

Henning Olsen

# De plukkes ikke som blomster på en mark...

## Tendenser i kvantitativ metode- litteratur vedr. konstruktion og kvalitetssikring af surveydata

På grundlag af et litteraturstudie af kvantitativ metode-  
litteratur handler artiklen om, hvordan surveydata  
bør konstrueres og kvalitetssikres, når metodeforfat-  
tere virker som "dommere": Hvad anbefales vedr. ud-  
arbejdelse af spørgeskemaer? Hvad er anbefalingerne  
vedr. udarbejdelse af spørgeskemaer med afsæt i  
forskningsresultater om respondenters løsning af  
kognitive opgaver? Hvad anbefales vedr. gennemfø-  
relse af surveyinterview? Hvad anbefales vedr. proce-  
durale ledetråde for kvalitetssikring af surveydata, fx  
prøve- og laboratorieinterview? Er surveymetodens  
datakonstruerende aspekter alt i alt så standardise-  
rede som almindeligvis antaget?

På grundlag af resultater fra et i Socialforskningsinstituttet gennemført litteraturstudie af engelsksproget og skandinavisk kvantitativ metodelitteratur sætter artiklen fokus på fire problemstillinger (se: *oversigt 1*).<sup>1</sup> Problemstillingerne omhandler, hvordan surveydata bør konstrueres og kvalitets-sikres, når metodeforfattere tildeles rollen som "dommere". Med Hansen & Andersen's (2000) velkendte og i Danmark meget anvendte introduktion til survey-metoden som komparativt grundlag kondenseres tendenser i kvantitativ metodelitteratur.

### Spørgeskemakonstruktion

Den første problemstilling handler om, hvordan i problemstillinger indeholder begreber defineres og operationaliseres, hvorefter opmærksomheden rettes imod metodeforfatteres typologisering – dvs. generaliserende kategorisering – af spørgsmål. Derpå diskuteres anbefalinger vedr. konstruktion af "gode" spørgeskemaer. Selv om det i praksis er vanskeligt at abstrahere fra respondenters opgaveløsninger, ses der indledningsvis bort fra respondenter som mulig årsag til generering af måleproblemer, dvs. reliabilitets- og validitetsproblemer. Derfor belyses den første problemstilling overvejende p.g.r.l.a. erfaringsbaseret *common sense*-litteratur.

Kvantitativ metodelitteratur tilbyder ikke konsensuelle svar på, hvad validitet og reliabilitet er, og hvorledes gyldighed og pålidelighed kan afprøves empirisk (se fx: Aldridge & Levine 2001; Babbie 1990; Sapsford 1999). Desuden undlader mange metodeforfattere at diskutere problematikken og anvender de to begreber ureflektet. Der er derimod udbredt enighed om de to begreber som konceptuelt grundlag for kvalitetssikring af spørgeskemakonstruktion, tilvirkning af data mv. Tilsvarende gælder ikke kvalitativ forskning, hvor mange metodeforfattere lægger afstand til "positivistiske" eller "empiristiske" begreber og substituerer dem med andre, fx "troværdighed" (se fx: Olsen 2002).

Ligeledes understreges, at validitet og reliabilitet ikke er symmetriske. Ifølge "standardlærebogen" kan målinger have en høj grad reliabilitet uden at være valide, mens det omvendte er umuligt (se fx: Nolinske 1998). Én betingelse for at data har høj grad af validitet er m.a.o., at også reliabiliteten er uanfægtelig. Som bl.a. påpeget af Hansen & Andersen (2000) kan det dog i praksis – fx pga. diffus spørgeintention – være svært at afgøre, om et måleproblem er et validitets- eller reliabilitetsproblem. Bl.a. derfor omtales de to begreber undertiden under fællesbetegnelsen *måleproblem* (fx: Olsen 2001) eller *measurement error* (fx: Groves 1989).<sup>2</sup>

#### *Oversigt 1. Problemstillinger.*

|                    |  |
|--------------------|--|
| 1. problemstilling | Hvad er metodeforfatteres anbefalinger vedr. udarbejdelse af spørgeskemaer?  |
| 2. problemstilling | Hvad er metodeforskeres resultater og anbefalinger vedr. udarbejdelse af spørgeskemaer med afsæt i respondenters løsning af kognitive opgaver? |
| 3. problemstilling | Hvad er metodeforfatteres anbefalinger vedr. gennemførelse af surveyinterview?   |
| 4. problemstilling | Hvad er metodeforfatteres anbefalinger vedr. kvalitetssikring af surveydata?   |

### *Fra problemstilling til operationalisering*

Hansen & Andersen påpeger, at spørgsmålsformulering, som problemstillinger rejser, "... skal have så præcis en formulering, at de kan danne grundlag for vurderingen af, *hvilken form for viden*, undersøgelsen skal frembringe ..." (Hansen & Andersen 2000:31). Hvordan problemstillinger formuleres og begreber operationaliseres, tildeles generelt en tilbagetrukket rolle i kvantitativ metodelitteratur: "... considerably less effort has been devoted to steps that precede questionnaire construction ..." (Schwarz, i: Lyberg et al. 1997:30). Flere metodeforfattere pointerer dog, at surveyundersøgelser forudsætter, at problemstillinger formuleres så præcist som muligt (se fx: Fowler 1993).

Ifølge Hansen & Andersen 2000 er *definitioner* af i problemstillinger indeholdte begreber hverken "sande" eller "falske", hvilket også påpeges af andre metodeforfattere: "... most concepts of interest to social researchers have *no real meaning*, no ultimate definitions." (Babbie 1990:119). Da kun få eller ingen begreber har "egentlig mening", bør de defineres og operationaliseres. At definere begreber "... involves concept formation, which establishes the meaning of a construct by elaborating the nomological network and defining important subdomains ..." (Hox, i: Lyberg et al. 1997:53). Kvantitativ metodelitteratur indeholder imidlertid kun få ledetråde for, *hvordan* begreber defineres. Ved at sætte fokus på en række eksempler understreger fx Andersson (1994) vigtigheden af at udpege til begreber knyttede komponenter eller dimensioner. Hertil kommer, at begrebsdefinitioner skal kunne relevansbegrundes.

Når begreber er defineret, skal de *operationaliseres*:

Operationalization is the process whereby researchers specify em-

alm. foto

Henning Olsen  
Seniorforsker ved  
Socialforsknings-  
instituttet,  
Adj. docent ved  
Växjö Universitetet.  
E-mail: ho@sfi.dk

pirical observations that can be taken as indicators of the attributes contained within a given concept (Babbie 1990:121; se også: Hox, i: Lyberg et al. 1997).

Derfor peger Hansen & Andersen (2000) og andre på en særlig operationaliseringsfase:

The process of moving from abstract concepts to the point where we can develop questionnaire items to tap the concept is called "descending the ladder of abstraction" (de Vaus 1998:50).

Efter manges opfattelse er fasen kritisk. Hverken problemstillinger eller definitioner bestemmer nemlig umiddelbart spørgsmål, som respondenter skal besvare. Der skal derfor argumenteres for, at spørgsmål er relateret til problemstillinger og hertil hørende begreber. Efter hvilke retningslinjer indikatorer udvikles, besvarer metodelitteraturen ikke éntydigt.

### *Spørgsmålstypologier*

Er der metodologisk konsensus om typologisering af surveyspørgsmål? Hansen & Andersen (2000) udpeger tre typologi-

serende dimensioner. Den første dimension – indholdsdimensionen – handler om, hvorvidt spørgsmål sigter mod at belyse faktiske forhold eller holdninger. Hansen & Andersen (2000) sondrer mellem faktuelle spørgsmål, kognitive spørgsmål, holdningsspørgsmål og evalueringsspørgsmål. Den anden dimension er tidsdimensionen (fortid, nutid, fremtid), mens den tredje er svardimensionen, dvs. lukkede versus åbne spørgsmål.

Mens de to sidste dimensioner tages *ad notam* i anden metodelitteratur, gælder det ikke indholdsdimensionen. Den overvejende tendens er dog dikotomien *faktuelle spørgsmål* versus *holdningsspørgsmål* (se fx: Babbie 1990; Judd et al. 1991). At der sondres mellem de to spørgsmålstyper, følger af sidstnævntes ikke-eksterne reference. Men hvad holdninger og dermed holdningsspørgsmål er, og hvilke komponenter – kognitive, affektive versus konative – holdninger inkluderer, er der langtfra konsensus om (se fx: Olsen 2001a). Ikke desto mindre bifalder de fleste metodeforfattere en basal sondring mellem faktuelle spørgsmål og holdningsspørgsmål. Det gælder fx Andersson (1994), der opdeler faktuelle spørgsmål i to kategorier: spørgsmål om faktiske forhold vs. kundskab. Et andet eksempel er Aldrige & Levine (2001), der sondrer mellem spørgsmål ang. kende-

tegn, adfærd og holdninger. Hermed anskueliggøres to væsentlige pointer: At der ikke er enighed om, hvordan spørgsmål typologiseres, og at Hansen & Andersen's (2000) typologi blot er én af mange opfattelser, hvis bredde overskrider nuanceforskelle (se: *oversigt 2*, der læses vertikalt).

### Fra ord til spørgeskema

Er konstruktion af spørgeskemaer med Payne's (1951) formulering *the art of asking questions*, eller er fordringen *the interdisciplinary science of asking questions*? Ifølge Hansen & Andersen (2000) bør spørgsmål "... formuleres i et sprog, der forstås af og falder naturligt for respondenterne." (Hansen & Andersen 2000: 108). Deres position er, at spørgsmålsformulering er et "forskningshåndværk". Hvis "grunden skrider" under surveyforskere og hverdagsprog er "ekspresivt" (Hansen & Andersen 2000), er det nærliggende at give afkald på kvalitetskriterier. Ikke desto mindre følger en lang række erfarings- og *common sense*-baserede anbefalinger med få referencer til metodeforskning. Forfatterne advarer fx imod anvendelse af dobbelte negationer, lange spørgsmål, "værdiladede" ord og "ubalancerede" spørgsmål.

Spørgeskemakonstruktion afhænger af måden, hvorpå surveydata tilvirkes, dvs. besøg, telefon eller selvbesvarelse.

### Oversigt 2. Spørgsmålstypologier i kvantitativ metodelitteratur (eksempler).

|  |   |   |   |   |   |   |   |
|--|---|---|---|---|---|---|---|
| Adfærdsspørgsmål .....                             |   |   |   | • | • | • |   |
| Faktuelle spørgsmål .....                          | • | • | • |   |   |   |   |
| Faktuelt-episodiske spørgsmål (tid/rum) .....      |   |   |   |   |   |   | • |
| Faktuelt-generelle spørgsmål .....                 |   |   |   |   |   |   | • |
| Holdningsspørgsmål .....                           | • | • | • | • |   | • | • |
| Kendetegnsspørgsmål ( <i>attributes</i> ) .....    |   |   |   | • | • |   |   |
| Kognitive spørgsmål (kundskabsspørgsmål).....      | • |   | • |   |   | • | • |
| Kombinationsspørgsmål (faktualitet/holdning) ..... |   |   |   |   |   |   | • |
| Opfattelsesspørgsmål ( <i>beliefs</i> ) .....      |   |   |   |   | • |   |   |
| Vurderingsspørgsmål (evalueringsspm.) .....        | • |   |   |   |   | • |   |

Den gennemgående, men ikke eneherkende anbefaling er, at der er taget stilling hertil, før et spørgeskema udarbejdes. Om valg af datatilvirkningsform findes en righoldig metodelitteratur. Nedenfor følger en oversigt over ofte nævnte *pro et contra*-anbefalinger (se: *oversigt 3*). Der sondres mellem besøgs-, telefon- og postspørgeskemabaseret tilvirkning, hvoraf førstnævnte er PAPI- eller CAPI-interview (*paper and pencil interview* henhv. *computer assisted interview*), mens telefoninterview almindeligvis er CATI-baserede (*computer assisted telephone interview*). Postspørgeskemaundersøgelser er normalt PAPI-baserede, om end også CASI (*computer assisted self interview*) bruges (se fx: de Vaus 2002).

De i *oversigt 3* viste fordele og ulemper anskueliggør tre pointer: Der er ikke overensstemmelse mellem Hansen & Andersen's (2000) overvejelser ang. datatilvirkningsformer og de påpegede *pro et contra*-aspekter. For det andet har andre forfattere ikke samme opfattelse af tilvirkningsformernes fordele og ulemper. Tværgående er der markante variationer i, hvad forfattere anbefaler. For det tredje retter mange anbefalinger sig imod andre aspekter end konstruktion af spørgeskemaer. Mange forfattere påpeger, at den "bedste" procedure til tilvirkning af surveydata ikke findes. Der skal fx tages hensyn til økonomi og datakvalitet.

Trods betydelige omkostninger er *besøgsinterview* "... the preferred method of data collection for some surveys because it lends well to certain questionnaire items and issues of data quality." (Czaja & Blair 1996:43). Blandt fordele nævnes, at mange spørgsmål med forskellige emner er mulige (se fx: Aldridge & Levine 2001). Det påpeges tillige, at åbne spørgsmål er anvendelige, fx også visuelle hjælpemidler. Til gengæld skal der under spørgeskemakonstruktionen tages to hensyn, nemlig til såvel respon-

denter som interviewere (se fx: Sapsford 1999). Hertil kommer, at spørgsmål om "socialt ønskede" handlinger bør begrænses, da hyppigheden af "socialt ønsket" adfærd ofte overvurderes (se fx: Nolinske 1998).

Navnlig fordi *telefoninterview* er tidsbesparende og fhv. omkostningslette, finder de voksende udbredelse: "... there is much to recommend the use of the telephone as a major form of interview research." (Crano & Brewer 2002:224) Også telefoninterview har fordele og ulemper. Som ved besøgsinterview skal spørgeskemaer konstrueres ud fra nævnte dobbelthensyn. Respondenter skal forstå spørgsmål og (re)konstruere informationer. Samtidig skal spørgsmål formuleres, så oplæsning muliggøres (se fx: Sapsford 1999). Modsat besøgsinterview peges der på, at sensitive spørgsmål er mere anvendelige. Derimod bør telefoninterview gennemføres med afsæt i spørgeskemaer med korte og sprogligt enkle spørgsmål med få svarkategorier (se fx: de Vaus 2002). Hertil kommer, at åbne spørgsmål er vanskelige at administrere.

