedriva något systematiskt medlyssningsarbete på det sättet.

Syftet med medlyssning är alltså ofta inte att utvärdera mätinstrumentet men det kan även användas för det ändamålet. Analysen av inspelningarna kan vara informell och kvalitativ – att man lyssnar igenom inspelningarna (eller läser utskriften av dem) och drar slutsatser

om eventuella problem med frågorna. Detta kan vara ett snabbt sätt att utnyttja medlyssningen till att förbättra frågorna.

Ibland kan dock det finnas behov av ett mer strukturerat förfarande och kvantitativ output. Förutom medlyssning sker därför ofta en beteendekodning av det inspelade materialet (Fowler, 1995) där intervjuarens och intervjupersonens beteende kodas utifrån ett fastställt kodschema. Detta för att få en systematisk problembild av intervjusituationen (Biemer & Lyberg, 2003).

Exempel på koder är:

- läsavvikelse av intervjuaren
- respondenten avbryter intervjuaren
- intervjuaren följer upp respondentens svar.

Resultatet

Beteendekodningen kan uppmärksamma och åskådliggöra problem med frågorna men främst explicita sådana (Fowler, 2009; Schaeffer, 2009). Med "explicit" menas här observerbart för tredje person, att intervjuare eller intervjupersonens agerande i intervjun avviker från en ideal fråga-svar-sekvens. Ett tydligt exempel är att respondenten på eget initiativ uttrycker att han eller hon inte förstår frågan. Implicita problem, att respondenten till exempel har missförstått frågan utan att det påverkar fråga-svar-sekvensen, är betydligt svårare att upptäcka med beteendekodning. För det krävs andra utredande metoder, till exempel kognitiva intervjuer (Willis, 2005).

Även om beteendekodning främst använts för att flagga för problematiska frågor kan metodiken även användas för att undersöka hur frågor förstås (Holbrook, Cho & Johnson, 2006). Det krävs då en koppling mellan frågeegenskaper och beteendekoder. Senare forskning visar dessutom att beteendemässiga koder som till exempel att respondenten vänder bort blicken eller har bra flyt i sina svar kan ge meningsfull information om frågorna och mätningen (Schober, Conrad, Dijkstra & Ongena, 2012). Den här typen av resultat ökar kunskapen om hur respondenternas tankar och känslor tar sig beteendemässiga uttryck. Sådan kunskap kan vara värdefull när frågor ska utvärderas.

Beteendekodningen fokuserar på avvikelser från en ideal fråga-svar-sekvens men avvikelserna i sig behöver inte nödvändigtvis innebära problem eller resultera i mätfel (Fowler, 2009). Vad som utgör en ideal

fråga-svar-sekvens varierar mellan undersökningar, bland annat beroende på nivån av standardisering (Schaeffer, 2009). Exempelvis kan konverserande intervjuer ge högre datakvalitet än standardiserade intervjuer trots stora avvikelser från manus (läs mer i kapitel 6).

Medlyssning och beteendekodning innebär sammanfattningsvis unika möjligheter att få insyn i och undersöka eventuella problem i intervjuarundersökningar – en typ av undersökningar som är mer komplexa än enkäter eftersom de baseras på ljud (inte text) och innefattar både en respondent och en intervjuare.

Sammanfattande punkter:

- Medlyssning innebär att tredje part lyssnar på intervjuer som genomförts i fält. Det ger en god inblick i intervjusituationen.
- Intervjuerna kan analyseras informellt eller mer strukturerat genom att intervjuarens och respondentens beteenden i intervjun kodoas (beteendekodning).
- Medlyssning och beteendekodning identifierar främst explicita problem – tydliga avbrott från hur frågan idealt ska administreras.
- Hur frågan idealt ska administreras beror på nivån av standardisering.

Experiment

Orsak → effekt

Experimentet är en utmärkt metod att studera till exempel effekten av en viss frågeformulering på svarsdata i jämförelse med en annan frågeformulering. Experimentets metodik syftar just till att ge förutsättningar för att kunna dra slutsatser om en viss orsak hade en viss effekt eller var skälet till ett visst utfall. Relationen orsak-effekt kallas för kausalitet och kunskapen om sådana relationer är mycket värdefull men också svårtillgänglig. Det kan till exempel finnas många potentiella orsaker till ett visst utfall och därför vara svårt att avgöra vilken av de möjliga orsakerna som är giltig.

Vetenskapsteoretiska och metodologiska utgångspunkter

John Stuart Mill (brittisk 1800-tals filosof och ekonom) funderade mycket på kausalitet (orsak-effekt) och formulerade tre villkor som måste uppfyllas för att man ska kunna dra slutsatser om orsak-effekt (Field & Hole, 2003):

1. att orsaken ska ske före effekten
2. att det ska finnas ett samband mellan orsak och effekt
3. att alla andra förklaringar till orsak-effekt-förhållandet kan förkastas.

Det är framförallt det tredje villkoret som är svårt att uppfylla. Mill föreslog dock tillvägagångssätt för att uppfylla det: 1) att en effekt alltid ska inträffa om orsaken är närvarande och 2) att om orsaken är frånvarande ska effekten också vara det. Mill föreslog alltså att kausala relationer kan undersökas genom att studera två betingelser: en där orsaken är närvarande och en annan där orsaken är frånvarande (Field & Hole, 2003). Om man till exempel vill undersöka hur en instruktion till en viss fråga påverkar datakvaliteten, ska man alltså jämföra en betingelse där en respondentgrupp får en version där instruktionen är närvarande och en annan respondentgrupp en version där den är frånvarande. Detta upplägg är experimentets kärna. Oavsett hur avancerat det experimentella upplägget är, är alltid utgångspunkten att jämföra olika betingelser i vilka man varierar ("manipulerar" enligt experimentets terminologi) det man vill undersöka.

Även om ovanstående upplägg är kärnan i experimentell metodologi är det dessvärre inte tillräckligt. Externa faktorer kan fortfarande påverka resultatet. Exempelvis är kanske inte våra testgrupper jämbördiga eller representativa. En viss typ av individer kanske är överrepresenterade i ena gruppen. Det kan leda till ett missvisande resultat. För att kunna dra slutsatser krävs därför också (Field & Hole, 2003):

- kontroll över experimentsituationen – det enda som ska skilja mellan betingelserna är det specifika man avser att undersöka och allt annat ska hållas konstant
- grupperna ska vara i princip likvärdiga.

Systematiska skillnader mellan grupperna bör alltså med alla medel motverkas. Ett sätt att undvika systematiska skillnader är att använda slumpen, så kallad randomisering – att man slumpmässigt delar in individer i experiment- och kontrollgrupp. Om man har tillräckligt många testpersoner kommer slumpen att se till att grupperna är relativt

lika. En annan strategi är matchning – att man utifrån förkunskap om testpersonerna (till exempel registeruppgifter) ser till att grupperna är likvärdiga. Oftast baseras matchningen på variabler som man misstänker är relevanta för det man vill undersöka.

Ovanstående är experimentproceduren beskriven ur ett principiellt perspektiv (Field & Hole, 2003; Christensen, 1994).

Sammanfattande punkter:

- Experimentets metodik syftar till att ge förutsättningar för att kunna dra slutsatser om en viss orsak hade en viss effekt, var skälet till ett visst utfall (s.k. kausalitet).
- Enligt Mill krävs tre betingelser för kausalitet: 1) att orsaken ska ske före effekten, 2) att det ska finnas ett samband mellan orsak och effekt och 3) att alla andra förklaringar till orsak-effekt-förhållandet kan förkastas.
- Kausala relationer kan undersökas genom att studera två betingelser: en där orsaken är närvarande och en annan där orsaken är frånvarande.
- För att kunna uppfylla punkt 3 ovan behövs också kontroll över testsituationen och att testgrupperna är likvärdiga.
- Likvärdiga grupper kan uppnås genom randomisering eller matchning.

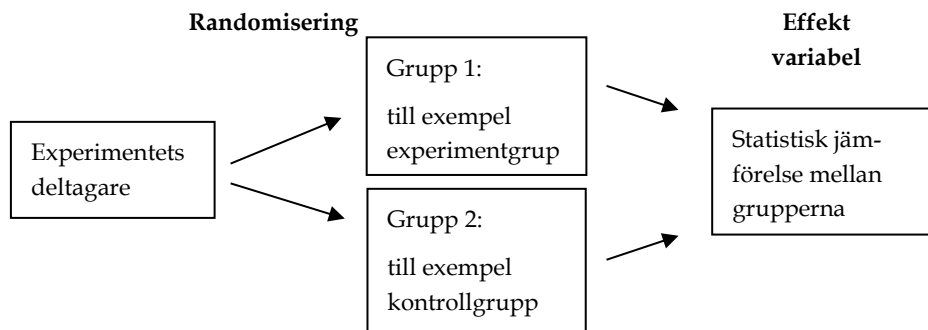
Experimentets logik och terminologi

Experimentets utgångspunkt är att man vill testa någon eller några hypoteser (Field & Hole, 2003; Christensen, 1994). Några exempel på påståenden som skulle kunna utvecklas till hypoteser är att

- en viss frågeformulering ger högre svarsfrekvens än en annan frågeformulering
- en frågas instruktion påverkar hur frågan besvaras
- ett frågeformulär som besvaras på papper ger en annan svarsfördelning än om frågeformuläret besvaras via webben.

Det alla ovanstående punkter har gemensamt är att någonting påverkar någonting annat – att något orsakar en effekt, till exempel att frågeformuleringen påverkar svarsfrekvensen. Modellen nedan visar experimentproceduren.

Figur 14.2



Experimentets deltagare delas in i grupper. För enkelhets skull beskrivs här en situation med två grupper (det finns flera olika experimentdesigner, läs nedan). Det är som sagt oerhört viktigt att detta sker slumpmässigt (så kallad randomisering) alternativt genom matchning, så att det inte råkar bli att en viss typ av individer är överrepresenterade i Grupp 1 eller 2. Om detta har utförts korrekt kan man anta att Grupp 1 och Grupp 2 är i princip lika.

De två grupperna utsätts för olika betingelser (detta kallas ibland för "behandlingar" från den medicinska forskningens "treatment"). Exempelvis kan den ena gruppen få besvara en fråga utan instruktion medan den andra besvarar frågan med en instruktion. Ett annat exempel kan vara att den ena gruppen får besvara ett frågeformulär på papper medan den andra får besvara det på webben.

Grupperna jämförs sedan statistiskt. Det man jämför dem i kallas effektvariabel (eller ibland för beroende variabel, eftersom den antas vara beroende av de olika behandlingarna, som i sin tur kallas för oberoende variabel).

Logiken är sålunda följande:

- Grupperna ska inledningsvis vara i princip lika/slumpmässiga.
- Grupperna ska behandlas exakt lika genom hela testproceduren förutom i de behandlingar där man är intresserad av att undersöka skillnaderna.
- Gruppernas respons jämförs statistiskt med effektvariabeln.

Om grupperna *var lika inledningsvis* men *skiljer sig åt efter* exponering av de olika behandlingarna, är den troligaste förklaringen att det är dessa behandlingar som orsakat skillnaderna.

Experimenten kan även göras mer komplexa genom att bland annat innefatta fler oberoende variabler (så kallad faktoriell design). Exempelvis kanske både ett nytt frågeformulär och en ny insamlingsmetod prövas i samma experiment. Den övergripande logiken är dock densamma.

Sammanfattande punkter:

- Experimentets deltagare delas upp i en experimentgrupp och en kontrollgrupp genom randomisering.
- Grupperna utsätts för olika "behandlingar" där man varierar ("manipulerar") den faktor man vill testa orsaken av.
- Grupperna jämförs sedan på ett utfallsmått ("effektvariabel").
- Om grupperna var lika innan men skiljer sig åt efter behandlingarna så är den troligaste förklaringen att det är behandlingarna som orsakat skillnaden.

Val av variabler och design

Det första man behöver göra då man planerar designen av ett experiment är att utveckla tre avgörande komponenter: 1) en hypotes, 2) de olika behandlingarna och 3) effektvariabeln. Hypotesen kan baseras på tidigare forskning eller på något man uppmärksammat i statistikproduktionen, till exempel utifrån processdata eller resultat från föregående insamling. Valet av behandlingar styrs av hypotesen. Det centrala är att de olika behandlingarna är lika i alla avseenden förutom den aspekt som man avser att testa. Om behandlingarna skiljer sig åt på många olika sätt går det inte att dra slutsatser om exakt vad som orsakat en eventuell effekt.

Nästa steg är att definiera effektvariabeln och välja design. Det är viktigt att definiera och operationalisera effektvariabeln i mätbara termer. Det är också viktigt att bestämma vilka typer av statistiska test man ska göra eftersom det är beroende av i vilken skaltyp effektvariabeln är. Vanliga statistiska analyser för att analysera resultatet från experiment är t-test och logistisk regression. Vid faktoriell design används ofta olika former av ANOVA.

