

Essay

“I meget høj grad” – Om udviklingen og udbredelsen af tommel- fingerregler til surveys

Jonathan Hermansen

I 2004 påpegede Henning Olsen i en artikel i *Dansk Sociologi*, at litteraturen om surveys på den tid var karakteriseret ved “common sense-regler”, der ofte var enten diffuse eller gensidigt eksklusive (Olsen 2004). Der er med andre ord gået 20 år, siden denne kritik af tommelfingerregler blev fremsat. Ønsket om at opstille tommelfingerregler for surveys er dog ikke en ny for teelse, og det er budskabet om, at det vanskeligt lader sig gøre, heller ikke (Alwin 1977, 2014). Til gengæld er der siden da sket en ændring i udbredelsen og teknologien til surveys. Det er ikke længe siden, at spørgeskemaer udsendt via internettet blev behandlet som en ny og særskilt metode. Her kunne man som læser opleve at blive advaret imod at udsende spørgeskemaer elektronisk, da det kunne opfattes som useriøst (se f.eks. Lungholt & Metelmann 2011; Madsen & Grønbæk 2010). De fleste surveys udsendes efterhånden elektronisk og med begrænset brug af personlige interviewere eller postomdeling. Spørgsmålet om, hvordan spørgeskemaer konstrueres mest hensigtsmæssigt, berøres også i mange bøger og artikler, der ikke har surveys som det primære emne. Det er ofte bøger om kvantitative metoder eller anvendt statistik mere generelt. Denne type af bøger går dog sjældent i dybden med, hvordan man konkret kan modvirke målefejl som følge af formuleringen af spørgsmål og svarkategorier. Man kan derfor undre sig over, hvorfor tommelfingerregler til surveys stadig findes.

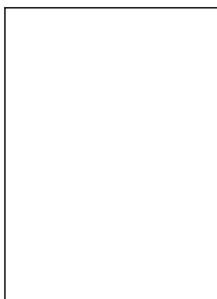
Et dominerende paradigme

Total Survey Error (herefter forkortet TSE) er fortsat det dominerende paradigme i surveyforskningen. TSE blev udviklet som paradigme i udviklingen af surveymetodologi fra 'kunst' eller 'håndværk' til en egentlig videnskab (Joye et al. 2016). Paradigmet er stadig dominerende i dansk og international forskning om surveys (Severin et al. 2020; Weisberg 2018), og det er kendetegnet ved at have en processuel forståelse af udviklingen af spørgeskemaundersøgelse (Groves & Lyberg 2010), hvorfor målefejl kan optræde fra en lang række af omstændigheder lige fra operationalisering, design, udsendelse til efterbehandling af data, hvilket skal holdes op imod de omkostninger, der ville være forbundet med at minimere risikoen for selvsamme målefejl (Weisberg 2009).

Generelt beskrives målefejl ofte i surveylitteraturen som afstanden mellem den ideelle måling og de observerede værdier (Groves et al. 2009). Denne forståelse er givetvis grunden til, at surveys bliver beskrevet som en metode traditionelt hjemmehørende i den videnskabsteoretiske realisme eller positivisme (f.eks. Boolsen & Jacobsen 2010; Thisted 2011), selvom surveys som metodisk redskab ikke i sig selv kan siges at abonnere på et særligt erkendelsesteoretisk ståsted (se også Nottelmann 2017). Det er generelt sjældent, at videnskabsteoretiske præmisser gøres til genstand for diskussion. Dog bliver surveys også undertiden kategoriseret som et eksempel på en arketypisk kvantitativ metode, der trækker på en postpositivistisk forståelse (se f.eks. Creswell 2014). Dette skyldes naturligt nok, at der kan opstilles testbare hypoteser i surveybaserede studier, samt at de anvendte analyser ofte vil bero på statistiske udregninger. Desuden tilsiger TSE-paradigmet, at der findes 'sande' værdier, der i målesammenhæng forstyrres af tilfældige fejl eller bias (se også Groves et al. 2009). Den udbredte rubricering af surveys som et kvantitativt foretagende er ikke nødvendigvis altid formålstjenlig, eftersom man risikerer at overse og drage nytte af erfaringer fra andre metodegrene med lignende problemstillinger. I tråd med dette har nogle advokeret for, at man ser bredt på forskningsstilarter frem for enkeltmetoder (se f.eks. Brewer & Hunter 2006; Frederiksen 2013).