I modsætning til besøgs- og telefoninterview skal der ved udarbejdelse af *postspørgeskemaer* kun tages hensyn til respondenter, ligesom både sensitive spørgsmål og visuelle hjælpemidler er mulige (se fx: Aldridge & Levine 2001; Bickman & Rog 1998). På den anden side skal postspørgeskemaer være "selvforklarende". Adskillige metodeforfattere fraråder derfor lange og sprogligt komplekse spørgsmål (se fx: Andersson 1994), mens andre advarer imod anvendelse af åbne spørgsmål. Endelig påpeges, at kundskabsspørgsmål bør ekskluderes.

Efter mange metodeforfatteres opfattelse er spørgeskemakonstruktion surveyundersøgelser "kerne". Der findes en righoldig metodelitteratur med veksellende anbefalinger om, hvordan "gode"

**Oversigt 3. Fordele og ulemper ved tre datatilvirkningsformer ifgl. metodelitteraturen.**

|                   | Fordele   | Ulemper  |
|-------------------|---|--|
| Besøgs-interview  | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Mange temaer</li> <li><input type="checkbox"/> Mange spørgsmål</li> <li><input type="checkbox"/> Komplekse spørgsmål</li> <li><input type="checkbox"/> Åbne spørgsmål</li> <li><input type="checkbox"/> Filterspørgsmål</li> <li><input type="checkbox"/> Visuelle hjælpemidler</li> </ul> <p><i>Andre fordele</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Svarmotivation</li> <li><input type="checkbox"/> Samarbejdsskabende</li> <li><input type="checkbox"/> Tillidsskabende</li> <li><input type="checkbox"/> Mulighed for svarkontrol</li> <li><input type="checkbox"/> Rettelse af misforståelser</li> <li><input type="checkbox"/> Gruppeinterview</li> </ul> | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Hensyn til respondent og interviewer</li> <li><input type="checkbox"/> Spørgsmål om det ”socialt ønskede”</li> <li><input type="checkbox"/> Ledende spørgsmål</li> </ul> <p><i>Andre ulemper</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Tidkrævende</li> <li><input type="checkbox"/> Omkostningskrævende</li> <li><input type="checkbox"/> Geografisk begrænsning</li> <li><input type="checkbox"/> Relativt små stikprøver</li> <li><input type="checkbox"/> Interviewereffekter</li> <li><input type="checkbox"/> Interaktionseffekter</li> <li><input type="checkbox"/> ”Udenomssnak”</li> </ul>  |
| Telefon-interview | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Bedre end besøg ved sensitive spørgsmål</li> <li><input type="checkbox"/> Filterspørgsmål</li> </ul> <p><i>Andre fordele</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Relativ anonymitet</li> <li><input type="checkbox"/> Personlig sikkerhed</li> <li><input type="checkbox"/> Tidsbesparende</li> <li><input type="checkbox"/> Relativt lave omkostninger</li> <li><input type="checkbox"/> Relativt få interviewere</li> <li><input type="checkbox"/> Store stikprøver</li> <li><input type="checkbox"/> Geografisk spredning</li> <li><input type="checkbox"/> Færre interviewereffekter end ved besøg</li> </ul>   | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Hensyn til respondent og interviewer</li> <li><input type="checkbox"/> Korte spørgsmål</li> <li><input type="checkbox"/> Sprogligt enkle spørgsmål</li> <li><input type="checkbox"/> Åbne spørgsmål</li> <li><input type="checkbox"/> Få svarkategorier</li> <li><input type="checkbox"/> Ingen visuelle hjælpemidler</li> </ul> <p><i>Andre ulemper</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Manglende telefondækning</li> <li><input type="checkbox"/> Manglende svarmotivation</li> <li><input type="checkbox"/> Begrænset tid</li> <li><input type="checkbox"/> Ufleksibelt ift. besøg</li> <li><input type="checkbox"/> Interviewereffekt, fx fonologisk</li> <li><input type="checkbox"/> Interaktionseffekter</li> </ul> |
| Post-spørgeskema  | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Kun hensyn til respondent</li> <li><input type="checkbox"/> Visuelle hjælpemidler</li> <li><input type="checkbox"/> Sensitive spørgsmål</li> </ul> <p><i>Andre fordele</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Anonymitet</li> <li><input type="checkbox"/> Tidsbesparende</li> <li><input type="checkbox"/> Små omkostninger</li> <li><input type="checkbox"/> Store stikprøver</li> <li><input type="checkbox"/> Geografisk spredning</li> <li><input type="checkbox"/> Eget svar tempo</li> <li><input type="checkbox"/> Ingen interviewer- eller interaktionseffekter</li> </ul>  | <p><i>Spørgeskema</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Selvforklarende</li> <li><input type="checkbox"/> Korte spørgsmål</li> <li><input type="checkbox"/> Sprogligt enkle spørgsmål</li> <li><input type="checkbox"/> Få åbne spørgsmål</li> <li><input type="checkbox"/> Ikke kundskabsspørgsmål</li> <li><input type="checkbox"/> Særlig formattering</li> </ul> <p><i>Andre ulemper</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Manglende svarmotivation</li> <li><input type="checkbox"/> Relativt stort bortfald</li> <li><input type="checkbox"/> Ingen kontrol af respondent</li> <li><input type="checkbox"/> Reduceret læsefærdighed</li> <li><input type="checkbox"/> Ingen svarkontrol</li> </ul>   |

spørgeskemaer udarbejdes (se: *oversigt 4*). Efter nogle forfatteres opfattelse er spørgeskemakonstruktion en vanskelig opgave, hvis løsning skal honorere adskillige krav: "We need to frame questions that are meaningful, sensitive, precise, searching, and salient to our respondents." (Aldridge & Levine 2001:95). Selv om der er konsensus om adskillige kriterier, er der langtfra enighed om, hvordan "gode" spørgeskemaer udarbejdes, ligesom mange kriterier er diffuse og vanskelige at implementere. Derfor er heller ikke Hansen & Andersen's (2000) anbefalinger, der baseres på mangeårige erfaringer med surveymetoden, "neutrale".

Om formulering af spørgsmål peger nogle forfattere på, at de skal tildeles en eksplicit relation til problemstillinger. Endvidere advokeres for, at spørgsmål formuleres grammatisk korrekt (se fx: Fink 1995; Fowler 1993), hvilket ikke nødvendigvis er kongruent med "hverdags-sprog". Ét aspekt, som flere advarer imod, er brug af negationer. I dele af litteraturen peges der endvidere på, at spørgsmål skal være fokuserende. Fokusering kræver, at spørgsmål kun retter sig imod ét emne. Spørgsmålslængde er yderligere et aspekt. Selv om mange forskningsresultater problematiserer anbefalingen af spørgsmål med få ord, er efter-

lysningen af *brevity* ofte en *common sense*-anbefaling: "The shorter the question the less confusing and ambiguous it will be." (de Vaus 1998:83). Det gælder også Hansen & Andersen (2000).

Ét krav, som mange tildeler forrang, er éntydighed: "Clarity demands that virtually everyone interprets the question in exactly the same way." (Alreck & Settle 1995:89). Der advares imod vaghed: "The best way to avoid ambiguity is to use short, crisp, simple questions." (de Vaus 1998:84). Hvordan "éntydige" spørgsmål formuleres, indeholder litteraturen dog kun få ledetråde for. Ét vigtigt aspekt er reflekteret ordvalg. Alreck & Settle (1995) anbefaler fx fravalg af "flertydige" ord. Andre peger på, at "tekniske" termer og fremmedord bør undgås. Atter andre pointerer, at ord bør være uden "bias". En tilsvarende anbefaling goes af Hansen & Andersen (2000), der advarer imod "værdiladede" ord.

Yderligere en *common sense*-anbefaling er neutrale og balancerede spørgsmål, dvs. fravalg af ledende og bias-behæftede spørgsmål. Problemet er, at afsender- og modtagerperspektivet – surveyforsker versus respondent – ikke nødvendigvis er sammenfaldende. Nogle forfattere anbefaler også, at spørgsmål formuleres på måder, der muliggør svar.

**Oversigt 4.** Aspekter vedr. konstruktion af spørgeskemaer ifgl. metodelitteraturen.

| Spørgsmål                                    | Svarkategorier   | Spørgeskema                                       |
|--|--|---|
| <input type="checkbox"/> Tematisk relatering | <input type="checkbox"/> Tilvalg/fravalg (lukkede versus åbne spørgsmål) | <input type="checkbox"/> Spørgsmålsrækkefølge     |
| <input type="checkbox"/> Grammatik           | <input type="checkbox"/> Spørgsmålsrelatering                            | <input type="checkbox"/> Omfang (antal spørgsmål) |
| <input type="checkbox"/> Fokusering          | <input type="checkbox"/> Antal kategorier                                | <input type="checkbox"/> Formatering (layout mv.) |
| <input type="checkbox"/> Længde (antal ord)  | <input type="checkbox"/> Balancering                                     | <input type="checkbox"/> Andre aspekter           |
| <input type="checkbox"/> Formulering         | <input type="checkbox"/> Ved ikke-kategorier                             |   |
| <input type="checkbox"/> Neutralitet         | <input type="checkbox"/> På midten-kategorier                            |   |
| <input type="checkbox"/> Respondentrelevans  | <input type="checkbox"/> Andre aspekter                                  |   |
| <input type="checkbox"/> Hukommelse          |  |   |
| <input type="checkbox"/> Tidsrum             |  |   |
| <input type="checkbox"/> Andre aspekter      |  |   |

Derfor bør spørgsmål rette sig imod respondenteres egne erfaringer: "The strength of survey research is asking people about their firsthand experiences." (Fowler, i: Bickman & Rog 1998: 365). Heraf følger, at fx hypotetiske spørgsmål og kausalspørgsmål fravælges. Blandt andre anbefalinger kan nævnes, at spørgsmål præciseres spatio-temporalt, dvs. tidsafgrænsende. Hertil kommer påpegnin af, at spørgsmål ikke bør være "truende", og at de ikke tildeles "prestige bias". Endelig anbefales, at spørgsmål afholder sig fra virkelighedsferne antagelser om respondenter.

Også når det handler om svarkategorier, indeholder metodelitteraturen vekslende anbefalinger. Ét problem er, om spørgsmål skal efterfølges af svarkategorier eller ikke. Mens nogle forfattere anbefaler lukkede spørgsmål (se fx: Babbie 1990), har andre mere nuancerede opfattelser (se fx: Converse & Presser 1986). Trods vekslende opfattelser er survey-spørgsmål dog i almindelighed helt eller delvist lukkede. I så fald bør svarkategorier relateres til og være sammenhængende med spørgsmålet. Én forudsætning, som også Hansen & Andersen (2000) påpeger, er, at svarkategorier er udtømmende og gensidigt eksklusive. Angående sidstnævnte kriterier synes der at være metodologisk konsensus, men derimod ikke om, hvordan det tilsikres, at spørgsmål og svarkategorier er kongruente, således at kategorier udgør "logiske" og betydningsmæssige forlængelser af spørgsmål.

Blandt andre problematikker er antallet af svarkategorier og balancen disse imellem. Hvad den første angår, er anbefalingerne tvetydige. Megen litteratur er uden anbefalinger, mens andre peger på, at mængden af svarkategorier "afhænger af", men dog bør være mindst mulig (se fx: Converse & Presser 1986). Atter andre pointerer, at antallet beror på, hvor-

dan data tilvirkes (se fx: de Vaus 2002). Mht. den anden problematik gør adskillige forfattere opmærksom på, at svarkategorier bør balanceres, når dette er påkrævet. Yderligere problematikker angår brug af *ved ikke*- og *på midten*-kategorier, hvoraf sidstnævnte især kommer på tale i tilknytning til holdningsspørgsmål med evaluerende kontinuua. Heller ikke her er anbefalingerne konsensuelle, men i højere grad "det kommer an på"-råd. Mens nogle metodeforfattere mener, at *ved ikke*-kategorier bør undgås, er andre af den opfattelse, at den slags svar er vigtige og anvendelige informationer. Meningerne er ligeledes delte mht. brug af *på midten*-kategorier. Hvad Hansen & Andersen (2000) angår, er de tvetydige mht. de nævnte kategorier.

Først når spørgsmål udgør en sammenhængende helhed, har forskere konstrueret et *spørgeskema*, der er mere end summen af de spørgsmål, der indgår i det: "It is a totality, a gestalt .." (Labaw, 1980:12). Hvad spørgsmålsrækkefølge angår, peger flere forfattere på, at den har afgørende betydning: "A poorly organized questionnaire can confuse respondents, bias their responses, and jeopardize the quality of the entire research effort." (Rea & Parker 1997:35). Én gennemgående opfattelse i ikke-forskningsbaseret metodelitteratur synes dog at være, at spørgsmål grupperes med samhørende emner som kriterium (se fx: Fink & Kosecoff 1998; Johannessen & Tuftte 2002). Ét argument, der også fremføres af Hansen & Andersen (2000), til støtte for tematiske spørgeforløb er, at spørgsmålsrækkefølgen skal virke "naturlig". Problemet er dog, at netop tematiske spørgeforløb undertiden genererer konteksteffekter.

Anbefalingerne vedr. "gode" spørgeskemaer viser, at der langt fra er konsensus herom: "No one has ever succeeded in compiling a perfect list of rules for writing good questions .." (Salant & Dillman

1994:91). Med afsæt i metodelitteratur, hvis forfattere anskuer spørgeskemakonstruktion som "håndværk" eller "kunst", er litteraturen karakteriseret ved *common sense*-regler. Blandt "reglerne" findes fx anbefalinger om brug af "hverdagsprog" og formulering af præcise/korte spørgsmål. Spørgsmålet er imidlertid, om den slags "regler" fører til "gode" spørgsmål. Mange regler er for diffuse til at kunne implementeres, hvortil kommer, at andre regler er gensidigt eksklusive. Når metodeforfattere fx anbefaler "præcise" spørgsmål, kan det næppe anfægtes, men derimod komme i konflikt med en anbefaling af korte spørgsmål.