En annan viktig sak att ta ställning till är hur stora grupperna (experiment- och kontrollgrupp) ska vara, det vill säga hur många individer som ska ingå i respektive grupp. Detta påverkas ofta av praktiska faktorer. Det är dyrt och kan vara svårt att rekrytera testpersoner till experimenten. Ibland används därför till exempel studenter som testpersoner och grupperna är inte sällan ganska små. Om man ska bygga in experimentet i pågående undersökningar finns det större möjligheter att ha större grupper. Generellt sett gäller att om man har stora grupper är det större chans att eventuella skillnader mellan experiment- och kontrollgrupp blir statistiskt säkerställda. Antalet individer påverkar nämligen testets "Power" (sannolikheten att urskilja skillnader mellan grupperna, om det finns några). Detta går att räkna ut på förhand och man slipper då trista konsekvenser av att man har haft för små experimentgrupper.

Sammanfattande punkter:

- Det är viktigt att noggrant designa en hypotes, behandlingarna och effektvariabeln. Annars blir det svårt att dra slutsatser från experimentet.
- Det är också viktigt att tänka på storleken på grupperna.

Experiment i praktiken

Ovanstående beskriver experimentet konceptuellt och i teorin. I praktiken finns det förstås också en hel del annat att tänka på för att allt ska flyta och för att resultatet ska bli bra. Speciellt i större organisationer kan det krävas att många olika roller och kompetenser samarbetar. Det kan vara lätt att vid skrivbordet designa finurliga experiment men det kan kräva mycket arbete för att få det att fungera i praktiken.

SCB har gett ut "[Experimentell utvärdering – en handbok](#)". Handboken finns att ladda ner från SCB:s webbplats.

Sammanfattande beskrivning ur surveyperspektiv

Experiment används sällan i explorativt syfte, till skillnad från de kvalitativa testmetoderna. Experimentet kräver dessutom en ordentlig kvantitativ datainsamling vilket i praktiken medför att det genomförs i samband med pilotundersökningar eller inbäddat i en pågående undersökning.

En annan skillnad jämfört med de kvalitativa testmetoderna är att experimentet ger ett kvantitativt resultat med ofta klara besked om en hypotes kan förkastas eller inte. Experimentet kan därför ses som en mer precis metod vars resultat kan vara lättare att sammanfatta och förmedla än resultat från kvalitativa metoder (ESS, 2006).

Om man ska utföra experimentella utvärderingar av frågorna i ett frågeformulär är det viktigt att noggrant tänka igenom vad som ska vara effektvariabel och om det krävs annat underlag för utvärderingen. Experimentet ger nämligen bara information om skillnader, inte vilken frågeformulering som är bäst. Ibland är det givet vilket av två alternativ som är mest önskvärt. Om vi till exempel vill veta om en alternativ frågeformulering ger lägre partiellt bortfall så räcker svarsandelen som utfallsmått. Men om vi är intresserade av vilken fråga som bäst mäter det vi vill mäta och får olika svarsfördelningar för två olika frågeformuleringar – vilken svarsfördelning är då bäst? Experimentet ger ingen information om det. Det krävs andra analyser och ibland också att man lägger till andra mått, som man egentligen kanske inte är intresserad av, för att kunna jämföra de två frågornas prestation. Detta är viktigt att tänka igenom på förhand.

Sammanfattande punkter:

- Experiment i statistiska undersökningar kräver ofta samarbete mellan olika personer med olika roller och kompetenser.
- Experimentet används inte explorativt utan för att testa en väl genomarbetad hypotes.
- Resultatet kan vara lättare att förmedla och basera beslut på än resultaten från kvalitativa undersökningar.
- När formulärfrågor ska utvärderas experimentellt krävs ofta att man har jämförelsemått – experimentet visar bara på skillnader, inte vad som är bäst.

Analys av data

Processdata

Processdata (eller paradata) är data som beskriver datainsamlingsprocessen och som oftast kan fås "på köpet", då det är en sidoprodukt av själva datainsamlingen. Exempel på processdata är svarstid, respondentens användning av instruktioner och intervjuarnoteringar.

Det främsta syftet med sådana data är att optimera datainsamlingen vad gäller datakvalitet, flöden och kostnader. Processdata kan även användas för att utvärdera frågeformuläret, till exempel för att få indikationer på problematiska frågor.

Processdata för att utvärdera mätinstrumentet förutsätter att datainsamlingen är datoriserad, åtminstone till viss grad. Den är därför främst relevant vid webbenkäter och intervjuer med datorstöd. Eftersom dessa insamlingsmetoder är datoriserade kan en mängd processdata samlas in automatiskt och det är då ett billigt sätt att skaffa sig information om frågeformuläret. Eftersom det är relativt enkelt att samla in mycket data är det emellertid lätt att drunkna i information. Det finns ingen anledning att samla in data som man inte har någon användning för. Det kräver alltså eftertanke vilken typ av data som ska samlas in och hur den ska tillämpas.

Processdata är ett trubbigt verktyg för att utvärdera frågor. Det är generellt sett svårt att utifrån processdata dra slutsatser om orsakerna till olika problem. Däremot är det ett kostnadseffektivt och snabbt sätt att få en överblick över vilka frågor som kan vara problematiska och som därmed kan utredas vidare med andra, skarpare testmetoder (till exempel kognitiva intervjuer).

Återstoden av avsnittet tar upp olika former av processdata.

Svarstid

Hur lång tid besvarandet tar kan vara värdefull information. Både orimligt kort och orimligt lång svarstid kan tyda på problem. Vad är då en orimlig svarstid? För att dra sådana slutsatser behövs någon form av jämförelse. Det enklaste sättet är att skapa en baslinje eller ett medelvärde på till exempel respondentgruppsnivå eller frågenivå. Utifrån denna kan sedan avvikelser studeras.

För att vara nyttig information krävs dock att data är mer specifik än total svarstid (vilket också kan vara intressant men då snarare i relation till eventuellt bortfall). Om ett sida-för-sida-upplägg används i webbenkäten kan till exempel laddningen av varje ny sida klockas. Man kan då få information om svarstiden per fråga eller per block av frågor (beroende på hur många frågor som visas på respektive sida). En väldigt snabb svarstid, jämfört med en baslinje, kan vara ett uttryck för att respondenten inte läser eller begrundar frågorna särskilt väl. Det kan

vara värt att känna till när data ska analyseras, men hur kan man få respondenter att sänka tempot?

Som frågekonstruktör kan man påverka en överdrivet snabb svarstid genom att till exempel ändra frågeformatet. Tabellfrågor är exempel på format som verkar leda till ett snabbt ifyllande och den typen kan då styckas upp i enskilda eller mindre block av frågor. Sådana förändringar kan dock öka respondentbördan och påverka jämförbarheten, så besluten bör vara väl genomtänkta.

Det är viktigt att tänka på att både kort och lång svarstid inte behöver betyda att frågan är förknippad med problem. Kort svarstid kan tyda på en enkel fråga och lång svarstid kan tyda på att respondenten begrundar frågan och tar den på allvar. Svarstider som man bedömer som orimligt korta eller långa motiverar dock en närmare granskning om det kan finnas problem.

Identifieringen och räkning av ifyllningsfel

I enkäter som använder logiska kontroller (att enkäten är programmerad att reagera om vissa villkor uppfylls eller inte uppfylls) kan i vilken grad kontrollerna aktiveras vara värdefull information. Om en enkät till exempel har en fråga där respondenten ska fördela 100 procent på olika aktiviteter och en kontroll ofta utlöses på grund av att respondenterna misslyckas med att totalsumman av de olika aktiviteterna blir exakt hundra, kan det signalera att uppgiften är för avancerad och kanske bör förenklas.

Navigering genom frågeformuläret

Hur respondenterna navigerar genom frågeformuläret kan vara bra information i förbättringsarbetet med frågorna. Hoppas de till exempel fram och tillbaka i frågeformuläret? I vilken ordning besvarar de frågorna? Hur lång tid tar det att komma vidare? Detta är bara några exempel på saker som kan vara värda att känna till och ibland utforska djupare.

Ändring av svar

Om många respondenter ändrar sina svar kan det vara ett tecken på att något är fel. Exempelvis att frågan är svår att förstå, att frågan påverkas av senare frågor som gör att respondenterna går tillbaka och ändrar sina svar eller att frågan är känslig på något sätt. Det är alltså en indikation på att frågan bör utredas.

Användning av instruktioner och annan hjälpinformation

Utrymmet på skärmen är begränsat, ofta *för* begränsat för all information som man vill ska kunna finnas tillgänglig, speciellt i komplicerade undersökningar. En kompromiss är då att göra informationen tillgänglig men endast om respondenten känner behov av det. Det sker oftast via länkar och popup-fönster. I vilken utsträckning sådana klickbara instruktioner används kan ge värdefull information om frågorna i frågeformuläret.

Intervjuarnoteringar

Intervjuundersökningar har en stor fördel jämfört med webbenkäter – det finns en annan människa med under besvarandet. Intervjuaren får förstås massor av information om hur frågorna fungerat när de intervjuar respondenter. Det är dock viktigt att uppmärksamma att intervjuaren, precis som respondenten och alla andra människor, kan missförstå eller göra felaktiga tolkningar om vad som utgör ett problem i en fråga. Om intervjuarnas erfarenheter samlas in på bred front och på ett systematiskt sätt utgör de dock en bra informationskälla om frågeformuläret. Erfarenheter kan samlas in genom att lämna utrymme för noteringar och kommentarer i intervjuarnas datorstöd. Ett annat sätt är att genomföra fokusgrupper med intervjuarna efter att de genomfört sina intervjuer.

Sammanfattande punkter:

- Processdata är data som beskriver datainsamlingsprocessen.
- Processdata kan användas för att utvärdera mätinstrumentet eller flagga för frågor som kan vara problematiska.
- Exempel på processdata är: svarstiden för respondenterna, räkne- och ifyllningsfel i enkäten, ändring av svar, navigering fram och tillbaka i frågeformuläret och användning av instruktioner.

Svarsdata

Det ultimata testet hur frågor presterar är att analysera svarsdata. Kvalitativa tester kan ge god förhandsinformation och kan vara ett kostnadseffektivt sätt att få en inblick, men i slutänden är det svarsdata som är det viktiga. Det är kvaliteten i svarsdata som avgör statistikens

kvalitet. Om svarsdata finns tillgängligt är det en utmärkt utgångspunkt för att utvärdera frågorna.

Svarsdata kan analyseras på olika sätt, till exempel enkelt eller avancerat. Enkla analyser av svarsdata innebär att man granskar fullständigheten och misstänkta felaktigheter i data. Exempelvis kan man studera utbreddheten av partiellt bortfall, ogiltiga svar och avbrutna sessioner. Man kan även studera innehållet på öppna svar för att fånga upp kommentarer och om till exempel några svarsalternativ borde läggas till som fasta svarsalternativ. Ett ytterligare tillvägagångssätt är att undersöka relationen mellan frågor med korrelation eller korstabell. Hänger frågor ihop såsom man kan förvänta sig baserat på förnuft, erfarenhet eller teori? Har respondenten svarat logiskt och konsekvent? Avvikelse behöver inte innebära att något är fel men det är ofta ett tillräckligt skäl för att gräva djupare.

Mer avancerade analysmetoder är att använda statistiska modeller eller annan data för att ta reda på hur frågorna fungerat. Faktoranalys och reliabilitetsanalys med till exempel Cronbach's Alpha är exempel på metoder som analyserar sambandet mellan frågor. Det finns många fler men detta område ligger utanför denna boks avgränsning (Biemer, Groves, Lyberg, Mathiowetz & Sudman, 1991).

Det starkaste testet av svarsdata är nog att jämföra det med det "sanna värdet". Problemet är att det "sanna värdet" allt som oftast är okänt. Om det vore känt skulle det inte vara lika stor vits att fråga om det i en surveyundersökning. Det händer dock att man har tillgång till andra data som man har så pass hög tilltro till att det kan fungera som en utgångspunkt (till exempel registeruppgifter). I sådana fall utgör det ett ypperligt underlag att jämföra svarsdata mot.

Sammanfattande punkter:

- Svarsdata kan användas för att utvärdera formulärfrågor – både genom enkla och avancerade analyser.
- Enkla analyser: 1) utbreddheten av partiellt bortfall, ogiltiga svar och avbrutna sessioner och 2) om respondenternas svar på frågor hänger ihop som de borde.
- Avancerade analyser: 1) reliabilitetsanalyser, 2) validitetsanalyser och 3) andra statistiska modeller.

Återkontakter

För att kunna skatta avvikelser från det sanna värdet (det så kallade mätfelet) i en statistisk undersökning krävs det att det sanna värdet är känt. Detta är sällan fallet. Hur ska man då göra? Ett sätt är att införskaffa de sanna värdena, eller värden som man har högre tilltro till än dem i originalundersökningen, till exempel genom en separat datainsamling. Ett sådant sätt är återintervjun. För ett exempel, se SCB:s mätfelsstudie i Arbetskraftsundersökningen (AKU) från 2014. [Rapporten](#) finns att ladda ner från SCB:s webbplats.