I en dansk sammenhæng er de epistemologiske uenigheder om, hvad der kendetegner surveys, kommet til udtryk i diskussioner om, hvorvidt ideallet om ensartet spørgestimuli grundlæggende er uhensigtsmæssigt (Olsen 2005a), eller det på den anden side er en umulighed at tilpasse sig den enkelte respondents sproglige og sociale univers i surveys (Hansen & Andersen 2009). TSE som paradigme anviser ikke specifikt til bestemte måder at foretage dataindsamling på, men kan bruges til at anskueliggøre eventuelle afvejninger mellem hensynet til datakvalitet over for omkostninger. De manglende anvisninger kan være årsag til, at man i surveylitteraturen sjældent definerer, hvori forskellen består mellem mere eller mindre standardiserede interviewformer

Jonathan Hermansen
Lektor, ph.d.
Forvaltning,
Organisation og Skat
Ledelse, Organisation og
Forvaltning
Videreuddannelse
johe@kp.dk



(se også Olson et al. 2020). Derfor er der også blevet rejst en kritik af, at den lineære og processuelle forståelse af surveys, som TSE repræsenterer, ikke tager større hensyn til de kontekstuelle forhold (se også Smith 2011).

De mange udbredte tommelfingerregler

TSE-paradigmets processuelle forståelse af surveys lægger op til, at man anskuer udarbejdelsen af spørgeskemaer som en overvejende standardiseret metode. På dette punkt adskiller spørgeskemaer sig fra de fleste andre gængse metoder i samfundsvidenskaberne. Selvom tommelfingerregler i sig selv natur bør have en almen karakter for at kunne være anvendelige, så er der en betydelig variation i antallet og detaljegraden af de tommelfingerregler. Der er dog en række anbefalinger, der tilsyneladende går igen. Tabellen sammenfatter en række af de mest udbredte tommelfingerregler omhandlende henholdsvis spørgsmålsformuleringer og svarkategorier.

Udbredte tommelfingerregler	
Udbredte tommelfingerregler om spørgsmålsformulering	Udbredte tommelfingerregler om svarkategorier
Undgå tvetydige spørgsmål	Svarkategorier skal være balancerede
Undgå ledende spørgsmål	Svarkategorier skal være udtømmende
Undgå abstrakte spørgsmål	Overvej brugen af midterkategori
Undgå lange spørgsmål	Overvej brugen af 'ved ikke' -svarmulighed
Undgå hukommelseskrævende spørgsmål	Gensidigt udelukkende – hvis kun ét svar er tilladt

(Note: egen sammenfatning af tommelfingerregler på baggrund af (Ahrenkiel 2020; Albrechtsen & Jakobsen 2018; Andersen 2013; Boolsen 2008; Clement 2017; K. M. Hansen 2012; Johansen & Friis 2013; Larsen 2010; Launsø et al. 2017; Lungholt & Metelmann 2011; B. S. Madsen 2021; M. L. Madsen & Grønbæk 2010; P. Nielsen 2021; Rosenmeier & Jeppesen 2008))

Overordnet set minder anbefalingerne en del om de tilsvarende fra engelsksproget surveylitteratur (f.eks. Krosnick & Presser 2010), om end der sjældent henvises til anden eller international litteratur ved fremsætningen af de gode råd. Ofte lader der til at være enighed om, at svarmulighederne skal være balancerede (ved bipolære svarskalaer) og udtømmende (ved lukkede spørgsmål). Til gengæld kan man støde på argumenter både for og imod brugen af 'ved ikke'-svarmuligheder og brugen af neutrale midterkategorier til holdningsspørgsmål.