### Spørgeskema og respondent

Andre positioneringer, som fx Hansen & Andersen (2000) er skeptiske overfor, afviser spørgeskemakonstruktion som "kunst" eller "håndværk". Der har i en årrække været bestræbelser i retning af forskningsbaseret spørgeskemakonstruktion ved at "... building bridges between various subdisciplines of psychology, linguistics and survey methodology ..." (Bradburn 2000:in). Det gælder fx CASM (*Cognitive Aspects of Survey Methodology*), der har haft stor indflydelse på amerikansk surveyforskning. Der findes en righoldig metodelitteratur, hvori det anbefales, at spørgeskemakonstruktion baseres på analyser af respondenters svarprocessuelle opgaveløsninger. Selv om der endnu ikke er udviklet en *grand theory*, som forklarer alle svarprocessuelle aspekter, imødeser mange forfattere, at et intensiveret samarbejde mellem surveyforskere, kognitionspsykologer og lingvister vil bidrage til reduktion af måleproblemer.

Artiklens anden problemstilling angår respondenters opgaveløsninger og måleproblemer som følge heraf. At respondenter forstår spørgsmål og (re)konstruerer informationer på samme måde

er urealistisk. Det er en position, som også Hansen & Andersen (2000) synes at bifalde, om end de fastholder stimulus-respons-modellen som ideal. Andre forfattere drager konsekvensen og erstatter stimulus-respons-modellen med modeller, der sætter fokus på respondenters opgaveløsninger. En ofte citeret model med komponenter, der ikke nødvendigvis er successivt sammenhængende, hidrører fra Tourangeau et al. (2000): (1) Spørgsmålsforståelse (*comprehension*) '!(2) Informationsgenskabelse (*retrieval*) '!(3) Svarbedømmelse (*judgment*) '!(4) Svarafgivelse (*map onto response category*).<sup>3</sup> Idet hver komponent kan generere måleproblemer, er modellens formål at forstå respondenters svarprocesser med sigte på at reducere måleproblemer. Modellen ekskluderer tendentielt kommunikation – interaktion mellem respondenter og interviewere – som årsag til mangelfulde data. Det er en urealistisk forudsætning, der ofte problematiseres, men som oprettholdes af hensyn til loyal litteraturkondensering: Hvordan forstår respondenter ord, spørgsmål og svarkategorier? Hvordan (re)konstruerer respondenter informationer? Hvilken rolle spiller konteksteffekter for måleproblemer? Hvordan bedømmer eller redigerer respondenter, hvad "passende" svar er?

### Spørgsmålsforståelse

Mange metodeforskere peger på, at respondenters spørgsmålsforståelse genererer måleproblemer. Metodeforskning, der underbygger, at valg af enkeltord påvirker respondenters forståelse og dermed genererer måleproblemer, er omfattende. Selv små ændringer i ordvalget ændrer respondenters forståelse af spørgsmål. En forklaring er: "Given a question with ambiguous words, respondents may not ask for clarification, but may interpret the question as best they can ..." (Bradburn & Sudman, i: Biemer et

al. 1991:33). Begrænsning af ords betydningsvidde kan kræve yderligere ord, hvilket kan føre til andre måleproblemer, fx pga. mangelfuld *working memory*, dvs. korttidshukommelse. Måleproblemer grundet ordforståelse er både påvist ved besvarelse af faktuelle spørgsmål og holdningsspørgsmål, men oftest ved måling af holdninger (se fx: Groves 1989; Schuman & Presser 1996).

I sprogpraksis er sondringen mellem ord- og spørgsmålsforståelse en abstraktion. Derfor sætter metodeforskere også fokus på forståelse af spørgsmålshelheder. Metodeforskning peger på, at måder, hvorpå respondenter forstår spørgsmål, genererer måleproblemer, og at spørgsmålsforståelse undertiden fører til *ikke-formresistente korrelationer*, dvs. over direkte indflydelse på statistiske sammenhænge variabler imellem (se fx: Schuman & Presser 1996).

I enkelte undersøgelser sættes der fokus på grammatik, mens størstedelen angår semantik. Hvad spørgsmåls grammatiske struktur angår, peger flere forskere på, at strukturen bør være sproglig konsistent og ekskludere dobbelte negationer. Mht. sammenhænge mellem spørgsmåls grammatik og spørgsmålsforståelse er det velkendt, at syntaks- kan ændre spørgsmåls betydning. Imidlertid har forskningen kun bidraget marginalt til belysning heraf.

Når det handler om spørgsmålssemantik, er metodeforskningen righoldig. Det fremgår, at respondenters forståelse af både faktuelle spørgsmål og holdningsspørgsmål genererer måleproblemer. Ligesom forståelse af enkeltord navnlig fremkalder måleproblemer ved besvarelse af holdningsspørgsmål, gælder tilsvarende forståelse af spørgsmåls-helheder. Flere undersøgelser peger på, at moderate ændringer af holdningsspørgsmål fremkalder signifikant forskellige svar og svarfordelinger. Siden

1970'erne har navnlig amerikanske metodeforskere gennemført undersøgelser af måleproblemer grundet spørgsmålsforståelse. 1970'ernes og 1980'ernes forskning bidrog til voksende problematisering af respondenter som medier for datatilvirkning og førte til ledetråde for udarbejdelse af spørgsmål. Converse & Presser (1986:in) anbefaler fx "simple language" og efterlyser "simple questions rather than general ones" tillige med "clear definitions for terms of shared meaning."

I de senere år har amerikanske og andre forskere ydet fortsatte bidrag til udforskning af, hvordan respondenter forstår spørgsmål. Eksempelvis påpeger Sudman et al. (1996), at spørgsmålsforståelse såvel fordrer fokus på *literal meaning* som *intended meaning*:

Even if comprehending the literal meaning of the question poses no problem, respondents still have to infer what the specific information is that the researchers are interested in (Sudman et al., 1996:63).

I så henseende gør respondenter fx brug af svarkategorier, ligesom også konteksten spiller en rolle. Derfor betyder tilsyneladende "ens" spørgsmål ikke det samme for forskellige respondenter (se fx: Lyberg et al. 1997). Ifølge Tourangeau et al. (2000) er det surveyforskeres opgave at reducere det "semantiske rum", hvilket fordrer, at spørgsmål og svarkategorier er så éntydige som muligt, fx ved fravalg af kompliceret syntaks, diffuse ord, vage adverbelle kvantifikatorer og utilsigtede forudsætninger (*presuppositions*). Men uanset bestræbelser i retning af sproglig "entydighed" kan spørgsmål næsten altid forstås på flere forskellige måder, hvilket problematiserer standardiseret datakonstruktion (se også: Olsen 2001).

Ét tredje aspekt er spørgsmålslængde. Ikke alle metodeforskere anbefaler spørgsmål med få ord, eftersom lange spørgsmål kan fremkalde mere valide målinger. Bateson påpeger, at lange spørgsmål, der fx præciserer vage ord, kan reducere antallet af mulige betydninger: "There seems no doubt that, under some circumstances at least, long questions produce better data than short ones..." (Bateson 1984:59). Andre eksperimenter fører til anbefaling af vekslen mellem lange og korte spørgsmål. Det er ikke overraskende, hvis lange spørgsmål undertiden er at foretrække. Det skyldes, at antallet af betydningsudfald kan reduceres ved at øge spørgsmåls længden, der – når den øges – til gengæld belaster *working memory*.

I surveyundersøgelser er spørgsmål normalt lukkede, dvs. efterfølges af svar-kategorier. De fleste metodeforskere synes enige om, at *åbne spørgsmål* bør anvendes i moderat omfang. Én årsag er surveys standardiserende "logik", en anden pragmatisk: "Data from open-ended questions are much more time consuming to read, reduce and analyze." (Nolinske 1998:in). Det spiller også en rolle, at "samme" åbne og lukkede spørgsmål fører til statistisk inkongruente svarfordelinger (se fx: Schwarz, i: Biemer et al. 1991). Der er ikke konsensus om, hvornår åbne spørgsmål bør anvendes. I nogle undersøgelser vises fx, at "socialt uønsket" adfærd bedst belyses ved brug af åbne spørgsmål (se fx: Bradburn & Sudman 1979).

Når respondenter gøres bekendt med *svarkategorier*, udgør de en del af forståelsesrammen. Antallet af ord, der skal fastholdes i *working memory*, øges, men samtidig muliggøres reduktion af mulige betydningsudfald. To "identiske" spørgsmål, hvoraf det ene er åbent og det andet lukket, er derfor semantisk uens. Svarkategorier bidrager til at lede respon-

denter i specifikke forståelses- og svarretninger. Ligesom forståelse af spørgsmål er socio-kulturelt specifik, gælder tilsvarende svarkategorier. At svarkategorier påvirker respondenters besvarelser har været kendt længe, hvorimod ikke mange forskere har forsøgt at belyse de bagvedliggende kognitive processer. Svarkategorier er ikke "tekniske", men har afgørende betydning for måleproblemer. Det er derfor afgørende, at svarkategorier nøje overvejes. Som foreslået af Tourangeau et al. (2000) kan problemet løses ved, at svarkategorier oplyses og reducerer det "semantiske rum", fx ved at kategorier udgør semantiske forlængelser af spørgsmål.

Angående *ved ikke*-svarkategorier er det ifgl. Hansen & Andersen (2000) "normal praksis", at muligheden ikke tilbydes. Når det handler om faktuelle spørgsmål, mener de fleste metodeforskere, at *ved ikke*-svar bør undgås. Ikke desto mindre peger undersøgelser på, at svarfordelinger ændres, når muligheden tilbydes, hvilket antyder, at respondenter "rekonstruerer" informationer, de ikke har genskabt (se fx: Poe et al. 1988). Mht. holdningsspørgsmål peger mange på, at *ved ikke*-kategorier bør tilbydes, fordi ikke alle har holdninger til hvad som helst. Angående neutrale *på midten*-svarkategorier tilbyder Hansen & Andersen (2000) en "det afhænger af"-position. Når på midten-kategorier tilføjes, ændres svarfordelinger. Ikke blot vælger flere på midten-svar, når muligheden foreslås, men samtidig fremkommer færre *pol-svar*.

Også svarkategoriernes *antal og rækkefølge* kan generere måleproblemer via *registereffekter*. Der peges på, at registereffekter skyldes et samspil mellem forståelse og vurdering af emnelister, som respondenter tilbydes. I spørgsmål, hvor respondenter skal vurdere forskellige emner, der præsenteres på en liste,- kan

forståelse-af emner, der står øverst på listen, danne en semantisk ramme, som påvirker forståelse af senere emner (se fx: Schwarz & Sudman 1992). Forståelsesoverførsel listede emner imellem kan bevirke, at det mest "tiltalende" emne tildeles det mest positive svar. Afhængigt af antal svarkategorier, kategoriernes art og af datatilvirkning (besøg, telefon, post), forekommer både *først-* og *sidsteffekter* (se fx: Bradburn 2000). Undersøgelser peger på, at førsteffekter findes, når respondenter får udleveret svarlister. Førsteffekter beror på, at respondenter vælger "rimelige" initialt præsenterede emner. Hvis alle emner synes rimelige, foretrækker respondenter ofte det eller de emner, der står øverst på listerne. Formodentlig skyldes førsteffekter primært dysfunktionel *working memory*. Sidsteffekter, derimod, forekommer, når svarkategorier læses op, fx ved telefoninterview.

### **Informationsgenskabelse**

Forskning om respondenteres informationsgenskabelse sætter fx fokus på:

.. such processes as adopting a retrieval strategy, generating specific retrieval cues to trigger recall, re-collecting individual memories, and filling in partial memories through inference (Tourangeau et al. 2000:9).

Mere præcist handler metodeforskningen, der forbigås i tavshed af Hansen & Andersen (2000), navnlig om følgende aspekter: episodisk versus generel hukommelse, forglemmelse, teleskopi, hændelsers karakteristiske træk (fx *turning points*), kognitiv magelighed (*satisficing*) og kognitive aflastninger (*landmarks, aided recall* etc.).

Hvad det første aspekt angår, er langtidshukommelse både situationel og ikke-situationel. Enkelte forskere har gen-

nemført undersøgelser, som peger på, at respondenteres generelle hukommelse undertiden har forrang ved besvarelse af episodiske spørgsmål. Burton & Blair (1991) peger fx på, at respondenter, der genskaber informationer episodisk, undlader at informere om alle begivenheder, mens respondenter, hvis svar navnlig beror på ikke-episodisk hukommelse, ofte overvurderer antallet. Rekonstruktionsprocesser, der sigter mod genskabelse af episodiske informationer, men som beror på generel hukommelse, forekommer især, når forskere ønsker informationer om gængse hændelser, og/eller når retrospektive tidshorisonter er lange. Derfor anbefaler mange, at tidshorisonter reduceres (se fx: Sudman et al. 1996).

Mht. det andet aspekt påvirker flere parametre episodisk genskabelse. Én er, at den ønskede information aldrig er tilgængelig og derfor ikke kan rekonstrueres. Hvis informationen er tilegnet, kan der foreligge *simple failure to recall*. Desuden kan informationen kontamineres med andre. Rekonstruktion er processer, hvor der successivt "fyldes på", drages slutninger og tilføjes plausible detaljer:

Survey questions about factual matters require respondents to recall and to report information about their own lives (...). Distorted memories about these matters usually imply response errors in survey data (Shum & Rips, i: Sirken et al. 1999:96).