Återintervjuer

Metodiken för återintervju utvecklades i USA på 1940-talet, framförallt vid Censusbyrån (Biemer & Forsman, 1992). Metoden har även använts vid SCB för flera av de stora individ- och hushållsundersökningarna (Wärneryd, 1989; Eriksson, 1989; Karlsson & Thudin, 1989).

Återintervju innebär, precis som det låter, att intervjuer genomförs ytterligare en gång utöver originalinsamlingen. Återintervjun kan ha flera syften, till exempel att kontrollera att intervjuerna ägt rum eller för att utvärdera fältarbetet (Forsman & Schreiner, 1991). Oftast är dock syftet att utvärdera mätningen.

Det finns två huvudtyper av återintervjustudie för utvärdering av mätningen (Biemer & Lyberg, 2003):

- replikering
- gold standard.

Replikering

Syftet med den replikerande intervjun är att utvärdera tillförlitligheten (reliabilitet) och slumpmässiga fel. För att kunna skatta tillförlitligheten på ett bra sätt krävs det att svarsprocessen är i princip identisk vid de två mätningarna. Målet med den replikerande varianten av återintervju är alltså att återskapa ordinarie intervju. Därför följer den ofta följande principer: 1) återintervjuerna ska vara på "samma nivå" som intervjuerna i originalintervjun, 2) proceduren ska så långt som möjligt vara identisk, 3) frågeformuläret ska så långt som möjligt vara exakt detsamma, 4) det ska vara samma respondenter och 5) originalintervju och återintervju ska vara oberoende. Tiden mellan intervjuerna ska vara tillräckligt kort för att undvika minnesfel men tillräckligt lång för att undvika inlärningseffekter

(att respondenterna minns frågorna och i värsta fall även sina svar) (Forsman, 1987; Biemer & Lyberg, 2003).

Gold standard

Återintervjun kan också närma sig en validitetsprövning – att man strävar efter att ytterligare höja kvaliteten på svaren för att få säkrare värden att jämföra de ordinarie svaren med (Bergman & Wärneryd, 1982). Denna variant kallas för gold standard. Målet är alltså att genomföra en mätning som ger sanna värden, eller värden som man har högre tilltro till än originalmätningen, för att kunna skatta systematiska fel.

För att kunna uppnå en gold standard krävs det att återintervjun har förbättrade procedurer jämfört med originalintervjun. Till skillnad från i den replikerande varianten ska alltså inte utförandet av återintervjuerna vara så snarlikt originalintervjuerna som möjligt utan tvärtom förbättrat (Biemer & Lyberg, 2003; Biemer & Forsman, 1992). Detta kan innebära att man använder mer kvalificerade intervjuare eller att man ger intervjuerna en särskild utbildning. Ramarna för intervjun kan också förbättras genom längre fältarbetsperiod och intervjutid. Det kan också innebära att man gör avvikelser i frågeformuläret, jämfört med originalinsamlingen, för att få en säkrare mätning. Forsman skriver:

”En återintervju innebär inte nödvändigtvis att man ställer samma fråga som vid första intervjutillfället. Man kan till exempel i återintervjun ersätta en fråga med två eller flera andra i syfte att eliminera risken för missförstånd och då få ett svar som förmodas ligga sanningen närmare” (Forsman, 1987, s. 10).

Det är inte heller nödvändigt att hela originalintervjun ingår i återintervjun utan fokus kan istället ligga på några nyckelvariabler.

Rekonciliering

Gold standard-varianten kan kompletteras med så kallad rekonciliering. Det innebär att intervjuerna utreder orsakerna till eventuella avvikelser mellan svaren i originalintervju och återintervju (Morton, Mullin & Biemer, 2008; Biemer & Lyberg, 2003). Rekoncilieringen är alltså en fristående del av intervjun som genomförs efter att återintervjun avslutats och endast om den avviker från ordinarie intervju. Syftet är främst att utreda orsaker till avvikelserna, vilket är värdefull information i förbättringsarbetet med undersökningen. Men det kan också vara ett

sätt att ytterligare försäkra sig om att återintervjun verkligen är korrekt, en gold standard (Forsman & Schreiner, 1991).

Rekonciliering kräver alltså att intervjuarna har tillgång till originalintervjuns svar men helst inte innan återintervjun är färdig eftersom det annars kan påverka utförandet och därmed resultatet. (Det finns även varianter utan att intervjuarna har tillgång till originalintervjuns svar, vilket kallas för oberoende rekonciliering. Syftet är då ett annat och påminner mer om en debriefing (Forsman, 1987)).

Rekoncilieringsintervjun är ofta en mjukare intervju som sker genom samtal snarare än standardiserat (Forsman, 1987; Bergman & Wärneryd, 1982). Rekoncilieringen ställer alltså högre krav på intervjuaren.

Det finns risker med rekonciliering: 1) att respondenten kan medvetet svara felaktigt på ett systematiskt sätt, 2) respondenten kan försöka "försvara" sina originalsvar för att framstå i bättre dager, 3) problem i kommunikationen mellan respondent och den som utför rekoncilieringen, 4) inlärningseffekter – att respondenten försöker svara som "förra gången" och 5) att intervjuaren medvetet eller omedvetet "tonar ner" avvikelser.

Många av dessa risker kan pareras genom undersökningsdesignen, förarbete och utbildning.

Sammanfattande punkter:

- Återkontakter är ett sätt att få information som undersöknings-svaren kan jämföras med.
- Återintervjun är en form av återkontakt.
- Återintervjun kan vara en replikering (exakt lika) av originalintervjun.
- Återintervjun kan också genomföras med en förbättrad procedur för att få så "sanna" svar som möjligt, en så kallad gold standard.
- Återintervjun kan kompletteras med en utredande rekonciliering för att få förståelse om orsakerna till skillnader i svar mellan originalintervjun och återintervju.

Kombinationer av testmetoder

Litteraturen om hur olika testmetoder överlappar eller kompletterar varandra är fattig (Yan, Kreuter & Tourangeau, 2012). Beslutet om vilken eller vilka metoder som passar bäst för en specifik undersökning är därför en bedömningsfråga. I praktiken är dessutom test av frågeformulär inte bara en metodologisk fråga, utan även en resursfråga (Persson, Björnram, Elvers & Erikson, 2015). Några avslutande tips är dock:

- Fundera över syftet med och behovet av utvärderingen. Behövs förståelse för bakomliggande orsaker till ett problem eller ett numeriskt mått på en frågas prestation?
- Finns det redan tillgänglig information som kan användas? Det är ett kostnadseffektivt sätt att åtminstone få indikationer på vad senare testmetoder kan fokusera på.
- I den mån flera testmetoder ska kombineras kan metoder som är empiriska och de som baseras på individuella bedömningar vara bra komplement till varandra.
- Det finns en naturlig kronologi mellan en del testmetoder (läs mer i kapitlets inledning).

Sammanfattning

Att utvärdera frågor – sammanfattning

Olika metoder för att utvärdera frågor:

- Det finns många skäl att utvärdera sina formulärfrågor (se punktlista ovan).
- Det finns många olika utvärderingsmetoder och de är naturligt knutna till olika faser i en undersökning (se figur ovan).
- Metoderna varierar också bland annat i resurskrav, om de är kvalitativa eller kvantitativa, om de ger ett statistiskt säkerställt värde eller en djupare förståelse.

Expertgranskning:

- Expertgranskning innebär att en tredje part, experter på frågekonstruktion, granskar frågeformuläret på ett systematiskt sätt.

- Syftet med expertgranskning är att identifiera möjliga problem och föreslå förbättringar.
- Expertgranskningen kan ske ostrukturerat och informellt eller mer strukturerat utifrån en checklista eller kodningsmall.
- Utvärderingar av metoden visar att den är framgångsrik när det gäller att identifiera problem, även om det kan variera vilka problem olika granskare uppmärksammar.

Kognitiva tekniker:

- Syftet med kognitiva tekniker är att få insikt i respondenternas tankeprocess när de besvarar frågor för att förbättra mätningen och minska bördan.
- Tänka högt innebär att testpersonen med ord återger sitt tankeflöde. Tekniken är explorativ och kräver inte så mycket förarbete men desto mer av respondenten.
- Probing är mer en traditionell intervjuetodik, där tankeprocessen utreds genom att ställa frågor. Tekniken är mer styrd än tänka högt.
- Vinjetter är fall- eller scenariobeskrivningar av till exempel en situation eller ett beteende som testpersonen ombeds ta ställning till. Tekniken går "rakt på sak" om ett redan identifierat problem.
- Kortsortering innebär att testpersoner får gruppera kort som representerar till exempel ord eller begrepp. Resultatet blir en begreppskartläggning.
- De kognitiva teknikerna kan användas i intervjuform med en testledare och en testperson – ett kognitivt test.
- Hur många intervjuer man ska göra varierar. Ju fler desto bättre, men även ett fåtal kan ge värdefull information.
- Vilken teknik som används och på vilken nivå resultatet ska analyseras bestäms utifrån testets syfte.

Medlyssning och beteendekodning:

- Medlyssning innebär att tredje part lyssnar på intervjuer som genomförs i fält. Det ger en god inblick i intervjusituationen.
- Intervjuerna kan analyseras informellt eller mer strukturerat genom att intervjuarens och respondentens beteenden i intervjun kodalas (beteendekodning).

- Medlyssning och beteendekodning identifierar främst explicita problem – tydliga avbrott från hur frågan idealt ska administreras.
- Hur frågan idealt ska administreras beror på nivån av standardisering.

Experiment:

- Experimentets metodik syftar till att ge förutsättningar för att kunna dra slutsatser om en viss orsak hade en viss effekt, var skälet till ett visst utfall (s.k. kausalitet).
- Enligt Mill krävs tre betingelser för kausalitet: 1) att orsaken ska ske före effekten, 2) att det ska finnas ett samband mellan orsak och effekt och 3) att alla andra förklaringar till orsak-effekt-förhållandet kan förkastas.
- Kausala relationer kan undersökas genom att studera två betingelser: en där orsaken är närvarande och en annan där orsaken är frånvarande.
- För att kunna uppfylla punkt 3 ovan behövs också kontroll över testsituationen och att testgrupperna är likvärdiga.
- Likvärdiga grupper kan uppnås genom randomisering eller matchning.
- Experimentets deltagare delas upp i en experimentgrupp och en kontrollgrupp genom randomisering.
- Grupperna utsätts för olika "behandlings" där man varierar ("manipulerar") den faktor man vill testa orsaken av.
- Grupperna jämförs sedan på ett utfallsmått ("effektvariabel").
- Om grupperna var lika innan men skiljer sig åt efter behandlingarna så är den troligaste förklaringen att det är behandlingarna som orsakat skillnaden.
- Det är viktigt att noggrant designa en hypotes, behandlingarna och effektvariabeln. Annars blir det svårt att dra slutsatser från experimentet.
- Det är också viktigt att tänka på storleken på grupperna.
- Experiment i statistiska undersökningar kräver ofta samarbete mellan olika personer med olika roller och kompetenser.
- Experimentet används inte explorativt utan för att testa en väl genomarbetad hypotes.

- Resultatet kan vara lättare att förmedla och basera beslut på än resultaten från kvalitativa undersökningar.
- När formulärfrågor ska utvärderas experimentellt krävs ofta att man har jämförelsemått – experimentet visar bara på skillnader, inte vad som är bäst.

Analys av data – processdata:

- Processdata är data som beskriver datainsamlingsprocessen.
- Processdata kan användas för att utvärdera mätinstrumentet eller flagga för frågor som kan vara problematiska.
- Exempel på processdata är: svarstiden för respondenterna, räkne- och ifyllningsfel i enkäten, ändring av svar, navigering fram och tillbaka i frågeformuläret och användning av instruktioner.

Svarsdata:

- Svarsdata kan användas för att utvärdera formulärfrågor – både genom enkla och avancerade analyser.
- Enkla analyser: 1) utbreddheten av partiellt bortfall, ogiltiga svar och avbrutna sessioner och 2) om respondenternas svar på frågor hänger ihop som de borde.
- Avancerade analyser: 1) reliabilitetsanalyser, 2) validitetsanalyser och 3) andra statistiska modeller.

Återkontakter:

- Återkontakter är ett sätt att få information som undersöknings-svaren kan jämföras med.
- Återintervjun är en form av återkontakt.
- Återintervjun kan vara en replikering (exakt lika) av originalintervjun.
- Återintervjun kan också genomföras med en förbättrad procedur för att få så "sanna" svar som möjligt, en så kallad gold standard.
- Återintervjun kan kompletteras med en utredande rekonciliering för att få förståelse om orsakerna till skillnader i svar mellan originalintervjun och återintervju.