Overordnet betragtet fremstår alle de mest udbredte tommelfingerregler som intuitivt fornuftige råd. Det er nok de færreste, der vil mene, at abstrakte eller tvetydige spørgsmål er at foretrække. Man kan dog muligvis som læser blive efterladt med følelsen af, at rådene er lettere sagt end gjort, når det eksempelvis opsummeres med en opfordring til at gøre "det simpelt" (Madsen 2021:84).

Hvorvidt et spørgsmål er abstrakt eller letforståeligt, afhænger i høj grad af sammenhængen og målgruppen for den pågældende survey, hvilket udfordrer selve præmissen for at opsætte tommelfingerregler. I lighed kan det være svært at fastslå, hvorvidt et spørgsmål er for langt, om end 20 ord bliver opstillet som et bud (Ahrenkiel 2020), og man kan finde opfordringer til at begrænse spørgeskemaets længde til eksempelvis 20 (Albrechtsen & Jakobsen 2018) eller 30 spørgsmål (Andersen 2013), hvilket nok kan være fornuftigt, men der angives sjældent noget belæg for denne type af konkrete anbefalinger om antallet. Tilsvarende er det indlysende, at det kan være problematisk at stille spørgsmål, der stiller urealistiske krav til respondentens hukommelse. Det er en problematik, der længe har været kendt i surveylitteraturen (se f.eks. Sudman & Bradburn 1973).

Begrundelserne for opstillingen af tommelfingerregler er forskellige, men de beskrives eksempelvis som tjek, man bør foretage (Johansen & Friis 2013) eller "what not to do" (Boolsen 2008). Det er påfaldende, at de gode råd i store træk er de samme, uagtet om de er fremsat ved årtusindskiftet eller i løbet af de seneste år. Man kunne forestille sig, at der var behov for at ændre anbefalingerne, set i lyset af at der er sket en markant ændring, idet langt flere surveys i dag bliver indsamlet online og besvaret ved hjælp af mobiltelefoner (se også Couper 2013). Dette kan skyldes, at der i forskningen ikke er identificeret nogen nævneværdig betydning af overgangen til nye teknologier for respondenternes besvarelser endnu (Clement & Shamshiri-Petersen 2015; Toninelli & Revilla 2016; Tourangeau et al. 2017), men det er vanskeligt at vide, når der sjældent henvises til forskningsresultater til at underbygge de gode råd.

Alternativer til TSE og de gængse tommelfingerregler

I tillæg til tommelfingerreglerne findes en række evalueringsværktøjer, der forsøger at sætte kvaliteten af spørgeskemaspørgsmål på formel. De kan bruges til at undersøge læsevenligheden af spørgsmål, eller hvorvidt begreber bliver forstået på ensartet måde på tværs af målgrupper eller sprog (se f.eks. Holesch 2021). Nogle onlineværktøjer til at evaluere spørgeskemaspørgsmål kan bruges inden udsendelsen af surveys til at forudsige kvaliteten af spørgsmålene gennem metaanalyser på baggrund af eksperimenter (såsom Survey Quality Predictor – se også Saris & Gallhofer 2007) eller ved at trække på en database af ord til at identificere komplicerede formuleringer (såsom QUAID – se også Graesser et al. 2006). I tillæg hertil kunne man anføre, at udviklingen i sprogmodeller kan gøre dem til en udbredt assistent i udarbejdelsen af surveys.