Tidsmæssige afstande mellem tilegnelses- og genskabelsestidspunkt påvirker genskabelse. Engang troede Ebbinghaus, at informationsgenskabelse kunne beskrives ved en logaritmisk glemselskurve. Imidlertid forsvarer kun få kognitionsforskere i dag kurven. At forglemmelse ikke er simple virkninger af "tidens

gang“ skyldes, at visse informationer til-egnes sikrere end andre. Informationer om emotionelt ladede begivenheder rekonstrueres ofte uproblematisk. Derimod glemmes stereotype begivenheder ofte hurtigt. Ebbinghaus' glemselskurve underbygges empirisk i nogle tilfælde, men ikke i andre. Alt i alt er der derfor grund til at drage Ebbinghaus' kurve i tvivl. Genskabelse af episodiske informationer er ikke nødvendigvis en simpel funktion af *recall period*, men “.. an interplay between the past, the present, and implicit theories of change, stability, and relationships among attributes.” (Tanur 1992: 83). At Ebbinghaus' kurve ikke underbygges modsigelsesfrit, implicerer dog ikke, at respondenter ikke *ceteris paribus* har dårligere hukommelse, jo længere den tidsmæssige afstand er fra tilegnelse til genskabelse. Pointen er, at alt andet sjældent er lige.

Med hensyn til det tredje aspekt er der yderligere en årsag til, at rekonstruktion af episodiske informationer ikke er en simpel funktion af tiden fra tilegnelse til genskabelse, *teleskopi*. I en ofte citeret artikel udviklede Sudman & Bradburn 1973 en forglemmelsesmodel, der omfatter forglemmelse og teleskopisk tidskompression. Ved test af modellen peges på forekomst af såvel simpel forglemmelse som tidskompression. Modellen har siden været været genstand for kritik. Teleskopi underbygges empirisk, mens andre undersøgelser peger på moderate eller ingen effekter. I nogle undersøgelser vises, at hændelsers hyppighed øger sandsynligheden for fejlagtige tidsmæssige placeringer. Fravær af begivenheders karakteristiske træk bevirker, at de placeres forkert. Også karakteristiske træk ved begivenheder kan føre til teleskopi. Der peges på, at korte retrospektive tidshorisonter øger sandsynligheden for teleskopi, men generelt er forskningen om, hvornår teleskopi optræder, mangelfuld.

Hvad det fjerde aspekt angår, er visse hændelsers *karakteristiske træk* yderligere en årsag til måleproblemer. Uafhængig af tid genskabes informationer med karakteristiske træk mest pålideligt (se fx: Crano & Brewer 2002). Begivenheder, der er forbundet med emotionelle reaktioner, huskes fx ofte glimrende. Det kan være *turning points*, fx overværelse af et barns fødsel. Tilsvarende gælder andre begivenheder, der overskrider kognitive drejebøger. Derimod glemmes stereotype begivenheder ofte hurtigere end usædvanlige. Den betydning, som vendepunkter har for genskabelse af informationer om begivenheder, synes overensstemmende med en skemateoretisk tilgang, men med den tilføjelse, at der bør bygges bro mellem respondenteres kognitive og emotionelle processer.

Det femte aspekt er *kognitiv magelighed*. Nogle respondenter genskabelse af episodiske og/eller generelle informationer, fx holdninger, hæmmes af kognitive belastninger. Uanset om spørgsmål er tilpasset respondentens hukommelse eller ej, kræver genskabelse kognitiv kompetence. Nogle spørgsmål kræver genskabelse af mere komplekse informationer end andre. Hertil kommer, at nogle respondenter er mere kognitivt øvede end andre. For mindre øvede bliver virkningen kognitiv magelighed, der fører til tilfældig “konstruktion” af informationer. Det gælder navnlig ved besvarelse af holdningsspørgsmål, hvor det har vist sig, at nogle respondenter har “holdninger” til fiktive fænomener (se fx: Schuman & Presser 1996).

Pseudo-genskabelsesstrategier kalder Krosnick (1991) for *satisficing*. Behævningsstrategisk “genskabelse” afhænger ifgl. Krosnick af følgende parametre:

.. the first is the inherent *difficulty of the task* that the respondents confronts; the second is the respon-

dent's ability to perform the required task; and the third is the respondent's motivation to perform the task (Krosnick 1991).

*Satisficing* gør svar til "mentale møntkast". Selektion af *satisficing*-respondenter er ikke stokastisk, da undersøgelser tyder på, at navnlig respondenter med begrænset kognitiv kapacitet er behagestrategiske. Kognitiv magelighed og behagestrategier indebærer mangelfuld informationsgenskabelse, idet respondenter vælger "den første den bedste" svarkategori. Svarkategorier af typen *ved ikke* eller *har ingen mening* synes i nogen grad at kunne forebygge kognitivt magelige svar. Men selv når den slags svarkategorier tilbydes, praktiserer nogle respondenter *satisficing*. Kognitiv magelighed forekommer også i tilknytning til besvarelse af holdningsspørgsmål med neutrale på midten-svarkategorier.

Ikke kun holdningsspørgsmål, men også episodiske spørgsmål fører til kognitivt magelige svar. Sværhedsgraden af spatio-temporale spørgsmål bør tilpasses respondenters hukommelse. Det kan gøres ved at reducere spørgsmåls tidshorisonter (se fx: Tanur 1992). Spørges der fx til respondentens fritidsaktiviteter inden for den seneste måned i stedet for inden for de sidste tolv, bliver genskabelse af informationer herom mindre mangelfuld. Som ledetråd for formulering af såvel faktuelle spørgsmål som holdningsspørgsmål bør gælde, at spørgsmåls kognitive sværhedsgrad reduceres mest muligt samtidig med, at deltagelsesmotivation maksimeres.

Det sidste aspekt angår *kognitive afslutninger*. Hvor Hansen & Andersen's (2000) lærebog undlader at pege på anbefalinger vedr. reduktion af genskabelsesproblemer, gælder det modsatte store dele af den øvrige metodelitteratur (se fx: Tanur 1992; Tourangeau et al. 2000). Én

anbefaling er, at "gode" spørgsmål og spørgeskemaer forudsætter, at forskere tilegner sig indsigt i, hvordan hukommelse fungerer, og i hvad man kan forvente af respondenter. Fremstillingen om episodisk versus generel hukommelse peger fx på, at spørgsmål bør formuleres på måder, der aktiverer relevante genskabelsesstrategier. En anden anbefaling er, at retrospektive spørgsmåls tidshorisonter reduceres. Yderligere en anbefaling angår hændelsers karakteristiske træk. Eftersom analoge hændelser øger sandsynligheden for mangelfuld genskabelse af informationer herom, bør dette indtænkes ved formulering af spørgsmål. Hvad holdningsspørgsmål angår, anbefaler mange metodeforskere anvendelse af *ved ikke*-svarkategorier, der ikke løser *satisficing*-problemet, men som i nogen grad forebygger måleproblemer.

Vilkår for informationsgenskabelse spiller *per se* en rolle for, om genskabelse fører til måleproblemer. Ikke kun det tilgængelige og måden, hvorpå det er tilegnet, spiller en rolle. Følgelig kan episodisk informationsgenskabelse fremmes af vink (*cues*), især hvis disse er beslægtet med omstændigheder, som fandtes under selve tilegnelsen. En type af vink er *landmarks*. Hvis det fx ønskes oplyst, hvad respondenter foretog sig i en uge, hvor en almindelig kendt begivenhed – fx angrebet på WTC – fandt sted, kan interviewere informere om begivenheden. Anvendelse af tidlige orienteringspunkter er navnlig påkrævet, når spørgsmål omhandler tidsrum over tolv måneder, da der ikke findes temporale skemaer, som overskrider årets skemaer.

Andre vink er *aided recall*, *bounded recall* og *record checks* (se fx: Dijkstra & van der Zouwen 1982). Genskabelse kan fremmes ved, at interviewere læser svarkategorier op, dvs. *aided recall*. Desuden kan *probes* fremme genskabelse. Hvis respondenter deltager i en longitudinel sur-

vey, kan genskabelse fremmes ved at gentage oplysninger, som respondenter tidligere har afgivet (*bounded recall*). Har respondenter fx for ti år siden oplyst antallet af egne tænder, og ønskes der påny informationer herom, er det nærliggende at informere om det dengang oplyste antal. Endelig kan også dokumentation – *record checks* – virke fremmende for genskabelse af informationer, fx selvangivelser, dagbøger, lejekontrakter o.l.

Yderligere aflastninger kan fremme genskabelse. Der peges fx på en sammenhæng mellem genskabelse og den svar-tid, som respondenter tildeles. Jo mere tid, respondenter råder over ved besvarelse af spørgsmål, desto gunstigere er udsigten til pålidelig genskabelse. I så henseende synes postspørgeskemaer fordelagtige, mens telefoninterview almindeligvis tilbyder ringe muligheder. Bl.a. grundet tidsaspektet anbefaler flere metodeforskere lange spørgsmål. Genskabelse kan ligeledes stimuleres af flere spørgsmål om samme anliggende. Sidstnævnte anbefaling kan dog komme i konflikt med for store krav til respondenteres *working memory*.

### **Konteksteffekter**

Hansen & Andersen 2000 peger på, at besvarelse af ét spørgsmål kan påvirkes af det hidtidige interviewforløb. Konteksteffekter bør "så vidt muligt" søges kontrolleret, påpeger forfatterne, men foreslår ikke hvordan. De mener, at spørgsmålsrækkefølge skal være "naturlig": "Derfor samles spørgsmål om samme emne eller tema så godt som altid i spørgsmålsserier." (Hansen & Andersen 2000:135). Men måske er tematiske spørgeforløb én årsag til konteksteffekter? Ikke kun forståelse af ord, spørgsmål og svar-kategorier fører til måleproblemer. Det gælder også konteksten, der som kognitiv overførsel genererer konteksteffekter. Konteksteffekter har været kendt længe,

men der er ikke enighed om hvorfor og hvor ofte, konteksteffekter genereres. Konteksteffekter indebærer, at spørgsmålsforståelse og andre kognitive processer afhænger af den sammenhæng, hvori spørgsmål besvares: "... question comprehension is not merely a function of the wording of the question itself." (Sudman et al. 1996:69). Spørgsmål forstås og svar afgives inden for kontekstens tavse baggrundsforståelse.<sup>4</sup>

Amerikansk og anden metodeforskning sætter fokus på kontekster som årsag til måleproblemer (se fx: Schwarz & Sudman 1992). Fokuseringen er ikke overraskende, eftersom konteksteffekter har vist sig at bevirke, at respondenter besvarer "samme" spørgsmål forskelligt i uens kontekster. Konteksteffekter kan med andre ord bevirke, at spørgsmål, der fører til valide målinger i én sammenhæng, ikke gør det i en anden. Metodeforskning underbygger, at konteksteffekter i større eller mindre omfang er årsag til måleproblemer. Med forbehold for, at der er gennemført flere eksperimenter med holdningsspørgsmål end med faktuelle, synes konteksteffekter især at forekomme i tilknytning til besvarelse af førstnævnte. Tourangeau og andre forklarer tendensen med, at holdninger i højere grad end faktualiteter er skrøbelige *ad hoc*-konstruktioner (Tourangeau, i: Sirken et al. 1999).

Konteksteffekter forekommer især i tilknytning til tematisk sammenhængende spørgeforløb, hvilket problematiserer Hansen & Andersen's (2000) anbefaling af spørgsmålsserier. Flere forskere peger på, at sandsynligheden for konteksteffekter er størst, når holdningsspørgsmål er tematisk beslægtede.<sup>5</sup> Afsmitning kaldes undertiden *carryover effect*. Overførsel finder sted, når spørgsmål danner en forståelsesramme, som påvirker efterfølgende svar:

Carryover effects occurs when the interpretation, retrieval, judgment, or response associated with prior items provides a respondent with an easily accessible cognitive structure or schema, by bringing the cognitive structure into short-term memory, into a temporary *workspace*, or to the top at the *storage bin* of relevant information. (Harrison & McLaughlin 1993: 53).

Uanset konteksteffektens art er det et problem, at der endnu ikke er udviklet effektive procedurer til forudsigelse og forebyggelse heraf. Nogle metodeforskere peger dog på en række forskningsbaserede anbefalinger, fx:

In general, the question or set of questions most vital to the research should be presented first, to avoid unwanted influence from preceding questions, due to their effects on information accessibility and current question interpretation (Tourangeau, i: Sirken et al. 1999:in).

Der peges også på moderat anvendelse af tematiske spørgeforløb, ligesom det understreges, at test af spørgeskemaer, fx ved gennemførelse af kognitive laboratorieinterview, kan bidrage til reduktion af konteksteffekter.

### *Svarredigering*

Selv hvis det usandsynlige skulle ske, at hverken spørgsmålsforståelse, informationsgenskabelse eller konteksten, hvori spørgsmål findes, genererer måleproblemer, er respondenteres *svarredigering* yderligere en årsag til måleproblemer. Ifølge Hansen & Andersen (2000) vækker visse spørgsmål erfaringsmæssigt stærke følelser, fx spørgsmål om rusmiddelbrug

og sort arbejde. Forfatterne synes af den opfattelse, at svarredigerings beskafenhed og omfang ikke kan forudsiges. Hvad viser metodeforskningen herom?

Flere undersøgelser peger på, at respondenter ofte redigerer genskabte informationer, før svar afgives (se fx: Schwarz, i: Lyberg et al. 1997; Tourangeau et al. 2000). Der peges på, at svarredigering ikke kun angår selvpræsentation af "socialt ønskede" adfærdsformer og holdninger, men er redigering i videre betydning, dvs. kondensering og formidling af, hvad respondenter opfatter som "passende" svar. Nogle respondenter vurderer nøje genskabte informations tilstrækkelighed, før svar afgives. Hvad besvarelse af holdningsspørgsmål angår, kan redigering spille en afgørende rolle for måleproblemer, da respondenter undertiden tilbageholder eller modificerer holdninger.

Sensitive spørgsmål er:

.. largely personal and directly threatening to the respondent's self-esteem or are anxiety-provoking – for the most part, questions about illegal or deviant behaviour, or about major health problems.. (Sudman & Bradburn 1974:84).

Når spørgsmål omhandler den slags anliggender, tilbyder nogle respondenter "socialt ønskede" svar. Selv om spørgsmål forstås og informationer rekonstrueres, afgives svar, der afspejler respondenterens kendskab til sociale normer, og som bl.a. afhænger af konkrete interaktioner mellem respondenter og interviewere. Tendensen er, at respondenter fremstiller sig som socialt sympatiske:

Studies of response accuracy suggest that there is a tendency among respondents to distort an-

swers in ways that will make them look better or will avoid making them look bad (Fowler, i: Bickman & Rog 1998:354).