Kombinationer av testmetoder:

- Fundera över syftet med och behovet av utvärderingen. Behövs förståelse för bakomliggande orsaker till ett problem eller ett numeriskt mått på en frågas prestation?
- Finns det redan tillgänglig information som kan användas? Det är ett kostnadseffektivt sätt att åtminstone få indikationer på vad senare testmetoder kan fokusera på.
- I den mån flera testmetoder ska kombineras kan metoder som är empiriska och de som baseras på individuella bedömningar vara bra komplement till varandra.
- Det finns en naturlig kronologi mellan en del testmetoder (läs mer i kapitlets inledning).

Ordlista

Undersökning

Undersökningar – att information om ett sakproblem samlas in och sedan analyseras – kan utföras på många olika sätt. Fokus i denna bok ligger på statistiska undersökningar med frågeformulär som antingen administreras i enkätform eller som intervju.

Survey

Survey är ett engelskt ord som även brukar användas på svenska för en statistisk undersökning. Oftast är det den typ av undersökning i vilken man drar ett urval från en population för att med statistiska metoder kunna uttala sig om populationen.

Enkät

I denna bok definieras enkät som en form av frågeformulär som respondenten administrerar själv, oftast i pappersform eller på webben. En enkät är skriftlig och består av visuell information – respondenten tillgodogör sig information genom läsning (till skillnad från i en intervju).

Intervju

Det finns många olika varianter av intervjuer men denna bok utgår från intervjuer i statistiska undersökningar som genomförs med hjälp av en intervjuare och ett intervjuformulär. Intervjuer kan variera i hur pass standardiserade de är men vanligtvis innehåller de till hög grad frågor med fasta svarsalternativ. Intervjuer till statistiska undersökningar utförs oftast via besök eller telefon. Till skillnad från i enkäter är det en intervjuare, inte respondenten själv, som administrerar frågorna. Respondenten tillgodogör sig informationen genom att lyssna på intervjuarens läsning istället för att läsa själv.

Frågeformulär

Frågeformulär används i denna bok som den mest generella benämningen av mätinstrument i en surveyundersökning. När endast endera av självadministrerad enkät eller intervjuformulär åsyftas i texten används de mer specifika beteckningarna – enkät eller intervjuformulär).

Vid mer generella diskussioner, som inte kräver någon åtskillnad mellan enkät och intervju, används "frågeformulär" för att inkludera dem båda.

Fråga

I denna bok är en fråga en språkhandling *riktad till respondenten med målet att få information som respons*. Frågor kan formuleras på olika sätt grammatiskt, till exempel kan de ställas som påståenden. I boken syftar ibland "frågor" och "fråga" på hela frågekomplexet med frågeformulering, instruktioner och svarsalternativ men främst syftar det på själva frågeformuleringen som respondenten ska ta ställning till.

Instruktion

En instruktion är tilläggsinformation till frågeformuleringen. Den kan bestå av en definition av ett viktigt begrepp eller regler för hur frågan ska besvaras.

Svarsalternativ

Svarsalternativen kommer alltid efter frågan och utgör de valmöjligheter som respondenten har till sitt förfogande för att rapportera sitt svar. Svarsalternativ kan vara öppna och syftar då till att respondenten ska formulera sitt svar på egen hand. Vanligare är dock att svarsalternativen är slutna i form av fasta beskrivningar. Respondentens roll är att välja vilken av beskrivningarna som passar bäst.

Skaltyper

Svarsalternativen kan vara av olika skaltyper. Med det menas att mätnivån är olika vad gäller om svarsalternativen kan särskiljas, rangordnas och om måttenheten är konstant mellan svarsalternativen. De olika skaltyperna är:

- **Kvotskala:** Svarsalternativen kan särskiljas, rangordnas och måttenheten är konstant. Det finns en absolut nollpunkt.
- **Intervallskala:** Svarsalternativen kan särskiljas, rangordnas och måttenheten är konstant.
- **Ordinalskala:** Svarsalternativen kan särskiljas och rangordnas.
- **Nominalskala:** Svarsalternativen kan särskiljas.

Respondent

Respondenten är mottagaren av frågeformuläret, den person vi ställer frågor till. Olika benämningar på denna roll finns i litteraturen, till exempel uppgiftslämnare eller svarsperson. I denna bok används "respondent" (vilket också är den engelska beteckningen).

Datainsamlingsmetod

Datainsamlingsmetoden är det tillvägagångssätt vi använder för att administrera frågeformuläret. Det är det medium via vilket respondenten svarar. Denna bok fokuserar på framförallt fyra datainsamlingsmetoder: pappersenkät, webbenkät, telefonintervjuer och besöksintervjuer.

Konversation/kommunikation

Denna bok har ett kommunikationsperspektiv. Kommunikation förutsätter att det finns en sändare, ett meddelande, ett medium för meddelandet och en mottagare. Översatt till denna boks ämnesområde är frågeställaren sändaren, frågeformuläret och dess frågor är meddelandet, datainsamlingsmetoden är mediet och respondenten är mottagaren.

Konversation syftar i denna bok specifikt på språklig kommunikation mellan människor, både i skrift och i tal.

Kognition

Kognition är ett viktigt perspektiv i denna bok. Det är ett samlingsnamn för mentala processer som har med informationshantering att göra. Det innefattar processer som vad vi uppmärksammar, hur vi förstår språk, hur minnet fungerar och hur vi drar slutsatser och tar beslut. Den kognitiva psykologin är ett av de främsta teoretiska ramverken för frågekonstruktion och förekommer också i denna bok.

Referenser

- Alwin, D. (2007). *Margins of Error: A Study of Reliability in Survey Measurement*. New York: Wiley.
- Alwin, D. F. & Krosnick, J. A. (1985). The measurement of values in surveys: A Comparison of ratings and rankings. *Public Opinion Quarterly* 49.
- Alwin, D. F. (1992). Information transmission in the survey interview: Number of response categories and the reliability of attitude measurement. *Sociological Methodology*.
- Alwin, D. F. & Krosnick, J. A. (1991). The reliability of survey attitude measurement: The influence of question and respondent attributes. *Sociological Methods and Research* , 20.
- Anderson, L. W., Krathwohl, D- R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., & C. Wittrock, M. (Red.). (2001). *A taxonomy for learning, teaching and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman, Inc.
- Andersson, L. (2011). *Mått på brott. Självdeklaration som metod att mäta brottslighet*. Akademisk avhandling, Kriminologiska institutionen, Stockholms universitet.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E. Pintrich, P. R., et al. (Eds.). (2001). *A taxonomy for learning, teaching and assessing. A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman, Inc.
- Andreasen, A.R. (1970). Personalizing mail questionnaire correspondance, *Public Opinion Quarterly*, 34(2), s. 273-277.
- Andrews, F. M. (1984). Construct validity and error components of survey measures: A structural modeling approach. *Public Opinion Quarterly*, 48, 409-442.
- Antil, J.H. (1984), "Conceptualization and Operationalization of Involvement". *Advances in Consumer Research*, Vol. 11, 1984
- Apodaca, R., Lea, S., & Edwards, B. (1998). The Effect of Longitudinal Burden on Survey Participation. *American Statistical Association Proceedings of the Survey Research Methods Section*, s. 906-10.

- Armstrong, J. S. (1975). Incentives in mail surveys. *Public Opinion Quarterly*, 39(1), s. 111–116.
- Ashcraft, M. H. (1994). *Cognition*. NJ: Prentice Hall Inc.
- Bavdaz, M. (2007). *Measurement errors and the response process in business surveys*. Avhandling. University of Ljubljana.
- Bayer, L. R., & Thomas, R. K. (2004, August). A comparison of sliding scales with other scale types in online surveys. *Paper presented at the RC33 International Conference on Social Science Methodology*, Amsterdam.
- Beatty, P., Fowler, F. J. & Cosenza, C. (2007). *Do Questionnaire Recommendations Lead to Measurable Improvements? Some Experiments with Alternate Versions of Complex Survey Questions*. Conference presentation, EARA.
- Belson, W. A. (1981). *The design and understanding of survey questions*. Gower Publishing Company, Aldershot, UK.
- Bergman, L., & Wärneryd, B. (1982). Om datainsamling i surveyundersökningar: vilken metod är bäst och vad får vi egentligen veta? Stockholm: Statistiska centralbyrån (SCB)
- Biemer, P. P. & Lyberg, L. E. (2003). *Introduction to Survey Quality*. NJ. Wiley.
- Biemer, P. P. & Forsman, G. (1992). On the Quality of Reinterview Data with Application to the Current Population Survey. *Journal of the American Statistical Association*. 87, s. 915–923.
- Biemer, P., Groves, R., Lyberg, L., Mathiowetz, N., & Sudman, S. (1991). *Measurement errors in surveys*. New York: John Wiley & Sons.
- Bilgen, I. & Belli, R. F. (2010). Comparison of Verbal Behaviors between Calendar and Standardized Conventional Questionnaires. *Journal of Official Statistics*, 26(3): s. 481–505.
- Björkman, N. (2000). *Fyra metodologiska teman*. Stockholm.
- Björnram, A., Erikson, J., Furubjelke, E., Hartwig, P., Henningsson, B., & Lorenc, B. (2007). *Riktlinjer för design av frågor och instruktioner och mallar för layout i företagsundersökningar*. SCB-rapport.
- Blair, J., & Conrad, F. (2011). Sample size for cognitive interviewing. *Public Opinion Quarterly*, 75, s. 636–658.
- Blasius, J. (2012). Comparing Ranking Techniques in Web surveys. *Field Methods*, 24.

- Blumer, H. (1969). *Symbolic interactionism: perspective and method*. Berkeley: University of California Press.
- Bohm, P. (1988). Ger opinionsundersökningar om den offentliga sektorn någon information? *Ekonomisk debatt*, 8, s. 605–614.
- Boulianne, S., Klofstad, C. A., & Basson, D. (2010). Sponsor prominence and response pattern to an online survey. *International Journal of Public Opinion Research*, 23(1).
- Bradburn N; Sudman S; Blair E & Stocking C (1978). Question threat and response bias. *Public Opinion Quarterly*, 42, 221-234.
- Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: the definitive guide to questionnaire design--for market research, political polls, and social and health questionnaires*. CA: Wiley.
- Bradburn, N. M. & Danis, C. (1984). Potential contributions of cognitive research to survey questionnaire design. In Jabine, T. M., Tanur, J. M. & Tourangeau, R. (eds.) *Cognitive aspects of survey methodology: Building a bridge between disciplines*. National Academy Press, Washington DC.
- Bradburn N. M., & Miles C. (1979). Vague quantifiers. *Public Opinion Quarterly*, 43.
- Bradburn, N. & Sudman, S. & Associates. (1979) *Improving Interview Method and Questionnaire Design: Response Effects to Threatening Questions in Survey Research*. Josse-Bass, San Francisco, CA.
- Brennan, M. & Charbonneau, J. (2009). Improving mail survey response rates using chocolate and replacement questionnaires. *Public Opinion Quarterly*. 73(2). s. 368–378.
- Brottsförebyggande rådet (2010 a). *Brott bland ungdomar i årskurs nio. Resultat från skolundersökningen om brott åren 1995–2008*. Rapport 2010:6. Stockholm: Brottsförebyggande rådet.
- Brülde, B. (2007). *Lycka och lidande: begrepp, metod och förklaring*. Lund: Studentlitteratur.
- Buskirk, T.D. (2015). Survey research using emerging technologies: considerations for design, data collection, sampling and recruitment for smart phone and tablet based surveys. Kurs. JOS conference. Stockholm.

- Buskirk, T. D. & Andrus, C. (2012). Smart surveys for smartphone: exploring various approaches for conducting online mobile surveys via smartphones. *Survey Practice*. Available at:<http://surveypractice.wordpress.com/2012/02/21/smart-surveys-for-smart-phones/>
- Burton, J., Laurie, H. & Lynn, P. (2006). The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *Journal of Royal Statistical Society*. 169 (3), s. 459–478.
- Burton, S & Blair, E. (1991). "Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys." *Public Opinion Quarterly* 55: s. 50–79.
- Callegaro, M. (2008). Seam Effects in Longitudinal Surveys, *Journal of Official Statistics*, 24 (3), s. 387–409.
- Cannell, C.F., Fowler, F.J., & Maequis, K.A. (1968). The influence of interviewer and respondent psychological and behavioral variables on the reporting in household interviews. Vital and health statistics – series 2, no 26, Public health service publication No 1000.
- Cannell, C., Miller, P., & Oksenberg, L. (1981). Research on interviewing techniques. I Leinhardt, S. (Ed.), *Sociological methodology* (s. 389– 437). San Francisco: Jossey-Bass.
- Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. London: SAGE Publications.
- Chang, L., & Krosnick, J. A.: (2009). National Surveys via RDD Telephone Interviewing versus the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly* 73(4): s. 641–78.
- Christian, L.M., Dillman, D.A., Smyth, J.D., 2008. The effects of mode and format on answers to scalar questions in telephone and web surveys. In: Lepkowski, J., Tucker, C., Brick, M., de Leeuw, E., Japac, L., Lavrakas, P., Link, M., Sangster, R. (Eds.), *Advances in Telephone Survey Methodology*. Wiley-Interscience, New York, pp. 250–275 (Chapter 12).
- Christianson, A. & Polfeldt, T. (1996). Response quality improvement initiatives at Statistics Sweden. *Proceedings of the 2nd International Conference on Methodological Issues in Official Statistics*. Stockholm.
- Christensen, L. B. (1994). *Experimental Methodology*. MA: Allyn and Bacon.