Dette kunne i princippet reducere behovet for tommelfingerregler. Det er dog en udfordring, at de forskellige programmer og værktøjer identificerer forskellige problemer ved de samme spørgsmål (Stenger et al. 2022), hvilket betyder, at man i så fald bør benytte sig af flere forskellige og supplerende værktøjer. Dette vil dog i mange tilfælde ikke være praktisk muligt. Et nyere forskningsfelt, der potentielt kunne reducere målefejl, er brugen af paradata til at undersøge datakvaliteten efter indsamlingen (McClain et al. 2019). Efter som de fleste surveys i dag bliver indsamlet online, vil man ofte kunne finde oplysninger om eksempelvis respondentens tidsforbrug, afbrydelser eller redigering i svarafgivelse. Antallet af studier med og om paradata i surveys er dog fortsat relativt begrænset.

En anden tradition, der fokuserer på de kognitive aspekter, bliver ofte forkortet som CASM (fra eng. Cognitive Aspects of Survey Methodology), og den er blevet beskrevet som et paradigmeskifte, der fra 1980'erne ændrede fokus fra effekterne af målefejl til årsagen af målefejl (Tourangeau 2003). Om end den grundlæggende ikke ændrede ved den processuelle forståelse af surveys udvikling og den traditionelle optagethed af at undgå målefejl, der kendetegner TSE. Ikke desto mindre har studier i kognitive aspekter af spørgeskemabesvarelse bidraget til en større forståelse af respondenters hukommelse og sprogforståelse (Schwarz 2007). Eksempelvis påpegnings i det problematiske ved at anvende negationer eller sågar dobbelt negationer i spørgsmålsformuleringer, da det er mere kognitivt belastende for respondenter (se f.eks. Baumgartner et al. 2018). CASM er også blevet beskrevet som en brobygning mellem statistisk orienteret surveyforskning, lingvistik og psykologi, der har bevirket, at surveyforskningen er mere blevet tværvidenskabelig (Olsen 2005b). Således er gennemførelse og analyser af kognitive interviews som pilottest af spørgeskemaer efterhånden blevet en selvstændig disciplin (se også Willis 2015).

Den fortsatte tomten for surveys

Tommelfingerreglerne er ofte formuleret i overskriftsform, hvilket tilsiger, at man skal have en vis ekspertise for at vide, hvordan de konkrete anbefalinger skal bringes i anvendelse eller oversættes til et konkret survey. Man kunne derfor lettere polemisk indvende, at tommelfingerregler burde være forbedt eksperter, men hvilken gavn har de så?

Det kan virke omsonst at forholde sig til tommelfingerreglers troværdighed, når selv dem, der kritiserer det manglende forskningsmæssige belæg, erkender, at en del tommelfingerregler rejser flere problemer, end de løser (se også Olsen 2006:28). Det kan derfor være gavnligt at huske på, at Schuman og Presser fremhævede, at det i studier om surveys sjældent gav mening at skelne mellem konklusioner og hypoteser, da konklusioner i vid udstrækning blot er hypoteser for fremtidig forskning (Schuman & Presser 1996:297). Det samme gælder formentlig de udbredte tommelfingerregler. De kan med tiden blive afvist, men indtil da har de vist sig brugbare.

En perspektivrig distinktion, der dog endnu fylder relativt lidt, er spørgsmålet om, hvilke antagelser der ligger til grund for idéen om tommelfingerregler til surveys. Willis (2020) skelner mellem hhv. et absolutistisk og et relativistisk udgangspunkt til spørgeskemadesign. Hvor førstnævnte synspunkt er, at man meningsfuldt kan skelne mellem grundlæggende gode og dårlige spørgeskemaspørgsmål, og hvor de kvalitetsmæssige egenskaber går på tværs af tid og kontekst. Det er denne tankegang, som særligt de standardiserede evalueringsværktøjer afspejler. Heroverfor kan man anlægge en relativistisk forståelse af spørgeskemadesign, hvor kvaliteten af et spørgeskemaspørgsmål og dets medfølgende svarkategorier altid skal ses i lyset af undersøgelsens formål og sociokulturelle kontekst. Dette standpunkt finder man repræsenteret i dele af CASM-forskningen – og i dansk sammenhæng i særdeleshed fremført af Henning Olsen.