Respondenters overvurdering af "socialt ønskede" tilstande og adfærdsformer kan være af marginal betydning, men er ofte markant (se fx: Groves 1989).

Ifølge Hansen & Andersen (2000) kan svarredigering ikke forudsiges og heraf resulterende måleproblemer derfor ikke forebygges. Dén opfattelse genfindes ikke i metodeforskningen, hvor der peges på forskellige procedurer til reduktion af over- og underrapportering (se fx: Groves 1989). Én anbefaling er tilsikring af anonymitet, idet nogle undersøgelser peger på, at målinger af socialt uønskede tilstande og adfærdsformer bliver mere valide, når respondenter ikke betvivler anonymiteten. Yderligere en anbefaling er at understrege vigtigheden af nøjagtige svar. Enkelte undersøgelser tyder på, at åbne spørgsmål muligvis er at foretrække ved besvarelse af sensitive spørgsmål. Imidlertid synes datatilvirkningsformen at være mere afgørende, eftersom undersøgelser peger på, at postspørgeskemaer er at foretrække.

## Spørgeskema og interview

Ved belysningen af den tredje problemstilling sættes der fokus på forskellige opfattelser af surveyinterview: Er opfattelserne konsensuelle mht., hvad der karakteriserer "gode" surveyinterview? Hvordan genererer interviewere måleproblemer, og lader problemerne sig forebygge? Hvordan gennemføres interviewerdannelse? Hansen & Andersen (2000) fastslår, at interviewereffekter kun er et problem i interviewundersøgelser: "Ved besøgsinterview er denne effekt normalt mest udtalt" (Hansen & Andersen 2000: 100). Når besøgsinterview ikke desto mindre ofte foretrækkes, skyldes det, at

spørgsmål om komplicerede emner belyses bedst på den måde, og at interviewere har mulighed for at observere, om respondenter føler sig usikre, og kan bidrage til at præcisere spørgsmåls betydning. Derudover indeholder interviewmetoden mulighed for brug af *probes* uanset besøgs- eller telefoninterview.

## Interviewopfattelser

I metodelitteraturen tildeles surveyinterview ofte prædikaterne *struktureret* eller *standardiseret*. Dette til trods er der forskellige interviewopfattelser (se: *oversigt 5*). Én opfattelse tager afsæt i stimulus-respons-modellen, der anses for at være et ideal, om end vanskeligt opnåeligt. Interviewopfattelser, hvis afsæt er stimulus-respons-modellen, anskuer – som fx Hansen & Andersen (2000) – det ideelle interview som genuint standardiseret. Kommissoriet bliver at gennemføre interview, der er så "neutrale" og påvirkningsfjerne som muligt: "... it is when interviewers *fail to be standardized* that they are responsible for error" (Fowler & Mangione 1990:13). I sidste instans forudsætter standardisering, at interviewere ikke påvirker afgivne svar. Opfattelsen, der findes i store dele af metodelitteraturen, har i årtier været gængs, men er i de senere år blevet udsat for kritik (se fx: Houtkoop-Steenstra 2000; Krosnick 1999).

## Oversigt 5. Opfattelser af surveyinterview i kvantitativ metodelitteratur.

- Stimulus-respons-modellen
- Uundgåelig socio-sproglig interaktion
- Særlig konversationsform
- Ikke-specifikationer

En anden opfattelse lægger afstand til stimulus-respons-modellen og tager *ad*

*notam*, at interview altid er socio-sproglig interaktion, der genererer måleproblemer. *Complete interviewer-uniformity* er aldrig mulig, men bør tilstræbes. Besøgs- og telefonintervieweres "fejl" og evt. "signalforvirringer" er ikke afvigelser fra et utopisk ideal, men genererer interviewer- og interaktionseffekter:

The informal behaviors an interviewer uses in eliciting an amplification of clarification of an unsatisfactory response, although relatively under-investigated, are of real importance (Crano & Brewer 2002:243).

Såvel ved besøgs- som telefoninterview påvirkes svarprocessen af intervieweres adfærd: "Interviewers affect responses through their individual performance patterns .." (Lyberg & Kasprzyk, i: Biemer et al. 1991:240).

Yderligere en opfattelse anskuer surveyinterview som en særlig form for talehandling. Ét eksempel er Houtkoop-Steenstra's afhandling (2000), der er kritisk over for Fowler & Mangione's (1990) standardiseringskrav: "... the participants in survey interviews orient toward many of the same interaction rules as participants in conversation" (Houtkoop-Steenstra 2000:12). For at kunne konstruere spørgeskemaer konversationelt, fordrer konversationsanalytisk viden, fx om at spørgsmål også er pragmatisk-interaktionelle "enheder", der bør formuleres oplæsningsvenligt, så de fremmer kommunikativt *flow*. Mange respondenter er ikke opmærksomme på de ekstremt "standardiserede" interaktionsregler, der gælder for surveyundersøgelser. En tilsvarende opfattelse gøres fx gældende af Schober: "Respondents make sense of survey questions in ways that are related to the ways they make sense of questions as addressees in ordinary conversation"

(Schober, i: Sirken et al. 1999:90).

Hertil kommer, at megen metodelitteratur ikke specificerer arten af interviewopfattelse. Heller ikke hvad opfattelser af surveyinterview angår, er der således konsensus i kvantitativ metodelitteratur og -forskning. Begrebet surveyinterview har derfor større betydningsvidde end almindeligvis antaget. Mens nogle forfattere forfægter mulighederne for gennemførelse af genuint standardiserede interview, er andre af den opfattelse, at afstanden herfra – fx ved at gøre interviewere til surveyforskere "talsmænd" – er en forudsætning for reduktion af måleproblemer. Bortset fra formelle kriterier har surveyinterview således ikke konsensuelle konstituerende træk. I sidste instans finder der næppe heller en præcis demarkationslinje mellem surveyinterview og andre forskningsinterview (se fx: Olsen 2002).

### *Interviewereffekter*

Hvordan genererer interviewere måleproblemer? Surveyinterview fordrer betydelige kvalifikationer, eftersom mangel herpå kan modarbejde selv den mest planlagte undersøgelse. Intervieweres oplæsningspraksisser, *probe*-strategier, svarregistreringsprocedurer og ubevidste signaler kan føre til måleproblemer. Men måske kan også interviewerkendetegn, fx køn eller etnisk baggrund, påvirke svarprocessen? Måleproblemer pga. *kendetegn* ved interviewere afhænger af, om data tilvirkes ved besøg eller telefonisk, idet fx etnisk baggrund næppe spiller en rolle ved telefoninterview. I dele af metodelitteraturen peges der på, at både intervieweres køn, alder, sociale status, uddannelsesmæssige baggrund, etniske tilhørsforhold og udseende i øvrigt kan påvirke svarprocessen og dermed generere måleproblemer.

Der findes en del forskning om, hvordan intervieweres demografiske og an-

dre kendetegn fører til måleproblemer, om end temaet er mangelfuldt belyst (se fx: Lyberg & Kasprzyk, i: Biemer et al. 1991). Når tendenserne er tvetydige, skyldes det især, at genuine kendetegneffekter forudsætter, at *respondenters* kendetegn ikke spiller en rolle, hvilket i mange tilfælde er urealistisk. Enkelte undersøgelser viser fx et samspil mellem interviewerens og respondenters køn og etniske baggrund:

.. characteristics of the interviewer and the respondent, such as race or sex, may influence the willingness to express attitudes or affect relations in ways that are independent of the interview and reflect larger social beliefs. (Bradburn 2000:in).

Selv om det *ceteris paribus* forventes, at interviewere og respondenter, der ligner hinanden, fx mht. etnisk baggrund, tilvirker de mest valide data, er der behov for yderligere forskning.

Da begrebet surveyinterview har markant betydningsvidde, er der ikke konsensus om arten af kompetencer, som interviewere skal mestre. Hansen & Andersen 2000 peger på, at surveyinterview fordrer brug af kvalificerede og erfarne interviewere, men præciserer ikke, hvad det indebærer. Hvad er den øvrige metodelitteraturs svar? Ligesom Hansen & Andersen understreger også andre, at spørgeskemaer bør være uden "fejl og svagheder". For andre er kravet så indlysende, at det end ikke nævnes. På den anden side er der ikke enighed om, hvornår spørgeskemaer er "mangelfulde". Det er derfor ikke overraskende, at metodelitteraturen er uden konsensus ang. kompetencekrav stillet til "gode" interviewere. De kompetencer, der præsenteres i *oversigt 6*, fremhæves i store dele af metodelitteraturen, men ikke alle kompe-

tencer nævnes i alle metodebøger, hvilket bl.a. skyldes, at kompetencer undertiden er gensidigt eksklusive, fx standardiseret ift. konversationskompetence.

#### *Oversigt 6. Interviewerkvalifikationer ifgl. kvantitativ metodelitteratur (eksempler).*

- Spørgeskemakompetence
- Interviewerinstruktionskompetence
- Neutralitetskompetence
- Flexibilitetskompetence
- Sprogkompetence
- Probe*-kompetence
- Konversationskompetence
- Svarregistreringskompetence

I dele af metodelitteraturen pointeres, at interviewere bør have *spørgeskemakompetence*, dvs. både ved læsning og i øvelsespraksis tilegne sig grundigt kendskab til spørgeskemaer (se fx: Babbie 1990). Kendskab til spørgeskemaer *per se* er imidlertid ikke tilstrækkeligt. Hvis interview skal "lykkes", kræves tillige, at interviewere har kendskab til og er i stand til at formidle undersøgelsens formål og dets sammenhæng med spørgeskemaer. Ét med spørgeskemakompetence beslægtet aspekt er *interviewerinstruktionskompetence* (se fx: Fowler, i: Bickman & Rog 1998). Interviewerinstruktioner omfatter både i spørgeskemaer indbyggede instruktioner og separate mere generelle instruktioner og oplysninger om en given undersøgelse. Alene dette stiller krav til forskere, der udarbejder spørgeskemaer.

Yderligere en kvalifikation – *neutralitetskompetence* – retter sig også imod såvel besøgs- som telefoninterview (se fx: Fowler & Mangione 1990). Kompetencen efterlyser neutral intervieweradfærd, "objektivitet" og ikke-vurderende ytringer. Selv om *complete standardized uniformity* er uopnåelig, anskues den af mange

som et ideal. Andre forfattere, derimod, mener, at interviewere nødvendigvis må praktisere ikke-standardiserede adfærdsformer, da respondenter er forskellige og fordrer uens praksis. Der er sammenhæng mellem, hvordan surveyinterview opfattes, og om neutralitetskompetence tildeles forrang. I litteraturen nævnes også, at interviewere bør tilegne sig *fleksibilitetskompetence*. Nogle forfattere efterlyser fx evne til under fire øjne at etablere en afslappet atmosfære. Kompetencekravet lægger ikke nødvendigvis afstand til standardisering som ideal, hvilket fx Sapsford (1999) fremhæver, når han efterlyser *standardized presentation*, men samtidig erkender vanskelighederne herved.

Da ytringer er datatilvirkningens medie, er *sprogkompetence* en vigtig kvalifikation, fx ved såvel ved gennemførelse af besøgs- som telefoninterview at fordrer generelle sproglige kvalifikationer, oplæsningskompetence og intonationsbevidsthed (se fx: Andersson 1994; Fink & Kosecoff 1998). Oplæsningskompetence er evne til oral formidling af spørgsmål, svar-kategorier mv. Forskningsresultater peger på, at interviewere har meget forskellige evner til at læse spørgsmål "korrekt" og "tydeligt" op med heraf følgende måleproblemer (se fx: Groves 1989). Oplæsningskompetence er også evne til at læse netop det op, som forskere anmoder om. Ikke overraskende understreger især forfattere, der bifalder stimulus-responsmodellen, behovet for neutral og standardiseret oplæsning. Dette stiller særlige krav til surveyforskere, da spørgsmål, som er vanskelige at læse op, lettere ændres med andre betydninger til følge. Oplæsningskompetence er også intonationsbevidsthed, da tonefald kan påvirke afgivne svar.

Yderligere et aspekt af sproglig kompetence er *probe-kompetence*, der sættes fokus på i store dele af metodelitteraturen, også af Hansen & Andersen (2000). Ikke

overraskende pointerer metodeforfattere, der bifalder standardiserede interview, at *probes* bør være neutrale (se fx: Babbie 1990). Ikke desto mindre peger forskningsresultater på, at *probe*-praksisser er markant forskellige. Nogle forfattere advarer imod "frie" situationelt bestemte *probes* og anbefaler, at kun *probes* anført i spørgeskemaer læses op. Begrundelsen er, at sidstnævnte almindeligvis fører til mindre interviewereffekter (se fx: Groves 1989). Andre mener, at interviewere bør praktisere forskellig *probe*-adfærd, da respondenter er forskellige (se fx: Sapsford 1999).

Yderligere en kvalifikation er *konversationskompetence*. Når interview opfattes som talehandlinger, bliver evne til at befordre kommunikativt *flow* afgørende. Da "standardiserede" interview er umulige, bør interviewere virke som surveyforskernes "talsmænd": „When some respondents are not able to extract the intended meaning from the question formulation, we need the interviewer to act as spokesman for the researcher.“ (Houtkoop-Steenstra 2000:181). Heraf følger særlige krav til interviewere om indsigt i surveyundersøgelsers sigte og dettes sammenhæng med spørgeskemaer. Endvidere bør interviewere acceptere uformaterede svar, der ikke matcher på forhånd udarbejdede svarkategorier. Interviewere bør også tildeles mulighed for at drage slutninger og afprøve disse gennem supplerende udspørgen. Endelig bør interviewere forklare respondenter surveyinterviews interaktionelle spilleregler.

Endelig er også *svarregistreringskompetence* vigtig. Ikke overraskende er metodeforfatteres anbefaling, at svarregistrering bør være præcis. Interviewere skal fx kunne bistå respondenter med at lokalisere "rigtige" svarkategorier. Anskuet fra respondenters synsvinkel er lukkede spørgsmål uden oplyste svarkategorier åbne, hvilket kan påføre inter-

viewere problemer, hvis svar ikke matcher svarkategorier. Hvad besvarelse af åbne spørgsmål angår, er metodelitteraturen tvetydig, idet det er uklart, hvor meget interviewere skal nedfælde. Svarregistreringskompetence stiller ikke kun krav til interviewere, men også til spørgeskemakonstruktører.