- Church, A. H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly*, 57, s. 62–79.
- Cialdini, R. B., Cacioppo, J. T., Bassett, R. & Miller, J. A. (1978). Low-Ball Procedure for Producing Compliance Commitment then Cost. *Journal of Personality and Social Psychology*, 36, s. 463–76.
- Clark, H. H. & Schober, M. F. (1992). Asking questions and influencing answers. In Tanur, J. M. (ed.) *Questions about questions*. NY: Russel Sage Foundation.
- Clark, H. H., Schreuder, R. & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22, s. 245–258.
- Clement, S. L., & Shamshiri-Petersen, D. (2013). Examination of a 'Web Mode Effect': An Experimental Comparison of Web and Paper Based Surveys. *Unpublished manuscript*.
- Cliff, N. (1959). Adverbs as multipliers. *Psychological Review* 66.
- Cook, C., Heath, F. & Thompson, R. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. *Educational & Psychological Measurement*, 60 (6), s. 821–36.
- Colasanto, D., Singer, E., & Rogers, T. F. (1992). Context Effects on Responses to Questions about Aids. *Public Opinion Quarterly* 56: s. 515–518.
- Collins, M, Sykes, W., Wilson, P. & Blackshaw, N. (1988). Nonresponse: The UK Experience. I *Telephone Survey Methodology*, red. Groves, R. M., Biemer, P. P., Lyberg, L. E., Massey, J. T. Nicholls, II, W. L. & Waksberg, J., s. 213–231. New York: Wiley.
- Conrad, F. G., Couper, M. P., Tourangeau, R. & Peytchev, A. (2006). Use and Non-use of Clarification Features in Web Surveys. *Journal of Official Statistics* 22: s. 245–269.
- Conrad, F. G. & Schober, M. F. (2005). Promoting Uniform Question Understanding in Today's and Tomorrow's Surveys. *Journal of Official Statistics*, vol.21, no. 2, s. 215–231.
- Conrad, F. G. & Schober M. F. (2000). Clarifying question meaning in a household telephone survey. *Public Opinion Quarterly*. 64.
- Converse, J. M. & Presser, S. (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills: Sage.

- Converse, P. (1970). Attitudes and non-attitudes: Continuation of dialogue. In Tufte, E. (Ed.), *The quantitative analysis of social problems*. Reading, MA: Addison-Wesley.
- Couper, M. P. (2009). The role of Internet in mixed-mode surveys. *Course material*. Stockholm.
- Couper, M. P. (2008). *Designing effective web surveys*. NY: Cambridge.
- Couper, M. P. (2005). Technology Trends in Survey Data Collection. *Social Science Computer Review*, 23 (4), s. 486–501.
- Couper, M. P. (1997). Survey introductions and data quality. *Public Opinion Quarterly*, 61, s. 317–338.
- Couper, M. P., Kennedy, C., Conrad, F. G. & Tourangeau, R. (2011). Designing Input Fields for Non-Narrative Open-Ended Responses in Web Surveys. *Journal of Official Statistics*, 27(1), s. 65–85.
- Couper, M. P., Conrad, F. G. & Tourangeau, R. (2007). Visual Context Effects in Web Surveys. *Public Opinion Quarterly*, 71 (4), s. 623–634.
- Couper, M. P., Tourangeau, R., & Conrad, F. G. (2006). Evaluating the Effectiveness of Visual Analog Scales. *Social Science Computer Review*, 24 (2), s. 227–245.
- Couper, M.P., Tourangeau, R. & Kenyon, K. (2004). Picture This! An Analysis of Visual Effects in Web Surveys. *Public Opinion Quarterly* 68(2): s. 255–266.
- Couper, M. P., Tourangeau, R., Conrad, F. G., & Crawford, S. (2004). What They See Is What We Get. *Social Science Computer Review*, 22 (1), s. 111–127.
- Couper, M. P., Tourangeau, R., Conrad, F. G., & Crawford, S. (2004). What they see is what we get: Response options for web surveys. *Social Science Computer Review*, 22(1), s. 111–127.
- Couper, M.P., Traugott, M.W. & Lamias, M.J. (2001). Web Survey Design and Administration. *Public Opinion Quarterly*, 65, 230–53.
- Cronbach, L.J. (1988). "Five perspectives on the Validity Argument" I Wainer, B. & Braun, H. I. (red.) *Test Validity*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc. Publishers. Messick, 1989 s. 13.
- Cronbach, L. J. & Meehl, P. E. (1955). "Construct validity in psychological tests". *Psychological Bulletin*, Vol. 52, No.4.
- Dahmström, K. (2005). *Från datainsamling till rapport – att göra en statistisk undersökning*. Lund: Studentlitteratur.

- Davidsson, G. & Wärneryd, B. (1990). Frågornas innehåll, struktur och ordning. I Wärneryd, B. (red.). *Att fråga*. Statistiska centralbyrån.
- Danielsson, C. & Levin, A. K. (2009). *Kartläggning av hemtjänstpersonals kunskaper för säker läkemedelshantering i samband med diabetes*. Uppsats i vårdvetenskap, avancerad nivå. Högskolan Dalarna: Akademin för hälsa och samhälle.
- De Leeuw, E.D. (2009). Mixed mode datacollection in surveys. *Kursmaterial*.
- De Leeuw, E.D. (2009). Mixing survey methods --- why do it and what are the consequences? *Kursmaterial*.
- De Leeuw, E.D. (2009). Mixing survey modes: challenges for questionnaire construction. *Kursmaterial*.
- De Leeuw, E.D. (2009). Types of mixed mode designs. *Kursmaterial*.
- De Leeuw, E.D. (2009). Why a changing world needs mixed mode designs. *Kursmaterial*.
- De Leeuw, E.D. (2008). Choosing the mode of data collection. . In: de Leeuw E.D., Hox J.J., Dillman D.A., editors. *International handbook of survey methodology*. New York: Lawrence Erlbaum Associates; 2008.
- De Leeuw, E. D. (2008). *The Effect of Computer-Assisted Interviewing on Data Quality: A Review of the Evidence*. Mimeo, Department of Methodology and Statistics, Utrecht University.
- De Leeuw, E. D. (1992). *Data Quality in Mail, Telephone, and Face-to-face Surveys*. Amsterdam: TT-Publicaties.
- De Leeuw, E. D., Hox, J. J., & Dillman, D. A. (2008). *International handbook of survey methodology*. NY: Lawrence Erlbaum Associates.
- De Leeuw, E. D., Callegaro, M., Hox, J., Korendijk, E. & Lendsveldt-Mulders, G. (2007). The influence of advance letters on response in telephone surveys - a meta-analysis. *Public Opinion Quarterly*, 71(3), s. 413-443.
- De Leeuw, E. D., Hox, J. J. & Huisman, M. (2003). Prevention and treatment of item nonresponse. *Journal of Official Statistics*. 19, s. 153-176.
- DeRouvray, C. & Couper M.P. (2002). Designing a Strategy for Reducing 'No Opinion' Responses in Web-Based Surveys. *Social Science Computer Review*, 20:3-9

- Dillman, D. A. (2007). *Mail and Internet Surveys: The Tailored Design Method*. (2:a upplagan, 2007 års utgivning). Hoboken, John Wiley Co., New Jersey.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). Internet, mail and mixed-mode surveys: the tailored design method. NJ: Wiley.
- Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J. Berck, J., & Messer, B.L. (2009). Response Rate and Measurement Differences in Mixed-mode Surveys: Using Mail, Telephone, Interactive Voice Response and the Internet. *Social Science Research* 38:s. 1-18.
- Dillman, D. A. & Smyth, J. D. (2007). Design effects in the transition to Web-based surveys. *American Journal of Preventive Medicine*, 32(5S), s. 90-96.
- Dillman, D. A., Gertseva, A. & Mahon-Haft, T. (2005). Achieving usability in establishment surveys through the application of visual design principles. *Journal of Official Statistics*, 21, s. 183-214.
- Dillman, D. A., Clark, J. R. & Sinclair, M. A. (1995). How Pronotice Letters, Stamped Return Envelopes and Reminder Postcards Affect Mailback Response Rates for Census Questionnaires. *Survey Methodology*, 21, s. 1-7.
- Dillman, D. A., West, K. K. & Clark, J. R. (1994). The Influence of an Invitation to Answer by Telephone on Response to Census Questionnaires. *Public Opinion Quarterly* 58: s. 557-568.
- Dillman, D. A. & Tarnai, J. (1991). Mode effects of cognitively-designed recall questions: A comparison of answers to telephone and mail surveys. I Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A. & Sudman, S. (Eds.), *Measurement errors in surveys* (s. 73-93).
- Dillman, D. A. & Mason, R. G. (1984). The Influence of Survey Method on Question Response. *Paper presented at the annual meeting of the American Association for Public opinion Research*, Wisconsin.
- Diment, K. & Garrett-Jones, S. (2007). How demographic characteristics affect mode preference in a postal/web mixed-mode survey of Australian researchers. *Social Science Computer Review*, 25, s. 420-427.
- Downing, S. M. (2006). Selected-response item formats in test development. I Downing, S. M. & Haladyna, T. M. (Eds.), *Handbook of test development* (s. 287-301). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

- Dykema, J., Diloreto, K., Price, J. L., White, E. & Schaeffer, N.C. (2012). ACASI gender-of interviewer voice effects on reports to questions about sensitive behaviors among young adults. *Public Opinion Quarterly*, 76, s. 311–325.
- Dykema J, Schaeffer NC. 2000. Events, instruments, and reporting errors. *American Sociological Review*, 65, 619-29
- Dykema, J. (1996). "Events, Instruments, and Reporting Errors: Combining Knowledge From Multiple Perspectives." *Unpublished Master's Thesis, Department of Sociology, University of Wisconsin-Madison.*
- Eckman, S., Kreuter, F., Kirchner, A., Jäckle, A., & Tourangeau, R. (2014). Assessing the mechanisms of misreporting to filter questions in surveys. *Public opinion quarterly*, 78, 721-733.
- Ekelund, C. (2010). *Tobaksvanor och kunskaper om tobak och dess effekter i munhålan En enkätstudie bland sjuksköterske-, tandhygienist- och folkhälso-pedagogstudenter vid Högskolan Kristianstad* (Examensarbete i Oral Hälsa). Högskolan Kristianstad: Sektionen för hälsa och samhälle, Tandhygienistprogrammet.
- Ekman, T. & Wohlgemuth, D. (2008). *Demokrati och mänskliga rättigheter. Svenska folkets kunskaper och värderingar.* (Ungdomsstyrelsens skrifter 2008:4). Stockholm: Fritzes.
- Eljertsson, G. (2005). *Enkäten i praktiken: en handbok i enkätmetodik.* Lund: Studentlitteratur.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Revised Edition). Cambridge, MA: MIT Press.
- Eriksson, J. (1991). Inkomstfördelningsundersökningen. Ur Tre återintervjustudier. *R&D report.* Statistiska centralbyrån.
- ESS (2006). *Handbook of Recommended Practices for Questionnaire Development and Testing in European Statistical System.* ESS.
- Fan, W. M. & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, 26(2), s. 132–139.
- Field, A. & Hole, G. (2003). *How to design and report experiments.* SAGE: London.
- Fink, A. (2003). *How to Ask Survey Questions. The Survey Kit 2.* Thousand Oaks, CA: SAGE Publications.