En rigid udlægning og anvendelse af TSE kunne tilsige, at dette paradigme lægger op til et absolutistisk syn på datakvalitet. Som det dog også påpeges, så kan TSE også forstås som en ramme, der tillader og inkluderer en flertydighed i forståelsen af validitet i survey sammenhæng (Sturgis et al. 2020). Det er således i sidste ende tolkningen og anvendelsen af TSE-paradigmet, der er afgørende for, hvordan man kan evaluere datakvaliteten i en survey. Hvad angår tommelfingerregler, er det dog rimeligt at hævde, at de altovervejende anlægger et absolutistisk synspunkt på udformningen af spørgsmålsformuleringer og svarkategorier.

Olsen skriver afslutningsvis i sin bog om holdningsbegreber og holdningsmålinger, hvilket spørgeskemaer ofte har til formål at undersøge, at surveys sjældent er uproblematisk, hvorfor det er nødvendigt at afprøve dem; i hans optik helst ved hjælp af laboratorieinterview (Olsen 2008:161). Dette råd er ikke overraskende med tanke på det relativistiske udgangspunkt, men her er

det måske værd at tage hensyn til den enkelte surveys formål og de ressourcer, der er til rådighed. Videnskabelige surveys bør forventeligt gøre sig den ulempe at afprøve og kvalitetssikre spørgeskemaet. Ydermere kan man, som Willis (2020) fremfører, foretage pilottest med såvel et absolutistisk som et relativistisk udgangspunkt.

Tommelfingerreglerne er næppe universelt gyldige, men det forhold, at de stadig findes og i vid udstrækning er uforandrede, indikerer, at de har en værdi. Et relevant argument imod tommelfingerregler er, at de ikke tager hensyn til konteksten. Der er i de seneste år kommet yderligere fokus på, hvordan forskellige aspekter af surveys kan tilpasses, så de i højere grad tager hensyn til direkte konkrete omstændigheder (Lynn 2017). I litteraturen og i praksis lader det dog til, at den absolutistiske, standardiserede og lineære forståelse af surveys fortsat er den dominerende. Derfor er der grund til at tro, at de mest udbredte tommelfingerregler også fremover vil forblive de samme.

Referencer

- Ahrenkiel, L. (2020). *Spørgeskema*. 2. udgave. Hentet den 26. april 2022 på: www.laeremiddel.dk/viden-og-vaerktoejer/videnskabsteori/videnskabelige-metoder/spoergeskema/
- Albrechtsen, C., & Jakobsen, I. C. R. (2018). *Spørgeskemaer der virker* (C. Albrechtsen & I. C. R. Jakobsen, Eds.; 1.). Frydenlund.
- Alwin, D. (1977). Making errors in surveys: An overview. *Sociological Methods & Research*, 6(2), 131–150.
- Alwin, D. (2014). Investigating response errors in survey data. *Sociological Methods & Research*, 43(1), 3–14.
- Andersen, I. (2013). *Den skinbarlige virkelighed. Vidensproduktion i samfundsvidenskaberne* (I. Andersen, Ed.; 5.). Samfundslitteratur.
- Baumgartner, H., Weijters, B., & Pieters, R. (2018). Misresponse to survey questions: A conceptual framework and empirical test of the effects of reversals, negations, and polar opposite core concepts. *Journal of Marketing Research*, 55(6), 869–883.
- Boolsen, M. W. (2008). *Spørgeskemaundersøgelser. Fra konstruktion af spørgsmål til analyse af svarene*. Hans Reitzels Forlag.
- Boolsen, M. W., & Jacobsen, M. H. (2010). Positivism. In M. H. Jacobsen, K. Lippert-Rasmussen, & P. Nedergaard (Eds.), *Videnskabsteori i statskundskab, sociologi og forvaltning* (1.). Hans Reitzels Forlag.
- Brewer, J., & Hunter, A. (2006). *Foundations of multimethod research: Synthesizing styles*. Sage.
- Clement, S. L. (2017). Formulering af surveyspørgsmål. In M. Frederiksen, P. Gundelach, & R. S. Nielsen (Eds.), *Survey. Design, stikprøve, spørgeskema, analyse* (1st ed., pp. 115–136). Hans Reitzels Forlag.
- Clement, S. L., & Shamshiri-Petersen, D. (2015). Web versus papir: En sammenligning af kvalitet i surveydata. *Metode & Forskningsdesign*, 2(2).