Med vekslende interviewopfattelser såvel som forskellige anbefalinger vedr. interviewer-kvalifikationer *in mente* er det næppe overraskende, at surveyinterview har vist sig at generere validitets- og reliabilitetsproblemer, der ikke kun påvirker målingers niveau, når variabler – fx procentfordelinger – anskues separat, men som undertiden også fører til ikke-formresistente korrelationer, (se fx: Groves 1989; Fowler & Mangione 1990; Lyberg et al. 1997). Hverken interviewopfattelser, interviewer-kvalifikationer eller andre ledetråde for gennemførelse af "gode" surveyinterview findes i en konsensuel "kogebog".

### **Intervieweruddannelse**

Med forskellige interviewopfattelser og krav til interviewkompetence *in mente* er der heller ikke konsensus om, hvad en "god" intervieweruddannelse er (se fx: de Vaus 2002; Fowler 1993; Tourangeau et al. 2000). I metodelitteraturen sættes der fokus på flere aspekter, hvoraf ét omhandler *indhold* og relaterer sig til de nævnte anbefalinger vedr. interviewer-kvalifikationer. Der er som nævnt ikke konsensus om krav, der befojet kan stilles til intervieweres kompetencer. Der er dog kompetencer, der synes at være tendentielt enighed om, fx spørgeskema- og interviewerinstruktionskompetence. Fordringen synes evident, da det er meningsløst at påbegynde interview uden kendskab spørgeskema mv.

Hertil kommer sprog- og *probe*-kompetence samt svarregistreringskompetence, om end opfattelserne af, hvad

kompetencerne mere præcist indebærer, er forskellige. Hvad derimod tilegnelse af neutralitets- og fleksibilitetskompetence angår, er opfattelserne modsætningsfyldte. Mange metodeforfattere peger på, at interviewuddannelse bør sigte imod tilegnelse af *complete standardized interview uniformity*. Ifølge forfattere, der bifalder standardisering, er fravær af personlige holdningstilkendegivelser fx ét vigtigt aspekt. Andre forfattere, derimod, opfatter standardisering som et umuligt forehavende og tildeler derfor uddannelse i fleksibilitetskompetence forrang.

Et andet aspekt angår *praktiske interviewøvelser mv*. At tilrettelægge intervieweruddannelse med loyalt afsæt i kvantitativ metodelitteratur er et vanskeligt forehavende, som med litteraturens tendenser *in mente* bl.a. fordrer stillingtagen til interviewopfattelse. Vanskelighederne bliver ikke mindre af, at der skal træffes afgørelse om arten og omfanget af praktiske interviewøvelser mv. Det er velkendt, at ingen bliver gode interviewere ved kun at læse bøger eller ved at deltage i undervisning herom. Ikke overraskende pointerer mange, at uddannelse også bør inkludere praktisk erfaring med gennemførelse af interview, men når metodelitteraturen anskues tværgående, afhænger også arten af anbefalede erfaringer af interviewopfattelse.

Interviewuddannelses tredje aspekt er *varighed*. Også her er erfaringerne delte. Fowler & Mangione (1990) peger fx på, at interviewere, der har deltaget i kurser af 2-4 dages varighed, tilvirker de bedste data, mens længerevarende uddannelse, fx 10-dages kurser, risikerer at bidrage til øgede måleproblemer, fordi interviewere bliver trætte og demotiveres. Også andre problematiserer, at lang intervieweruddannelse er en fordel, fx fordi erfarne interviewere kan have udviklet "dårlige vaner" (se fx: Alreck & Settle 1995).

Anskuet tværgående viser metodelitteraturen, at surveymetoden heller ikke mht. den her behandlede problemstilling er så standardiseret som almindeligvis antaget. Selv om det standardiserede "neutrale" interview tildeles en dominerende position i litteraturen, har andre interviewopfattelser i voksende omfang gjort sig gældende i de senere år. Heraf følger, at også krav til interviewerkompetence og uddannelse af interviewere er forskellige.

### **Kvalitetssikrende procedurer**

Artiklens sidste problemstilling omhandler procedurer til kvalitetssikring af surveydata. Kvalitetssikring forudsætter et konceptuelt grundlag herfor. Behovet skyldes, at ingen kan frigøre sig fra meta-teoretiske forudsætninger, og at "kvalitet" er et vanskeligt definer- og operationaliserbart begreb (se fx: Olsen 2002). Ligesom anden empiri ikke lader sig kvalitetssikre uden konceptuelt afsæt, gælder tilsvarende surveydata. Hvad kvalitetssikring af surveydata angår, peger næsten alle på validitet og reliabilitet som bærende kvalitetsbegreber, men uden konsensus om, hvordan de defineres og operationaliseres. Implementering af procedurerne skal således bidrage til at sikre surveydatas validitet og reliabilitet, dvs. til at reducere potentielle måleproblemer. Til virkeliggørelse heraf peger flere forfattere på, at reduktion af måleproblemer kun er mulig, hvis procedurer iværksættes med afsæt i tydeligt formulerede formål (se fx: Converse & Presser 1986; Czaja 1998). Ét formål kan fx være at afprøve, om respondenter har mulighed for og er villige til at løse tildelte kognitive opgaver. Et andet formål kan være at belyse, om interviewere er i stand til at udføre deres arbejde.

#### ***Prøveinterview***

Blandt kvalitetssikrende procedurer sæt-

ter Hansen & Andersen (2000) næsten udelukkende fokus på prøveinterview. Ved gennemførelse af prøveundersøgelser konfronteres spørgeskemaer med "virkelighedens verden". Proveundersøgelser er "miniundersøgelser", der – gerne med forskerdeltagelse – bør gennemføres af interviewere, som medvirker i egentlige undersøgelser. Hvad er opfattelserne i den øvrige metodelitteratur? Ligesom Hansen & Andersen (2000) anbefaler mange andre gennemførelse af prøveinterview. Proveundersøgelser sigter imod at afprøve, hvordan spørgsmål og spørgeskemaer fungerer "i praksis": "Once a questionnaire has been developed, each question and the questionnaire as a whole must be evaluated rigorously before final administration" (de Vaus 1998:99). Men hvad er en prøveundersøgelse? Herom er der ikke konsensus i metodelitteraturen, hvilket vil sige, at surveymetoden heller ikke inden for dette område er så standardiseret som almindeligvis antaget.

Hvad er prøveundersøgelsers fordele og ulemper? (se fx: Aldridge & Levine 2001; Fowler, i: Bickman & Rog 1998). Til fordelene hører, at de som regel giver anledning til påpege svagheder ved spørgeskemaer. Proveundersøgelser er ingen garanti for reduktion af måleproblemer, men bedre end ingenting. Der er dog også kritiske røster, idet flere påpeger prøveundersøgelsers mangler. Prøveinterview er ofte for få, usystematiske og for overfladiske, fx pga. hastværk og begrænsede økonomiske midler. Ydermere tilbyder prøveundersøgelser begrænset indsigt i respondenterne problemer i forhold til at besvare spørgsmål. Undersøgelser underbygger kritikken, da det viser sig, at prøveundersøgelser fx er uegnede til at afdekke manglende eller forskellig forståelse af ord og spørgsmål. Hertil kommer, at resultater af prøveinterview ikke kan anvendes i generaliserende øjemed.

### **Fokusgrupper**

Yderligere en procedure til sikring af surveydatas kvalitet er sammensætning af fokusgrupper og gennemførelse af fokusgruppediskussioner. Proceduren, der almindeligvis benyttes i datatilvirkningens initiale fase, anbefales i dele af metodelitteraturen og finder voksende udbredelse, fx i USA. Når der advokeres for fokusgrupper, er det bl.a. med henvisning til, at gruppemedlemmernes interaktion tilbyder forskere informationer, som ikke kan opnås på andre måder. Ved gennemførelse af fokusgruppediskussioner selekteres og mødes et begrænset antal respondenter for at lade sig interviewe om og drøfte specifikke på forhånd fastlagte og situationelt opståede problematikker. Det kan være afprøvning af spørgeskemaudkast, fx mht. respondenteres forståelse af spørgsmål og genskabelse af informationer. Fokusgruppediskussioner kan også belyse, om respondenter er villige til at afgive nøjagtige svar. Der er imidlertid ikke enighed om, hvordan den slags diskussioner gennemføres: "Perhaps one of the most intriguing aspects of working with focus groups is that, as yet, there are no definite rules for their use" (Nassar-McMillan & Borders 2002:in).

Til procedurens fordele hører, at den ofte er karakteriseret ved fleksibilitet og deltaljerede svar, at omfattende informationsmængder kan genereres på kort tid, at fremgangsmåden som regel fører til uforudsete og vigtige oplysninger, og at anvendte resultater af gruppediskussioner derfor bidrager til at forbedre spørgeskemaer (se fx: Czaja 1998). Som ulemper påpeges, at fokusgrupperesultater ikke kan overskride deltageres ytringer, dvs. ikke lader sig generalisere, og at udbyttet undertiden er mangelfuldt, fordi mange surveyforskere synes uden erfaring med gennemførelse af fokusgruppeinterview. Det nævnes også, at enkelte respondenteres gruppedomnans og af-

visning af sensitive emner er mulige dysfunktioner.

### **Kognitive laboratorieinterview**

Kognitive laboratorieinterview kan også anvendes til forebyggelse af måleproblemer. Herom findes en righoldig litteratur, der for en dels vedkommende betoner prøveundersøgelers og fokusgruppers utilstrækkelighed. Laboratorieinterview, der drager nytte af lingvistik og kognitionspsykologi:

... are increasingly seen as an essential step in the design and evaluation of a survey instrument. Questions that are not consistently understood or answered in a laboratory setting certainly will not work any better in an actual survey (Fowler 1993:98).

Hvad er kognitive laboratorieinterview, og hvordan gennemføres de? Ligesom andre spørgsmål besvares heller ikke dette konsensuelt, men interviewene anskues som verbale "hjernebesøg", hvis formål er at belyse svarprocessuelle aspekter: "... respondents are instructed to think out loud as they answer each question" (Czaja & Blair 1996:99). En anden opfattelse er Fowler's, der peger på tre fællestæk: Interviewene sigter imod at belyse spørgsmålsforståelse og andre svarprocessuelle aspekter, de foregår almindeligvis under laboratorielignende forhold, og ved at "tænke højt" bidrager respondenter til forståelse af svarprocessen (i: Bickman & Rog 1998). "Tænke-højt-interview" findes i to versioner: retrospektive vs. kontinuerlige. Når laboratorieinterview er kontinuerlige, tænkes højt, mens spørgsmål besvares, mens retrospektive interview er tænken højt *ex post*.

I interviewpraksis findes yderligere procedurale aspekter. Det drejer sig bl.a. om tænke-højt-fremmede brug af *probes*,

fx: "Hvordan forstår du ordet demokrati i det spørgsmål, jeg netop stillede?". Efter metodelitteraturen at domme gennemføres tænke-højt-interview sjældent i "ren" form, men suppleres med *probes* og andre procedurale aspekter, fx parafrasering, *confidence ratings* (svarsikkerhed), *response latency* (tidsanvendelse fra spørgsmål til svar), *memory cues* og genbesvarelse af surveyspørgsmål.

Ifølge metodelitteraturen har laboratorieinterview fordele mht. reduktion af måleproblemer. Ved gennemførelse af velforbredte "hjernebesøg" tilvirkes informationer, bl.a. om respondenters spørgsmålsforståelse og hukommelsesproblemer, som fx prøveundersøgelser ikke tilbyder (se fx: Sudman et al. 1996). Der er dog også advarende røster (se fx: Czaja 1998). Én forbeholden stemme ytrer, at laboratorieinterview adskiller sig fra gængs *common ground* respondenter og interviewere imellem, og at ikke alle respondenter er lige egnede til at *ad hoc*-verbalisere "hjernebesøg". Andre påpeger laboratorieinterviews manglende generaliseringsmuligheder. Enkelte er særdeles skeptiske over for laboratorieinterview og spørger, hvad fx amerikanske surveyforskere generaliserende har lært om spørgeskemakonstruktion.

### **Interaktionskodning**

Mens laboratorieinterview sigter imod at opnå adgang til respondenters kognitive svarprocesser, er formålet med interaktionskodning registrering af samspillet mellem respondenter og interviewere. Herom findes en righoldig metodelitteratur, hvori beskrives, hvad interaktionskodning er, hvordan interaktionskodning gennemføres og hvilke fordele og ulemper, der knytter sig til proceduren.

Da kodning af interaktioner respondenter og interviewere imellem viser, hvordan spørgeskemaudkast fungerer i praksis, er interaktionskodning et vigtigt

supplement til gængse prøveundersøgelser. Proceduren, der forudsætter båndoptagelse og kodning af prøveinterview, sigter imod at registrere antallet af "afvigelser" fra en "ideel" model:

In interaction coding, a third party uses a simple set of codes to indicate ... nonverbatim reading of the question, respondents' requests for clarification, or inadequate answers (Czaja & Blair 1996:101).

"Afvigelser" kan også være afbrydelse af spørgsmålsoplæsning, afgivelse af *ved ikke*-svar og svarnægtelse. Nogle metodeforfattere peger på, at andelen af "afvigelser" ikke bør overstige et på forhånd berammet niveau, fx 15 pct.

Navnlig metodeforfattere, der bifalder stimulus-respons-modellen, mener, at interaktionskodning er en enkel, fleksibel og billig procedure til forebyggelse af interviewereffekter (se fx: Fowler & Mangione 1990). Endvidere pointeres, at kodning er brugbar, fordi den tilføjer viden om, hvordan spørgsmål fungerer, og at kodning derfor kan yde væsentlige bidrag til forbedring af spørgeskemaer (se fx: Groves 1989). Som yderligere fordele nævnes objektivitet, gentagelighed og sammenlignelighed. Desuden kan resultater sammenlignes på tværs af surveyundersøgelser, hvorved systematiske relationer mellem spørgsmåls træk og respondenters svarproblemer kan afdækkes. Der er dog også kritiske røster, hvoraf én angår interaktionsregistrering, som aldrig er "objektiv". Et mere afgørende *aber dabei* er, at interaktionskodning normalt ikke bidrager til at belyse evt. modelafvigelsers årsager (se fx: Czaja 1998). Trods ulempen har enkelte metodeforfattere dog forsøgt at udvikle kontekstoverskridende udsagn med afsæt i tværgående analyser af interaktionskodning.