- Foddy, W. (1994). *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*. Cambridge University Press, Cambridge.
- Forsman, G. (1987). "Analys av återintervjudata", *Urval*, **19**, SCB.
- Forsman, G. & Schreiner, I. (1991). The design and analysis of reinterview: An overview, in *Measurement errors in surveys*, eds. P. Biemer et al., New York: John Wiley & Sons, s. 280–301.
- Fowler, F. J. (2009). *Survey Research Methods*. Thousand Oaks, CA: Sage Publications.
- Fowler, F. J. (1995). *Improving Survey Questions: Design and evaluation*. Thousand Oaks, CA: Sage Publications.
- Fowler, F. J. & Mangione, T. W. (1990). *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park, CA: Sage.
- Frankfort-Nachmias, C. & Nachmias, D. (1996). *Research Methods in the Social Sciences*. London: Arnold.
- Frary, R. B. (1995). More multiple-choice item writing do's and don'ts. *Practical Assessment, Research & Evaluation*, **4**(11).
- Friedman, W. J. (1987). A follow-up to "Scale effects in memory for the time of events". *Memory and cognition*, **15**, s. 518–520.
- Fritz, M. & Lundström, S. (2007). *Vilka faktorer påverkar svarsfrekvensen i enkätundersökningar?* SCB-rapport.
- Fråga rätt (2001). SCB-publikation. Örebro: Statistiska centralbyrån.
- Galesic, M. & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey, *Public Opinion Quarterly*, **73**(2), s. 349–360.
- Galesic, M. & Tourangeau, R. (2007). What Is Sexual Harassment? It Depends on Who Asks! Framing Effects on Survey Responses. *Applied Cognitive Psychology* **21**(2): s. 89–202.
- Galesic, M., Tourangeau, R. & Couper, M. P. (2006). Complementing Random-Digit-Dial Telephone Surveys with Other Approaches to Collecting Sensitive Data. *American Journal of Preventive Medicine*, Volume 31, Issue 5.
- Galesic, M., Tourangeau, R., Couper, M. P. & Conrad, F. G. (2008). Eye-tracking data: New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, **72**, s. 892–913.

- Gallagher, P. M., Fowler, F. J. & Stringfellow, V. L. (2005). The nature of nonresponse in a medicaid survey: *Causes and consequences*. *Journal of Official Statistics*, 21(1), s. 73–87.
- Greene, J., Speizer, H. & Wiitala, W. (2008). 'Telephone and Web: Mixed-Mode Challenge'. *Health Services Research* 43: s. 230–248.
- Gren, A. & Juklen, A. (2010). *Sjuksköterskestudenters kunskap om oral hälsa – En jämförelse mellan studenter i årskurs 1 och årskurs 3*. (Examensarbete i Oral Hälsa). Högskolan Kristianstad: Sektionen för hälsa och samhälle, Tandhygienistprogrammet.
- Griffis, S., Goldsby, T. & Cooper, M. (2003). Web-Based Surveys: A Comparison of Response, Data, and Cost. *Journal of Business Logistics*, 24: s. 237–258.
- Grice, H.P. (1975). Logic and Conversation. I Cole, P., & Morgan, J. L. (eds.). *Syntax and Semantics: Vol. 3. Speech arts*. NY: Seminar Press.
- Groves, R. M. (1989). *Survey Errors and Survey Costs*. NY: Wiley.
- Groves, R. M., Cialdini, R.B. & Couper, M.P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 4, s. 475–495.
- Groves, R. M. & Kahn, R. L. (1979). *Surveys by telephone*. NY: Wiley.
- Groves, R.M., Singer, E., Corning, A.D., & Bowers, A. (1999). A Laboratory Approach to Measuring the Joint Effects of Interview Length, Incentives, Differential Incentives, and Refusal Conversion on Survey Participation. *Journal of Official Statistics*, 15, 251–68.
- Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. (2002). A review of multiple choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), s. 309–334.
- Hansagi, H. & Allebeck, P. (1994). *Enkät och intervjù inom hälso- och sjukvård*. Lund: Studentlitteratur.
- Hansen, M. H., Hurwitz, W. N., Marks, E. S. & Mauldin, W. P. (1951). Response errors in surveys. *Journal of American Statistical Association*, 46, s. 147–190.
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T., Lyberg, L., Mohler, P., Pennell, B. & Smith, T. (2010). *Survey methods in multinational, multiregional and multicultural contexts*. New Jersey: John Wiley & Sons Inc.

- Harley, T. A. (2008). *The psychology of language – from data to theory*. NY: Psychology Press.
- Henley, J. R. H. (1976). "Response Rate to Mail Questionnaires With a Return Deadline" *The Public Opinion Quarterly*, Vol. 40, No. 3, s. 374–375.
- Hernández, A., Espejo, B. & González-Romá, V. (2006). The functioning of central categories middle level and sometimes in graded response scales: Does the label matter? *Psicothema*, 18.
- Herzog, A. R., & Bachman, J. G. (1981). Effects of Questionnaire Length on Response Quality. *Public Opinion Quarterly* 45(4): s. 549-559.
- Hippler, H. J. & Schwarz, N. (1986). *Not forbidding isn't allowing. The Cognitive Basis for the Forbid-Allow Assymetri*. *Public Opinion Quarterly*, 50:1.
- Holbrook, A.L., S. Anand, T.P. Johnson, Y.I. Cho, S. Shavitt, N. Chavez and S. Weiner. 2014. Response heaping in interviewer-administered surveys: is it really a form of satisficing? *Public Opinion Quarterly* 78(3): 591–633.
- Holbrook, A. & Krosnick, J. A. (2010). "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74: s.37–67.
- Holbrook, A. L., Krosnick, J. A., Moore, D., & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally: The impact of question and respondent attributes. *Public Opinion Quarterly*. 71(3): s. 325–348.
- Holbrook, A. L., Cho, Y. & Johnson, T. (2006). The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties. *Public Opinion Quarterly* 70(4): s. 565–595.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67, 79–125.
- Holbrook, A. L., Krosnick, J. A., Carson, R. T. & Mitchell, R. C. (2000). Violating conversational conventions disrupts cognitive processing of attitude questions. *Journal of Experimental Social Psychology* 36, s. 465–494.

- Holmberg, A., Lorenc, B., & Werner, P. (2010). 'Contact Strategies to Improve Participation via the Web in a Mixed-Mode Mail and Web Survey'. *Journal of Official Statistics*, 26(3): s. 465–480.
- Hox, J. J. (1997). "From Theoretical Concept to Survey Question" I Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwarz, N. & Trewin, D. (red.) *Survey Measurement and Process Quality*. New York: John Wiley & Sons, Inc.
- Hutchinson, E. (1995). *Cognitions in the Wild*. Cambridge, MA: MIT Press.
- Huttenlocher, J., Hedges, L., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of experimental Psychology: Learning, Memory, and Cognition*, 16.
- Iredell H., Shaw, T., Howatt, P., James, R. & Granich, J. 2004. Introductory Postcards: Do They Increase Response Rate in a Telephone Survey of Older Persons? *Health Education Research* 19(2): s. 159–164.
- Jann, B., Jerke, J. & Krumpal, I. (2012). Asking sensitive questions using the crosswise model: An experimental survey measuring plagiarism. *Public Opinion Quarterly*, 76, s. 1–18.
- Japac, L. (2005). *Quality issues in interview surveys: Some contributions*. Dissertation. Stockholm University, Department of statistics. Stockholm.
- Junger-Tas, J. & Marshall, I. (1999). "The Self-Report Methodology in Crime Research." I Tonry, M. (red.), *Crime and Justice. A Review of Research*. Chicago: University of Chicago Press.
- Jäckle, A. (2008). Dependent Interviewing: Effects on Respondent Burden and Efficiency of Data Collection. *Journal of Official Statistics*, Vol.24, No.3, s. 411–430 Karlsson, M. & Thudin, S. (1991). Årsarbetskraften. Ur Tre återintervjustider. R&D rapport Statistiska Centralbyrån.
- Kerr, M. (2005). *Scientific Writing*, forskarkurs, Örebro universitet.
- Knäuper, B. (1999). The impact of age and education on response order effects in attitude measurement. *Public Opinion Quarterly*. 63: s. 347–370.
- Knäuper, B., Schwarz, N., Park, D. & Fritsch, A. (2007). The Perils of Interpreting Age Differences in Attitude Reports: Question Order Effects Decrease with Age. *Journal of Official Statistics* 23: s. 515–528.

- Kreuter, F., Presser, S. & Tourangeau, R. (2008). Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly* 72(5): s. 847–865.
- Krosnick, J. A. (2000). The threat of satisficing in surveys: The shortcuts respondents take in answering questions. *Survey Methods Newsletter*, 20, s. 4–8.
- Krosnick, J. A., & Tahk, A. M. (2008). The Optimal Length of Rating Scales to Maximize Reliability and Validity. Unpublished manuscript. Stanford University.
- Krosnick, J. A., Judd, C. M. & Wittenbrink, B. (2005). Attitude measurement . I Albarracín, D., Johnson, B. T. & Zanna, M. P. (Eds.), *Handbook of attitudes and attitude change*. Mahwah, NJ: Erlbaum.
- Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., et al. (2002). The impact of 'no opinion' response options on data quality: non-attitude reduction or invitation to satisfice? *Public Opinion Quarterly*. 66.
- Krosnick, J. A., & Fabrigar, L. R., (1997). Designing rating scales for effective measurement in surveys. In Lyberg, Biemer, Collins, de Leeuw, Dippo, Schwarz & Trewin, *Survey Measurement and Process Quality*, Wiley, New York.
- Krosnick, J. A., & Berent, M. K. (1993). Comparisons of Party Identification and Policy Preferences: The Impact of Survey Question Format. *American Journal of Political Science* 37.
- Krosnick, J. A., (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology* 5.
- Krosnick, J. & D. Alwin. (1987). An Evaluation of a Cognitive Theory of Response Order Effects in Survey Measurement. *Public Opinion Quarterly* 52: s. 526–538.
- Kvale, S. (1997). *Den kvalitativa forskningsintervjun*. Lund: Studentlitteratur.
- Langlet, P. (1990). Att ta hänsyn till minnets begränsningar. I Wärneryd, B. (red.). *Att fråga*. Statistiska centralbyrån.
- Langlet, P. & Wärneryd, B. (1990). Att arbeta med språket. I Wärneryd, B. (red.). *Att fråga*. Statistiska centralbyrån.

- Lee, L., Brittingham, A., Tourangeau, R., Rasinski, K., Willis, G. et al. (1999). Are reporting errors due to encoding limitations or retrieval failure? *Journal of Applied Cognitive Psychology*, 13: s. 43–63.
- Lenzner, T., Kaczmirek, L., & Galesic, M. (2011). Seeing through the eyes of the respondent: An eye-tracking study on survey question comprehension. *International Journal on Public Opinion Research*, 23(3).
- Lessler, J. & Forsyth, B. (1996). A Coding System for Appraising Questionnaires. In *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, Eds. Norbert Schwartz & Seymoure Sudman, s. 259–292. San Francisco: Jossey-Bass.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
- Link, M. W. & Mokdad, A. (2006). Can web and mail survey modes improve participation in an RDD-based national health surveillance? *Journal of Official Statistics*, 22, 293–312.
- Lorenc, B. (2010). Item Nonresponse Analysis for a Mixed-Mode Survey. *Official Statistics – Methodology and Applications in Honour of Daniel Thorburn*, s. 117–136
- Lorenc, B. (2006a). Survey response process in establishments: The socially distributed cognition perspective. *Research Report RR 2006:5*, Department of Statistics, Stockholm University.
- Lorenc, B. (2006b). Social distribution of the response process in a survey of schools. *Research Report RR 2006:6*, Department of Statistics, Stockholm University.
- Lundh, L.G., Montgomery, H. & Waern, Y. (1992). *Kognitiv Psykologi*. Lund: Studentlitteratur.
- Luppés, M. (1995). A content analysis of advance letters from expenditure surveys of seven countries. *Journal of Official Statistics*, 11(4), s. 461–480.
- Lutig, P., Lensvelt-Mulders, G. J. L. M., Frerichs, R. & Greven, A. (2011). Estimating nonresponse bias and mode effects in a mixed-mode survey. *International Journal of Market Research*, 53(5), s. 669–686.
- Madans, J., Miller, K., Maitland, A. & Willis, G. (2011). *Question Evaluation Methods: Contributing to the Science of Data Quality*. Wiley, New York.
- Magnusson, D. (2003). *Testteori*. Stockholm: Psykologiförlaget AB.

- Maguire, M. (1997). "Crime Statistics, Patterns, and Trends: Changing Perceptions and their Implications". I Maguire, M., Morgan, R. & Reiner, R. (red.) *The Oxford Handbook of Criminology*. Oxford: Oxford University Press.
- Malholtra, N. (2009). Order Effects in Complex and Simple Tasks. *Public Opinion Quarterly*, 73: s. 180–198.
- Malhotra, N. (2008). Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly* 72: s. 914–934.
- Malhotra, N., Krosnick, J. A., & Thomas, R. K. (2009). Optimal Design of Branching Questions to Measure Bipolar Constructs. *Public Opinion Quarterly*. 73(2): s. 304–324.
- Mason, R., Carlson, J. E. & Tourangeau, R. (1994). Contrast Effects and Subtraction in Part- Whole Questions. *Public Opinion Quarterly* 58: s. 569–578.
- McClendon, M. J. & Alwin, D. F. (1993). No opinion filters and attitude measurement reliability. *Sociological Methods and Research* 21.
- McCabe S. E., Diez, A., Boyd, C.J. et al. (2006). Comparing web and mail responses in a mixed mode survey in college alcohol use research. *Addictive Behavior*; 31: s. 1619–1627.
- Mead, G. H. (1934). "Play, the Game, and the Generalized Other" I Coser L. A. & Rosenberg B. (red.) (1964). *Sociological Theory. A Book of Readings*. New York: Macmillian Company.
- Medway, R. L. & Fulton, J. (2012). When more gets you less: a meta-analysis of the effect of concurrent web options on mail survey response rates, *Public Opinion Quarterly* 76(4): s. 733–746.
- Menold, N., Kaczmirek, L., Lenzner, T., & Neusar, A. (2014). How do respondents attend to verbal labels in rating scales? *Field Methods*, 26, 21–39.
- Messick, S. (1989). "Validity". I Linn, R. (red.) *Educational Measurement*. New York: Macmillian Publishing Company.
- Millar M. M., & Dillman, D. A. (2011). Improving Response to Web and Mixed-Mode Surveys. *Public Opinion Quarterly* 75(2): s. 249–269.
- Millar, M.M., O'Neill, A.C., & Dillman, D.A. (2009). Are Mode Preferences Real? Technical Report 09-003. Pullman, WA: Social & Economic Sciences Research Center, Washington State University.

- Miller, M. K., & Orr, S. K. (2008). Experimenting with a "third way" in political knowledge estimation. *Public Opinion Quarterly*, 72(4), s. 768–780.
- Mondak, J. (2001). Developing valid knowledge scales. *American Journal of Political Science*, 45(1), s. 224–238.
- Mondak, J. (2000). Reconsidering measurement of political knowledge. *Political Analysis*, 8(1), s. 57–82.
- Moreno, R., Martinez, R. J., & Muniz, J. (2006). New guidelines for developing multiple-choice items. *Methodology: European Journal of Research Methods for the Behavioural and Social Sciences*, 2(2), s. 65–72.
- Morton, J., Mullin, P. & Biemer, P. (2008). Using reinterview and reconciliation methods to design and evaluate survey questions. *Survey Research Methods*, 2(2), s. 75–82.
- Morton-Williams, J. (1993). *Interviewer approaches*, Aldershot: Dartmouth Pub.
- Mu, X. (1999). IVR and Distribution of Responses: An Evaluation for the Effects of IVR on Collecting and Interpreting Survey Data. *Unpublished Paper*. The Gallup Organization, Princeton, NJ.
- Munson, J. M. & McIntyre, S. H. (1979). Developing Practical Procedures for the Measurement of Personal Values in Cross-Cultural Marketing. *Journal of Marketing Research*, 16.
- Myers, D. G. (2002). *Social Psychology*. NY: McGraw-Hill Higher Education.
- Norenzayan, A., & Schwarz, N. (1999). Telling what they want to know: Participants tailor causal attributions to researchers' interests. *European Journal of Social Psychology*, 29, 1011–1020.
- Norman, D. A. (1988). *The design of everyday things*. Basic Books.
- Oscarsson, H. & Holmberg, S. (2008). *Väljarkunskap*. Uppdaterad tabellsammanställning byggd på resultat redovisade i kapitlet Okunniga väljare? i boken Väljare.
- Olson, K. (2010). An Examination of Questionnaire Evaluation by Expert Reviewers. *Field methods*, 22(4), s. 295–318.

- Olson, K., Smyth, J.D., & Wood, H. (2010). Does Giving People Their Preferred Survey Mode Actually Increase Survey Participation? *An Experimental Examination. Presented at the annual meeting of the American Association for Public Opinion Research, Chicago, IL. May 13-16, 2010.*
- Olson, K., Smyth, J. D. & Wood, H. M. (2012). Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination, *Public Opinion Quarterly* 76(4): s. 611-635.
- Oskamp, S. & Schultz, P. W. (2005). *Attitudes and opinions*. NJ: Lawrence Erlbaum Associates.
- Parsons, T. (1951). "Institutionalization and Anomie". I Coser, L. A. & Rosenberg, B. (red.) (1964). *Sociological Theory. A Book of Readings*. New York: Macmillian Company.
- Passer, M. W. & Smith, R. E. (2004). *Psychology: The science of mind and behavior. 2:a upplagan*. New York: McGraw-Hill.
- Payne, S. L. (1951). *The Art of Asking Questions*. Princeton, NJ: Princeton University Press.
- Peter, J., & Valkenburg, P. M. (2011). The impact of "Forgiving" Introductions on the reporting of sensitive behavior in surveys. *Public Opinion Quarterly*, 75, 779-787.
- Persson, A., Björnram, A., Elvers, E., & Erikson, J. (2015). A strategy to test questionnaires at a national statistical office. *Statistical Journal of the IAOS*, 31, 2, 297-304
- Peytchev, A., Conrad, F. G., Couper, M. P. & Tourangeau, R. (2010). Increasing Respondents' Use of Definitions in Web Surveys. *Journal of Official Statistics*, Vol. 26, No. 4, s. 633-650.
- Ping, T.L.(2005). Does the survey response format matter? *Paper presented at IMTA*.
- Poe, G. S., Seeman, I., McLaughlin, I., Mehl, E. & Dietz, M. (1988). Don't Know' Boxes in Factual Questions in a Mail Questionnaire: Effects on Level and Quality of Response. *Public Opinion Quarterly* 52: s. 212-222.
- Presser, S., Rothgeb, J. M., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., Singer, E. (2004). *Methods for Testing and Evaluating Survey Questionnaires*. Wiley, New York.

- Presser, S. & Zhao, S. (1992). Attributes of questions and interviewers as correlates of interviewing performance. *Public Opinion Quarterly*, 56, s. 236–240.
- Presser, S. & Blair, J. (1994). Survey pretesting: Do different methods produce different results? In P. V. Marsden (Ed.), *Sociological Methodology*: Vol. 24, s. 73–100. Washington, DC: American Sociological Association.
- Presser, S. & Schuman, H. (1980). *Questions and Answers in Attitude Surveys*. NY: Academic Press.
- Preston, C. C. & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica* 104.
- Prüfer, P. & Rexroth, M. (2005). *Cognitive interviews (Cognitive interviews)*. ZUMA How to – No. 15.
- Quinney, R. (1970). *The Social Reality of Crime*. Boston: Little, Brown and Company.
- Rasinski, K. A., Mingay, D. & Bradburn, N.M. (1994). Do respondents really "mark all that apply" on self-administered questions? *Public Opinion Quarterly*, 58.
- Rasinski, K. A. (1989). The Effect of Question Wording on Public Support for Government Spending. *Public Opinion Quarterly* 53.
- Redline, C., Dillman, D., Smiley, R., Carley-Baxter, L., and Jackson, A., 2001, Making visible the invisible: an experiment with skip instructions on paper questionnaires, *Proceedings of the ASA Section on Survey Research Methods*. Alexandria, American Statistical Association.
- Regeringskansliet (2011). Svarta listan. Ord och fraser som kan ersättas i författningsspråk. SB PM 2011:1. Stockholm.
- Revilla, M. A., Saris, W. E. & Krosnick, J. A. (2014). Choosing the Number of Categories in Agree–Disagree Scales. *Sociological Methods & Research*, Vol 43(1), s. 73–97.
- Revilla, M. A., & Saris, W. E. (2013). A comparison of the quality of questions in a face-to-face and a web survey. *International Journal of Public Opinion Research*, 25, 242–253.
- Revilla, M., Saris, W. E. & Krosnick, J. A. (2009). Choosing the number of categories in agree/disagree scales. *RECSM working paper number 5*.

- Rips, L., Conrad, F. & Fricker, S. (2003). Straightening the Seam Effect in Panel Surveys. *Public Opinion Quarterly*, 67: s. 522–554.
- Rosing, H. (2002). *Vetenskapen logiska grunder*. Helsingfors: Schildts.
- Rothgeb, J., Willis, G. & Forsyth, B. H. (2001). *Questionnaire pretesting methods: do different techniques and different organizations produce similar results?* In Annual Meeting of the American Statistical Association.
- Sánchez-Fernández, J., Muñoz-Leiva, F. & Montoro-Ríos, F. J. (2012). Improving retention rate and response quality in web-based surveys. *Computers in Human Behavior*, 28(2).
- Sala, S. C. Uhrig, N. & Lynn, P. (2009). It is Time Computers do Clever Things! The Impact of Dependent Interviewing on Interviewer Burden. *Field Methods* 23: s. 3–23.
- SAOL (2013). Svenska akademiens ordlista över svenska språket. Svenska akademien.
- Saris, W. E. & Gallhofer, I. N. (2007) *Design, Evaluation and Analysis of Questionnaires for Survey Research*. New York: Wiley.
- Saris, W. E., Krosnick, J. A. & Schaeffer, E. M. (2005). *Comparing questions with agree/disagree response options to questions with construct-specific response options*. Unpublished manuscript, Political, Social, Cultural Sciences, University of Amsterdam.
- Sayles, H., Belli, R. F., & Serrano, E. (2010). Interviewer Variance Between Event History Calendar and Conventional Questionnaire Interviews. *Public Opinion Quarterly*, 74: s. 140–153.
- SCB (1997). *Välfärd och ojämlikhet i 20-årsperspektiv 1975–1995*. Serien Levnadsförhållanden, Rapport 91.
- Scherpenzeel, A. C., & Saris, W. E. (1997). The validity and reliability of survey questions a meta-analysis of MTMM studies. *Sociological Methods & Research*, 25(3).
- Schaeffer, E. M., Krosnick, J. A., Langer, G. E. & Merkle, D. M. (2005). Comparing the Quality of Data Obtained by Minimally Balanced and Fully Balanced Attitude Questions. *Public Opinion Quarterly*, 69:3.
- Schaeffer, N.C. (2010). Presentation at the *Question Evaluation Methods conference* in Hyatsville.
- Schaeffer, N. C. & Presser, S. (2003). The Science of Asking Questions. *Annual Review of Sociology*, 29.

- Schaeffer, N. C. (1991). Hardly Ever or Constantly? Group Comparisons using Vague Quantifiers. *Public Opinion Quarterly* 55.
- Schaeffer, N. C. & Bradburn, N. M. (1989). Respondent behavior in magnitude estimation. *Journal of the American Statistical Association* 84.
- Schober, M. F., Conrad, F. G., Dijkstra, W. & Ongena, Y. P. (2012). Disfluencies and Gaze Aversion in Unreliable Responses to Survey Questions. *Journal of Official Statistics*, 28(4), s. 555–582.
- Schober, M. F. (1999). Making Sense of Questions: An Interactional Approach. I Sirken, M. G., Herrmann, D. J., Schechter, S., Schwarz, N., Tanur, J. M. & Tourangeau, R. (eds.) *Cognition and Survey Research*. NY: John Wiley & Sons Inc.
- Schober M. F., & Conrad F. G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*. 60.
- Schuman, H., & Presser, S. eds. (1981). *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Context*. New York: Academic.
- Schuman, H., Presser, S. & Ludwig, J. (1981). Context effects on survey responses to questions about abortion. *Public Opinion Quarterly*, 45, s. 216–223.
- Schuman, H. (1972). Attitudes vs. actions versus attitudes vs. attitudes. *Public Opinion Quarterly*. 36: s. 347–354.
- Schwarz, N. (1999). Self-reports: How the questions shapes the answers. *American psychologist*, 54, p. 93-105.
- Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Erlbaum.
- Schwarz, N., Knäuper, B., Oyserman, D., & Stich, C. The psychology of asking questions. In: de Leeuw E.D., Hox J.J., Dillman D.A., editors. *International handbook of survey methodology*. New York: Lawrence Erlbaum Associates; 2008. p. 18–34.
- Schwarz, N. & Bohner, G. (2001). The construction of attitudes. I Tesser, A. & Schwarz, N. (Red.), *Blackwell handbook of social psychology: Intraindividual processes* (Vol.1, s. 436–457). Oxford, UK: Blackwell.
- Schwarz, N. & Hippler, H. (1995). Subsequent questions may influence answers to preceding questions in mail surveys. *Public Opinion Quarterly*, 59, s. 93–97.