### *Split sample-eksperimenter*

Som fx påpeget af Hansen & Andersen (2000) lader målingers stabilitet sig bl.a. afprøve ved gennemførelse af *split sample*-eksperimenter, der er en hyppigt om-talt og anvendt procedure. Mens resulta-ter af prøveundersøgelser, fokusgrupper, laboratorieinterview og interaktionskod-ning ikke tilvejebringer generaliserende viden, gælder det modsatte *split sample*-eksperimenter, der tillader konklusioner om specifikke måleproblemers samlede udbredelse. *Split sample*-procedurer inde-bærer, at stikprøver dekomponeres i to eller flere segmenter, hvoraf det ene segment interviewes med afsæt i én spør-geskemaversion, mens det andet besva-rer spørgsmål tilhørende en anden ver-sion (se fx: Aldridge & Levine 2001; Gro-ves 1989). Kun forskeres fantasi sætter grænser for, hvordan der fx manipuleres med syntaks, ordvalg, transformation af spørgsmål, ændrede svarkategorier og spørgsmåls- rækkefølge. Efterfølgende kan der ved statistisk inferens drages slutninger om målingers signifikante for-skelle to eller flere versioner imellem, men almindeligvis ikke om målinger, der bør foretrækkes.

*Split sample*-eksperimenters markante fordel er, at specifikke målingers stabili-tet/konsistens eller mangel på samme in-duktivt-statistisk overskrider responden-ter besvarelser og kan generaliseres til populationsniveau. Imidlertid har proceduren også en svaghed: at det som regel ikke er muligt at pege på spørge-skemaversioner, der genererer færrest måleproblemer. Kun hvis *split sample*-re-sultater sammenholdes med andre infor-mationer, er lokalisering af det "mindst ringe" spørgsmål, de "mindst ringe" svarkategorier eller de færreste kontekst-effekter mulig. Andre informationer kan være valide registeroplysninger, fx om indkomst eller lægebesøg.

### *Andre procedurer*

Der findes yderligere procedurer, hvis iværksættelse kan bidrage til reduktion af måleproblemer. Hvad *debriefing af re-spondenter* angår, er proceduren beslægtet med kognitive laboratorieinterview, men i modsætning til laboratorieinter-view, der som betegnelsen tilsiger, gen-nemføres under laboratorielignende for-hold, foregår respondentdebriefing i "fel-tet", normalt i forlængelse af prøveinter-view. Én fordel ved debriefing af respon-denter er, at proceduren i højere grad end kognitive laboratorieinterview muliggør generaliseringer. Det skyldes, at procedu-ren er fhv. struktureret og almindeligvis gennemføres med flere respondenter end ved laboratorieinterview. Der er imidler-tid også advarende roster, hvoraf én an-går mulige efterrationaliseringer.

*Monitorering af interviewere* findes i for-skellige former, fx overvågning af inter-viewøvelser under uddannelse og inter-aktionskodning. Men overvågning af egentlige surveyinterview er også en an-befalet procedure. I forbindelse hermed nævnes, at telefoninterview er lettere at supervisere end besøgsinterview, idet dog sidstnævnte fx kan optages på bånd og "superviseres" *ex post*. Også direkte ob-servation af besøgsintervieweres adfærd er en mulighed. Ligesom debriefing af re-spondenter er én mulig kvalitetssikrende procedure, gælder tilsvarende *debriefing af interviewere*, fx efter gennemførte prøve-interview. Flere forfattere nævner eller underforstår debriefing af inter-viewere som en integreret del af "klassiske" prøve-undersøgelser. Ved besvarelse af standar-diserede *follow up*-spørgsmål evt. supple-ret med diskussioner i fokusgrupper kan interviewere fx bidrage til formulering af "gode" spørgsmål. Der er imidlertid også advarende roster. Der peges fx på, at det for mange interviewere kan være proble-matisk både at virke som interviewere og spørgeskemaevaluatorer.

Endelig kan anvendelse af *ekspertpaneler*, der ifgl. flere forfattere ikke er et substitut fx for kognitive laboratorieinterview, bidrage til reduktion af måleproblemer, fx ved at forudsige og dermed forebygge kognitive tvivlsspørgsmål. Der peges på ekspertvurderinger af spørgeskemaer som en procedure med stort udbytte og fhv. lave omkostninger: "Expert panels can detect problems not found by other techniques and have the added advantage of being relatively inexpensive" (Czaja 1998:in). Ekspertvurderinger kan være individuelle eller gennemføres som fokusgruppediskussioner, fx med afsæt i på forhånd besvarede evalueringsskemaer.

## Perspektivering

Artiklen anskueliggør, at kvantitativ metodelitteraturs divergenser er markante, fx med hensyn til anbefalinger vedr. spørgeskemakonstruktion og opfattelser af surveyinterview, at surveymetodens datakonstruerende aspekter ifølge metodelitteraturen ikke er så standardiserede som almindeligvis antaget, og at adskillige parametre genererer måleproblemer, som anfægter surveybaserede forskningsresultater.<sup>6</sup> Det er desuden anskueliggjort, at Hansen & Andersen's (2000) lærebog blot er én af mange – ofte rivaliserende – opfattelser af, hvordan surveydata bør konstrueres og kvalitetssikres, når metodeforfattere tildeles roller som "dommere". Det er velkendt, at der inden for *koalitativ* interviewforskning ikke findes en datakonstruerende og kvalitetssikrende konsensuel "køgebog" (se fx: Olsen 2002). Mere overraskende er det muligvis, at noget tilsvarende gør sig gældende inden for *kvantitativ* interviewforskning.

Kvantitativ metodelitteratur, der sætter fokus på konstruktion af spørgeskemaer som "kunst" eller "håndværk", er for en stor dels vedkommende *common*

*sense*-litteratur, der tendentielt abstraherer fra respondenters kognitive opgaveløsninger som årsag til måleproblemer (se: *oversigt* 7, 1. sektion). Mange metodeforfattere påpeger, at spørgsmål bør "forstås af" og "falde naturligt" for respondenter, men undlader at udvikle forskningsbaserede procedurer til opnåelse heraf. Det problematiseres fx ikke, om det er muligt at formulere "entydige" spørgsmål, eller om grammatisk korrekte spørgsmål med så få ord som muligt nødvendigvis er at foretrække. Mange *common sense*-forfattere peger også på, at spørgsmåls rækkefølge bør virke tematisk "naturlig", men reflekterer ikke over, at netop den slags successioner har vist sig at fremme konteksteffekter med heraf følgende måleproblemer.

Idet andre metodeforfattere afviser, at spørgeskemakonstruktion er "kunst" eller "håndværk", er deres stræben – ikke mindst CASM-bevægelsens medlemmer – at basere spørgeskemakonstruktion på analyser af respondenters opgaveløsninger med henblik på reduktion af måleproblemer (se: *oversigt* 7, 2. sektion). Med afsæt i beslægtede svarprocessuelle modeller analyseres fx, hvordan respondenter forstår spørgsmål, svarkategorier og kontekster, hvordan respondenter (re)konstruerer informationer, og hvordan (gen)skabte informationer redigeres. Selv om en svarprocessuel *grand theory* endnu ikke er udviklet, er amerikanske og andre metodeforskere indsat for at videnskabeliggøre spørgeskemakonstruktion lovende og søgt overført til Danmark (se fx: Olsen 2001). Hvis konstruktion af spørgeskemaer – med Payne's (1951) problematiske formulering – vedbliver at være *the art of asking questions*, er en væsentlig reduktion af surveymetodens markante empiriske skrøbelighed udelukket. En alternativ fordring, der adskiller sig væsentligt fra *common sense*-forfatteres anbefalinger, og som artiklens for-

fatter bifalder, er *the interdisciplinary science of asking questions*, hvilket bl.a. kræver, at surveyforskere – som fx i USA – iværksætter et samarbejde med lingvister og kognitionspsykologer.

Der er tillige behov for forskningsbaseret afklaring af, hvad surveyinterview egentlig er, og hvordan surveyinterviews socio-verbale interaktioner genererer måleproblemer (se: *oversigt 7*, 3. sektion). I så henseende er CASM-bevægelsen mindre attraktiv, fordi dens medlemmer tendentielt abstraherer fra interaktion som årsag til måleproblemer. I store dele af metodelitteraturen – det gælder navnlig *common sense*-litteraturen – opfattes surveyinterview som standardiserede med

stimulus respons-modellen som bagvedliggende ideal, der ganske vist er vanskeligt at realisere. Andre metodeforfattere tager afstand fra stimulus respons-model og standardisering som idealer og peger fx på surveyinterview som særlige former for talehandlinger, hvor *complete interviewer-uniformity* hverken er mulig eller ønskelig. I netop den sammenhæng er Houtkoop-Steenstra's (2000) afhandling et lovende bidrag til ny erkendelse om surveyinterviews komplekse beskaffenhed. Opfattelserne af, hvad surveyinterview er, er med andre ord så uensartede og rivaliserende, at de overskrider simple nuanceforskelle.

Endelig er der behov for yderligere

**Oversigt 7. Kvalitetssikrende spørgsmål vedr. konstruktion og kvalitetssikring af surveydata.**

|  |  |
|--|--|
| <p>[1]</p> <p>Spørgeskema-konstruktion</p> | <ul style="list-style-type: none"> <li><input type="checkbox"/> Er undersøgelsens problemstillinger præcise og motiverede?</li> <li><input type="checkbox"/> Er begreber defineret præcist og hensigtsmæssigt?</li> <li><input type="checkbox"/> Er begreber operationaliseret præcist og hensigtsmæssigt?</li> <li><input type="checkbox"/> Baseres spørgsmålsformulering på en motiveret spørgsmålstypologi?</li> <li><input type="checkbox"/> Er forskellige datatilvirkningsformers fordele og ulemper overvejet?</li> <li><input type="checkbox"/> Er valget af datatilvirkningsform eksplicit og motiveret?</li> <li><input type="checkbox"/> Er alle spørgsmål eksplicit relateret til problemstillinger?</li> <li><input type="checkbox"/> Er spørgsmåls grammatiske struktur overvejet?</li> <li><input type="checkbox"/> Er spørgsmål fokuserede og éndimensionelle?</li> <li><input type="checkbox"/> Er spørgsmålslængde overvejet <i>pro et contra</i>?</li> <li><input type="checkbox"/> Er spørgsmål tilstrækkeligt éntydige?</li> <li><input type="checkbox"/> Er valg af ord overvejet mht. grad af éntydighed?</li> <li><input type="checkbox"/> Er spørgsmål tilstrækkeligt neutrale og balancerede?</li> <li><input type="checkbox"/> Retter spørgsmål sig imod respondenters egne erfaringer?</li> <li><input type="checkbox"/> Er tidsmæssige afgrænsninger af spørgsmål overvejet?</li> <li><input type="checkbox"/> Er spørgsmål ikke-truende og uden bias?</li> <li><input type="checkbox"/> Er spørgsmål med realistiske antagelser om respondenter?</li> <li><input type="checkbox"/> Er formulering af åbne vs. lukkede spørgsmål overvejet?</li> <li><input type="checkbox"/> Relateres svarkategorier til og er sammenhængende med spørgsmål?</li> <li><input type="checkbox"/> Er svarkategorier udtømmende og gensidigt eksklusive?</li> <li><input type="checkbox"/> Er antallet af svarkategorier overvejet?</li> <li><input type="checkbox"/> Er svarkategorier balancerede?</li> <li><input type="checkbox"/> Er <i>ved ikke-</i> og <i>på midten</i>-svarkategorier overvejet?</li> <li><input type="checkbox"/> Er spørgsmåls rækkefølge overvejet?</li> <li><input type="checkbox"/> Er antallet af spørgsmål overvejet?</li> <li><input type="checkbox"/> Er formatering, layout mv. overvejet?</li> </ul> |
|--|--|