- Schwarz, N. & Hippler, H. (1995). The numeric values of rating scales: A comparison of their impact in mail surveys and telephone interviews. *International Journal of Public Opinion Research*, 7, s. 72–74.
- Schwarz, N., Knäuper, B., Hippler, H., Noelle-Neumann, E. & Clark, L. (1991). Rating Scales: Numeric Values may Change the Meaning of Scale Labels. *Public Opinion Quarterly* 55.
- Schwarz, N., Strack, F., Mai, H-P. (1991). Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis. *The Public Opinion Quarterly*, Vol. 55, No. 1 (Spring, 1991), s. 3–23.
- Schwarz, N., Strack, F., Muller, G. & Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: Further evidence on informative functions of response alternatives. *Social Cognition*, 6, s. 107–117.
- Schwarz, N., Hippler, H.J., Deutsch, B. & Strack, F. (1985). Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgments. *Public Opinion Quarterly* 49.
- Shum, M.S. & Rips, L.J. (1999). The respondent's confession: autobiographical memory in the context of surveys. In Sirken, M.G., Herrmann, D.J., Schechter, S., Schwarz, N., Tanur, J.M., & Tourangeau, R. (eds) *Cognition and survey research*. Wiley.
- Singer, E. & Bossarte, R. (2006). "Incentives for survey participation: when are they coercive?" *American Journal of Preventive Medicine*. Vol. 31, s. 411–418.
- Singer, E., Van Hoewyk, J. & Maher, M. P. (2000). Experiments With Incentives in Telephone Surveys. *Public Opinion Quarterly* 64: s. 171–188.
- Singer, E., Von Thurn, D. R. & Miller, E. R. (1995). Confidentiality Assurances and Response: A Quantitative Review of the Experimental Literature. *Public Opinion Quarterly* 59: s. 66–77.
- Singer, E., Hippler, H. & Schwarz, N. (1992). Confidentiality assurances in surveys: Reassurance or threat? *International Journal of Public Opinion Research*, 4.
- Sjöberg, K. (red.) (1999). *Mer än kalla fakta. Kvalitativ forskning i praktiken*. Lund: Studentlitteratur.
- Skolverket (2003). *Ung i demokratin. Gymnasieelevers kunskaper och attityder i demokrati och samhällsfrågor*, Stockholm, Skolverket.

- Smith, W. R. (1992). The impact of question order on cognitive effort in survey responding. *Proceedings NCUR VI*.
- Smyth, J. D., Dillman, D. A., Christian, L. M. & McBride, M. (2009). Open ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality. *Public Opinion Quarterly*, 73 (2), s. 325–337.
- Smyth, J. D., Christian, L. M. & Dillman, D. A. (2008). Does yes or no on the telephone mean the same as check-all-that-apply on the Web? *Public Opinion Quarterly*, 72(1).
- Smyth J. D., Dillman D. A., Christian L. M. & Stern M. J. (2005). Comparing check-all and forced-choice question formats in web surveys. *Public Opinion Quarterly* 71: s. 66.
- Sohlberg, P. & Sohlberg, B-M. (2008). *Kunskapens former: vetenskapsteori och forskningsmetod*. Malmö:Liber.
- Spector, P. E. (1992). *Summated rating scale construction. An introduction*. Newbury Park: SAGE Publications.
- Språkrådet (2014). *Myndigheternas skrivregler*. Stockholm: Norstedts Juridik/Fritzes.
- Stiftelsen för opinionsanalyser (1983). *Enkät*, 5.
- Sturgis, P., Allum, N. & Smith, P. (2007). An experiment on the measurement of political knowledge in surveys. *Public Opinion Quarterly*, 85(1), s. 90-102.
- Sudman, S., Willimack, D. K., Nichols, E. & Mesenbourg, T. L. (2000). Exploratory Research at the U.S. Census Bureau on the Survey Response Process in Large Companies. *Proceedings of the Second International Conference on Establishment Surveys, American Statistical Association*, s. 327–337.
- Sudman, S., Bradburn, N. M. & Schwarz, N. (1996). *Thinking about answers – The application of cognitive processes to survey methodology*. SF: Jossey-Bass Publishers.
- Sudman, S. & Bradburn, N. M. (1974). *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Sudman, S. & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 68, s. 805–815.

- Sundström, A., Lyrén, P-E, & Alger, S. (2010). Att mäta allmänhetens kunskaper om trafiksäkerhet: Utveckling och utprövning av instrument. *BVM-rapport nr 46*. Umeå universitet, Institutionen för Beteendevetenskapliga mätningar.
- Sztabinski, P. B., Dyjas-Pokorska, A., & Zmijewska-Jedrzejczyk, T. (2006). Understanding refusals. *Proceedings at the 19th International Workshop on Household Survey Nonresponse*. SIS: Nonresponse in comparative surveys.
- Tarnai, J. & Dillman, D. A. (1992). Questionnaire context as a source of response differences in mail versus telephone surveys. In N. Schwarz & S. Sudman (Eds.), *Context Effects in Social and Psychological Research* (pp. 115–129). New York: Springer-Verlag.
- Taylor, H. (1998). "Rationing crime: the political economy of criminal statistics since the 1850s". *Economic History Review*, Vol. 51, No. 3, s. 569–590.
- Toepoel, V., Das, M. & van Soest, A. (2008). "Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents." *Public Opinion Quarterly* 72: s. 985–1007.
- Toepoel, V., & Couper, M. (2011). Can verbal instructions counteract visual context effects in web surveys? *Public Opinion Quarterly*, s. 1–18.
- Tourangeau, R., Conrad, F.G., Couper, M.P., & Ye, C. (2014). The Effects of providing examples in survey questions. *Public opinion quarterly*, 78, 100-125.
- Tourangeau, R., Groves, R. M., Kennedy, C., Yan, T. (2009). The Presentation of a Web Survey, Nonresponse, and Measurement Error among Members of Web Panel. *Journal of Official Statistics*. 25:s. 299–321.
- Tourangeau, R., Couper, M. P., Conrad, F. (2004). Spacing, position, & order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly* 68 (3).
- Tourangeau, R., Miller-Steiger, D., & Wilson, D. (2002). Self-administered questions by telephone: Evaluating interactive voice response. *Public Opinion Quarterly*. 66: s. pp. 265-278.
- Tourangeau, R., Rips, L. J. & Rasinski, K. (2000). *The Psychology of Survey Response*. NY: Cambridge University Press.

- Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. I Stone, A. A., Turkkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S. & Cain, V.S. (eds.). *The science of self-reports*. NJ: Lawrence Erlbaum associates.
- Tourangeau, R., Rasinski, K. & Bradburn, N. (1991). Measuring happiness in surveys: a test of the subtraction hypothesis. *Public Opinion Quarterly*. 55: s. 255–266.
- Tourangeau, R., Rasinski, K., Bradburn, N. & D'Andrade, R. (1989). Carryover Effects in Attitude Surveys. *Public Opinion Quarterly* 53: s. 495–524.
- Tourangeau, R. (1984). Cognitive sciences and survey methods. I Jabine T., Straf M., Tanur J., Tourangeau R., (red.) *Cognitive aspects of survey methodology: building a bridge between disciplines*. Washington DC: National Academy Press, 1984.
- Trankell, A. (1971). Vittnespsykologins arbetsmetoder. Utbildningsförlaget, Örebro.
- Trost, J. (1994). *Enkätboken*. Lund: Studentlitteratur.
- Trost, J. & Levin, I. (2004). *Att förstå vardagen – med ett symboliskt interaktionsistiskt perspektiv*. Lund: Studentlitteratur.
- Tsuchiya, T., Hirai, Y. & Ono, S. (2007). A Study of the Properties of the Item Count Technique, *Public Opinion Quarterly* 71.
- Tulving, E. (1987). Multiple memory systems and consciousness. *Human Neurobiology*, 6, s. 67–80.
- Tversky, A. (1977). On the elicitation of preferences: Descriptive and prescriptive considerations. In D. Bell, R. L. Kenney, & H. Raiffa (Eds.), *Conflicting objectives in decisions*. *International Series on Applied Systems Analysis* (s. 209–222). New York: Wiley.
- Velez, P. & Ashworth, S. D. (2007). The impact of item readability on the endorsement of the midpoint response in surveys. *Survey Research Method*, 1(2).
- Wanke, M., Schwarz, N. & Noelle-Neumann, E. (1995). Asking Comparative Questions: The Impact of the Direction of Comparison. *Public Opinion Quarterly* 59(3): s. 347–372.

- Wenemark, M., Persson, A., Noorlind Brage, H., Svensson, T. & Kristenson, M. (2011). Applying motivation theory to achieve increased response rates, respondent satisfaction and data quality. *Journal of Official Statistics*, 27(2), s. 393-414.
- Wenemark, M. (utkommer våren 2017) Enkätmetodik med respondenten i fokus. Studentlitteratur.
- Wenemark, M. (2010). *The Respondent's Perspective in Health-Related Surveys: The Role of Motivation*. Avhandling. Linköping. Linköpings Universitet.
- Winkielman, P., Knauper, B., Schwarz, N. (1998). Looking back at anger: Reference periods change the interpretation of emotion frequency questions. *Journal of Personality and Social Psychology*, 75, 719-28.
- Wikman, A. (1990). Att studera attityder och opinioner. I Wärneryd, B. (red.). *Att fråga*. Statistiska centralbyrån.
- Wikman, A. & Wärneryd, B. (1990). Measurement errors in survey questions: Explaining response variability. *Social Indicators Research*, 22, 199-212
- Willimack, D. & Nichols, E. (2010). Hybrid Response Process Model for Business Surveys. *Journal of Official Statistics*, Vol. 26, No. 1, 2010, s. 3-24.
- Willimack, D. K. & Nichols, E. (2001). Building an Alternative Response Process Model for Business Surveys. *Proceedings of the American Statistical Association*, Section on Survey Research Methods.
- Willis, G. (2005). *Cognitive Interviewing*. Thousand Oaks, CA: Sage.
- Willis, G. B., Lessler, J. T. (1999). *Question Appraisal System QAS-99*. Rockville, MD: Research Triangle Institute.
- Wilson, B. F., Altman, B. M., Whitaker, K. & Callegaro, M. (2004). How Good is Good? Comparing Numerical Ratings of Response Options for Two Versions of the Self-assessed Health Status Question. *Journal of Official Statistics*. 20:2.
- Wolming, S. (1998). "Validitet - ett traditionellt begrepp i modern tillämpning". *Pedagogisk Forskning*. Vol. 3. Nr. 2.
- Wärneryd, B. m.fl. (1990). *Att fråga*. Statistiska centralbyrån.
- Wärneryd, B. (1990). Utformningen av svarsalternativ. I Wärneryd, B. (red.). *Att fråga*. Statistiska centralbyrån
- Wärneryd, B. (1989). Levnadsförhållanden. Återintervjustudie i undersökningen av levnadsförhållanden (ULF). (In Swedish Living

- conditions. Reinterviews in ULF 1989). Stockholm: Statistics Sweden, 1991. Appendix 12.
- Yan, T. & Keusch, F. (2015). The effects of direction of rating scales on survey responses in telephone survey. *Public opinion quarterly*, 79, 145-165.
- Yan, T., Kreuter, F. & Tourangeau, R. (2012). Evaluating Survey Questions: A Comparison of Methods. *Journal of Official Statistics*, 28, s. 503-529.
- Yan T., Conrad F. G., Tourangeau R. & Couper M. P. (2010). Should I stay or should I go. *International Journal of Public Opinion Research*. 23 2, s. 131-47.
- Ye, C, Fulton, J. & Tourangeau, R. (2011). More Positive or More Extreme? A Meta Analysis of Mode Differences in Response Choice. *Public Opinion Quarterly* 75(2): s. 349-365.
- Yeager, D. S. & Krosnick, J. A. (2012). Does Mentioning "Some People" and "Other People" in an Opinion Question Improve Measurement Quality? *Public Opinion Quarterly*, 76:1.
- Zigerell, L. J. (2011). Midpoint misperceptions on 1-10 scales. *International Journal of Public Opinion Research*. 23, 2.

Tabell- och figurförteckning

Figur 1.1	12
Figur 5.1	165
Figur 5.2	165
Figur 5.3	166
Figur 5.4	166
Figur 5.5	167
Figur 5.6	167
Figur 6.1	219
Tabell 6.1	253
Figur 7.1	273
Tabell 8.1	286
Figur 8.1	301
Figur 8.2	304
Figur 10.1	327
Figur 11.1	349
Figur 14.1	403
Figur 14.2	420

”Som man frågar får man svar” är ett talesätt som i högsta grad är relevant för statistiska undersökningar. Att ställa frågor där svaren syftar till att skapa underlag för statistik ställer höga krav på den som konstruerar frågorna.

”Frågor och svar – om frågekonstruktion vid enkät- och intervjuundersökningar” vänder sig till dig som vill veta mer om att ställa frågor i statistiska undersökningar. Boken ger både en introduktion till frågekonstruktion och en fördjupning i ämnet. Den ger praktiska tips för den som vill ha konkret stöd i formuleringen av en fråga, men innehåller även teoretiska resonemang utifrån forskning och praktiska erfarenheter.

Huvudförfattare och redaktör är Andreas Persson vid Statistiska centralbyrån och medförfattare är Lina Fjelkegård, Statistiska centralbyrån, Pia Hartwig, Statistiska centralbyrån och Anna Sundström, Umeå universitet.

ISBN 978-91-618-1653-8 (print)

All officiell statistik finns på: **www.scb.se**
Statistikservice: tfn 010-479 50 00

All official statistics can be found at: **www.scb.se**
Statistics Service, phone +46 10 479 50 00