|   |  |
|---|--|
| <p>[2]</p> <p>Spørgeskema og respondent</p>           | <ul style="list-style-type: none"> <li><input type="checkbox"/> Baseres spørgeskemakonstruktion på en svarprocessuel model?</li> <li><input type="checkbox"/> Er ords betydningsvidde søgt reduceret?</li> <li><input type="checkbox"/> Er kvantificerende adverbier søgt ekskluderet?</li> <li><input type="checkbox"/> Er sammenhænge mellem grammatik og semantik overvejet?</li> <li><input type="checkbox"/> Er spørgsmåls betydningsvidde søgt reduceret?</li> <li><input type="checkbox"/> Er spørgsmåls længde overvejet i forhold til svarprocessen?</li> <li><input type="checkbox"/> Er anvendelse af åbne spørgsmål moderat og motiveret?</li> <li><input type="checkbox"/> Bidrager svarkategorier til reduktion af spørgsmåls betydningsvidde?</li> <li><input type="checkbox"/> Anvendes <i>ved ikke</i>-kategorier i tilknytning til holdningsspørgsmål?</li> <li><input type="checkbox"/> Er anvendelse af <i>på midten</i>-kategorier overvejet ift. svarprocessen?</li> <li><input type="checkbox"/> Er antallet af svarkategorier og deres rækkefølge overvejet?</li> <li><input type="checkbox"/> Er aktivering af spørgsmålskongruent hukommelse fremmet?</li> <li><input type="checkbox"/> Er respondenters episodiske forglemmelse søgt forebygget?</li> <li><input type="checkbox"/> Er respondenters teleskopiske informationsgenskabelse søgt forebygget?</li> <li><input type="checkbox"/> Er begivenheders potentielt karakteristiske træk overvejet?</li> <li><input type="checkbox"/> Er respondenters kognitive magelighed søgt forebygget?</li> <li><input type="checkbox"/> Er kognitive aflastninger inddraget i tilstrækkeligt omfang?</li> <li><input type="checkbox"/> Er negative konteksteffekter søgt forebygget?</li> <li><input type="checkbox"/> Er over- eller undervurderende svarredigering søgt forebygget?</li> </ul> |
| <p>[3]</p> <p>Spørgeskema og interview</p>            | <ul style="list-style-type: none"> <li><input type="checkbox"/> Baseres datatilvirkningen på en eksplicit og motiveret interviewopfattelse?</li> <li><input type="checkbox"/> Er potentielle interviewereffekter søgt forebygget?</li> <li><input type="checkbox"/> Er interviewerkendetegn overvejet i forhold til respondenterne?</li> <li><input type="checkbox"/> Muliggør spørgeskemaet gennemførelse af interview?</li> <li><input type="checkbox"/> Har interviewere tilstrækkelig uddannelse og interviewerfaring?</li> <li><input type="checkbox"/> Har interviewere spørgeskema- og interviewerinstruktionskompetence?</li> <li><input type="checkbox"/> Har interviewere sproglig kompetence og <i>probe</i>-kompetence?</li> <li><input type="checkbox"/> Har interviewere svarregistreringskompetence?</li> <li><input type="checkbox"/> Har interviewere andre nødvendige kompetencer?</li> </ul>  |
| <p>[4]</p> <p>Andre kvalitets-sikrende procedurer</p> | <ul style="list-style-type: none"> <li><input type="checkbox"/> Har kvalitetssikringen et eksplicit og motiveret konceptuelt grundlag?</li> <li><input type="checkbox"/> Har kvalitetssikringen et eksplicit og motiveret formål?</li> <li><input type="checkbox"/> Er spørgeskemaet afprøvet ved gennemførelse af prøveinterview?</li> <li><input type="checkbox"/> Er fordele og ulemper ved prøveundersøgelsen påpeget?</li> <li><input type="checkbox"/> Er der redegjort for prøveundersøgelsens bidrag til skemarevision?</li> <li><input type="checkbox"/> Er der gennemført fokusgruppediskussioner eller er de fravalgt motiveret?</li> <li><input type="checkbox"/> Er der gennemført laboratorieinterview eller er de fravalgt motiveret?</li> <li><input type="checkbox"/> Er der gennemført interaktionskodning eller er den fravalgt motiveret?</li> <li><input type="checkbox"/> Er undersøgelsen et <i>split sample</i>-eksperiment eller er det fravalgt motiveret?</li> <li><input type="checkbox"/> Er der gennemført respondentdebriefing eller er den fravalgt motiveret?</li> <li><input type="checkbox"/> Er der gennemført monitorering af interviewere eller er den fravalgt motiveret?</li> <li><input type="checkbox"/> Er der gennemført interviewerdebriefing eller er den fravalgt motiveret?</li> <li><input type="checkbox"/> Er der anvendt ekspertpanel eller er proceduren fravalgt motiveret?</li> </ul>   |

forskning om, hvordan forskellige procedurer til sikring af surveydatas kvalitet – fx prøveinterview, kognitive laboratorieinterview, *split sample*-eksperimenter og ekspertpaneler – fungerer mht. reduktion af måleproblemer (se: *oversigt 7*, 4. sektion).

For Danmarks vedkommende er der endvidere behov for implementering af procedurer, der hidtil kun undtagelsesvist er bragt i anvendelse, fx kognitive laboratorieinterview. Når Hansen & Andersen (2000) i deres lærebog næsten udelukkende sætter fokus på prøveundersøgelser, der ifølge amerikansk og anden metodeforskning i marginalt omfang fungerer som garant for reduktion af måleproblemer, er det en virkelighedstro afspejling af, at andre af de nævnte procedurer næsten aldrig anvendes ved kvalitetssikring af danske surveydata.

## Noter

1. Litteraturstudiet – Konstruktion og kvalitetssikring af surveydata – er endnu ikke offentliggjort, men er under omarbejdelse mhp. offentliggørelse som lærebog (interesserede kan rekvirere litteraturstudiet ved henvendelse til artiklens forfatter). Da arbejdspapirets "empiriske" grundlag er tekster, har den kondenserende tekstanalyse forudsat omfattende søgning, fremskaffelse og granskning af metodelitteratur. Litteraturen er fortrinsvis søgt på Danbib, Bibsys og Libris ved brug af relevante, trunkerede søgeord, fx kvantitat?, metode?, interview?, statisti? og survey?. Endvidere er gennemført supplerende søgning på Google, fx af pressemeddelelser, anmeldelser mv. Efter at litteraturen – ca. 150 bogtitler og diverse artikler – var fremskaffet af Socialforskningsinstituttets Bibliotek, blev der udarbejdet læsenoter (cirka 1.000 sider). Efterfølgende blev læsenoterne kodet og kategoriseret tematisk svarende til arbejdspapirets problemstillinger og hertil hørende subproblemstillinger.
2. Artiklens forfatter definerer måleproblem således: Når en målings troværdighed anfægtes pga. mangelfuld konstruktionsvaliditet, manglende muligheder for test af eks-

tern validitet, sandsynligt fravær af ekstern validitet og/eller pga. sandsynliggjorte reliabilitets-problemer, foreligger der et måleproblem (Olsen 2001, bd. I).

3. En CASM-inspireret model med fem komponenter er udviklet og afprøvet i en afhandling udarbejdet af artiklens forfatter (Olsen 2001, bd. I-II).

4. Læseren bedes forestille sig et skilt med opfordringen Badedragt forbudt! Som opfordringens modtager forstås afsenderens ærinde ikke umiddelbart. Er skiltet et krav om, at være klædt nydeligt på, eller skal læseren afklæde sig det sidste slør? Forst i en specifik kontekst – fx foran en kirke eller ved indgangen til en strand for nudister – beskrives opfordringens semantik. Sådan forholder det sig også, når respondenter besvarer spørgsmål.

5. Om konteksteffekter blandt danske respondenter konkluderer Olsen (2001): "Danske respondentes kontekstuelle reception og genskabelse af informationer er ... årsag til måleproblemer. Desuden er sandsynligheden for konteksteffekter størst mellem indholdsrelaterede spørgsmål i tematisk sammenhængende spørgeforløb. Endelig forekommer afsmitningsspørgsmål indimellem navnlig ved besvarelse af holdnings-spørgsmål" (Olsen 2001, bd. II:187f).

6. At der også i Danmark er behov for reduktion af måleproblemer ved konstruktion af surveydata påvises i en afhandling baseret på 32 laboratorieinterview og et split sample-eksperiment: "Hvis fremtidige danske sociologiske og andre samfundsvidenskabelige surveyundersøgelser befojet skal kunne hævdes at tilhøre det videnskabelige samfund, bør tilvirkningen af data gennemføres på måder, der dokumenterer, at måleproblemer er reduceret mest muligt" (Olsen 2001a, bd. II:200; vedr. split sample-eksperimentet, se fx: Olsen 2002a).

7. En komplet litteraturliste vedr. det endnu ikke-offentliggjorte arbejdspapir, som artiklen baseres på, kan rekvireres ved henvendelse til: ho@sfi.dk

## Litteratur<sup>7</sup>

- Aldridge, A. & Levine, K. 2001: *Surveying the Social World: Principles and Practice in Survey Research*. Philadelphia: Open University Press.
- Alreck, P.L. & Settle, R.B. 1995: *The Survey Research Handbook*. Chicago: Irwin Publications.
- Andersson, B-E. 1994: *Som man frågor får man svar*. Stockholm: Liber.
- Babbie, E.R. 1990: *Survey Research Methods*. Belmont: Wadsworth Publications.
- Bateson, N. 1984: *Data Construction in Social Surveys*. London: Harpercollins.
- Bickman, L. & Rog, D.J. (eds.) 1998: *Handbook of Applied Social Research Methods*. Thousand Oaks: Sage Publications.
- Biemer, P.P. et al. (eds.) 1991: *Measurement Errors in Surveys*. New York: John Wiley & Sons.
- Bradburn, N.M. 2000: *Questionnaire Design: from Art into Science*. Fifth International Conference on Social Science Methodology. Cologne, Germany, October 2000. National Science Foundation. ([www.norc.uchicago.edu/online/Cologne2.pdf](http://www.norc.uchicago.edu/online/Cologne2.pdf))
- Bradburn, N.M. & Sudman, S. 1979: *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass Publ.
- Burton, S. & Blair, E. 1991: "Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions. *Public Opinion Quarterly*. Vol. 55, 1991/3.
- Converse, J.M. & Presser, S. 1986: *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills: Sage Publications.
- Crano, W.D. & Brewer, M.B. 2002: *Principles and Methods of Social Research*. Lawrence Erlbaum Ass.
- Czaja, R. & Blair, J. 1996: *Designing surveys: A guide to decisions and procedures*. Thousand Oaks: Pine Forge Press.
- Czaja, R. 1998: "Questionnaire Pretesting Comes of Age". *Marketing Bulletin*, 1998/9:52-66.
- de Vaus, D.A. 1998: *Surveys in Social Research*. London: UCL Press.
- de Vaus, D.A. 2002: *Surveys in Social Research*. London: UCL Press.
- Dijkstra, N. & van der Zouwen, J. 1982: *Response Behaviour in the Survey Interview*. London: Academic Press.
- Fink, A. 1995: *How to Ask Survey Questions*. Thousand Oaks: Sage Publications.
- Fink, A. & Kosecoff, J. 1998: *How to Conduct Surveys: A Step-by-Step Guide*. Thousand Oaks: Sage Publications.
- Fowler, F. J. 1993: *Survey Research Methods*. Thousand Oaks: Sage Publications.
- Fowler, F. J. & Mangione, T. W. 1990: *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park: Sage Publications.
- Groves, R. M. 1989: *Survey Errors and Survey Costs*. New York: Wiley.
- Hansen, E. J. & Andersen, B. H. 2000: *Et sociologisk værktøj: Introduktion til den kvantitative metode*. København: Hans Reitzels Forlag.
- Hellevik, O. 1993: *Forskningsmetode i sosiologi og statsvitenskap*. Oslo: Universitetsforlaget.
- Houtkoop-Steenstra, H. 2000: *Interaction and the Standardized Survey Interview: The living questionnaire*. Cambridge: Cambridge University Press.
- Johannessen, A. & Tufte, P. A. 2002: *Introduksjon til samfunnsvitenskapelig metode*. Oslo: Abstrakt.
- Judd, C. M. et al. 1991: *Research Methods in Social Relations*. Philadelphia: Hartcourt Brace.
- Krosnick, J. A. 1991: "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys". *Applied Cognitive Psychology*. Vol. 5, 1991/2.

- Krosnick, J. A. 1999: "Survey Research". *Annual Review of Psychology*. (www.arjournals.annualreviews.org/doi/abs/10.1146/annurev.psych.50.1.537?cookieSet=1)
- Labaw, P. J. 1980: *Advanced Questionnaire Design*. Cambridge: Abt. Books.
- Lyberg, L. et al. (eds.) 1997: *Survey Measurement and Process Quality*. Chichester: Wiley.
- Mordal, T. L. 2000: *Som man spør, får man svar*. Oslo: Tano.
- Nassar-McMillan, S. C. & Borders, L. D. 2002: "Use of Focus Groups in Survey Item Development". *The Qualitative Report*, vol. 7., 2002/1. (www.nova.edu/ssss/QR/QR7-1/nassar.html)
- Nolinske, T. 1998: *Survey Research and Measurement Error*. Washington. (www.oandp.org/jpo/library/1995\_02\_068.asp)
- Olsen, H. 1998: *Talenes talende tavshed (II). Måleproblemer i surveyundersøgelser*. København: Akademisk Forlag.
- Olsen, H. 2001: *Sprogforståelse og hukommelse i danske surveyundersøgelser (bd. I-II)*. København: Socialforskningsinstituttet.
- Olsen, H. 2001a: "Hvor enig eller uenig er du?" Holdningsspørgsmål og holdningsmålinger i danske samfundsvidenskabelige surveyundersøgelser. *Nordiske Udkast* 2001/2:59-78.
- Olsen, H. 2002: *Kvalitative kvaler. Kvalitative metoder og danske kvalitative interviewundersøgelers kvalitet*. København: Akademisk Forlag.
- Olsen, H. 2002a: "The Silence of Numbers. A Split-sample Experiment concerning Respondents' Linguistic Sensitivity in Danish Survey Studies". *International Journal of Social Research Methodology*, 2002/4:293-312.
- Payne, S.L. 1951: *The Art of Asking Questions*. New Jersey: Princeton University.
- Poe, G. S. et al. 1988: "'Don't Know' Boxes in Factual Questions". *Public Opinion Quarterly*. Vol. 55:212-22.
- Rea, L. M. & Parker, R. A. 1997: *Designing and Conducting Survey Research*. San Francisco: Jossey-Bass Publications.
- Salant, P. & Dillman, D. A. 1994: *How to Conduct your own Survey*. New York: Wiley Publ.
- Sapsford, R. 1999: *Survey Research*. London: Sage Publications.
- Schuman, H. & Presser, S. 1996: *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. Thousand Oaks: Sage Publications.
- Schwarz, N. & Sudman, S. (eds.) 1992: *Context Effects in Social and Psychological Research*. New York: Jossey-Bass Publications.
- Schwarz, N. & Sudman, S. (eds.) 1996: *Answering Questions. Methodology for determining cognitive and communicative processes in survey research*. New York: Jossey-Bass Publications.
- Sirken, M. G. et al. (ed.) 1999: *Cognition and Survey Research*. New York: Wiley.
- Sudman, S. & Bradburn N. M. 1973: "Effects of Time and Memory Factors on Response in Surveys". *Journal of American Statistical Association*. Vol. 68, 1973/344.
- Sudman, S. & Bradburn, N. M. 1974: *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.
- Sudman, S. et al. 1996: *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers.
- Tanur, J. M. (ed.) 1992: *Questions about Questions*. New York: Russel Sage Foundation.
- Tourangeau, R. et al. 2000: *The Psychology of Survey Response*. Cambridge University Press.