- Couper, M. P. (2013). Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys. *Survey Research Methods*, 7(3), 145–156.
- Creswell, J. W. (2014). *Research Design. Qualitative, Quantitative, and Mixed Methods Approaches*. (4.). SAGE.
- Frederiksen, M. (2013). Integration i mixed methods' forskning: Metode eller design? *Metode & Forskningsdesign*, 1(1).
- Graesser, A. C., Cai, Z., Louwerse, M. M., & Daniel, F. (2006). Question Understanding Aid (QUAID): A web facility that tests question comprehensibility. *Public Opinion Quarterly*, 70, 3–22.
- Groves, R., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (Second). Wiley.
- Hansen, E. J., & Andersen, B. H. (2009). *Et sociologik værktøj. Introduktion til den kvantitative metode*. (2nd ed., Vol. 1). Hans Reitzels Forlag.
- Hansen, K. M. (2012). Spørgeskemadesign. In L. B. Andersen, K. M. Hansen, & R. Klemmensen (Eds.), *Metoder i statskundskab* (2., pp. 302–323). Hans Reitzels Forlag.
- Holesch, A. (2021). Measuring solidarity: towards a survey question on fiscal solidarity in the European Union. *Acta Politica*, 56(2), 376–394. <https://doi.org/10.1057/S41269-021-00198-4>
- Johansen, K. R., & Friis, L. (2013). *En ny undersøgelse viser - eller gør den?* (Second). Forlaget Ajour.
- Joye, D., Wolf, C., Smith, T., & Fu, Y. (2016). Survey methodology: challenges and principles. In D. Joye, C. Wolf, T. Smith, & Y. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 3–15). SAGE Publications.
- Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In P. v. Marsden & J. D. Wright (Eds.), *Handbook of Survey Research* (2., pp. 263–312). Emerald Publishing.
- Larsen, A. K. (2010). *En enklere metode. Vejledning i samfundsvidenskabelig forskningsmetode*. Akademisk Forlag.
- Launsø, L., Rieper, O., & Olsen, L. (2017). *Forskning om og med mennesker : Vol. 2.K* (7.). Munksgaard.
- Lungholt, H., & Metelmann, P. (2011). *Hvem spiser boller i karry? En introduktion til samfundsvidenskabelig metode i teori og praksis* (First). COK - Center for Offentlig Kompetenceudvikling.
- Lynn, P. (2017). From standardised to targeted survey procedures for tackling non-response and attrition. *Survey Research Methods*, 11(1), 93–103.
- Madsen, B. S. (2021). *Statistik for ikke-statistikere* (3.). Samfundslitteratur.
- Madsen, M. L., & Grønbæk, M. (2010). E-surveys. In L. Fuglsang, P. Hagedorn-Rasmussen, & P. B. Olsen (Eds.), *Teknikker i samfundsvidenskaberne: Vol. 3*. (1., pp. 83–102). Roskilde Universitetsforlag.
- McClain, C. A., Couper, M. P., Hupp, A. L., Keusch, F., Peterson, G., Piskorowski, A. D., & West, B. T. (2019). A typology of web survey paradata for assessing total survey error. *Social Science Computer Review*, 37(2), 196–213. <https://doi.org/10.1177/0894439318759670>
- Nielsen, P. (2021). *Produktion af viden. En praktisk guide til samfundsvidenskabelig metode.*: Vol. 4. (1.). Praxis.

- Nottelmann, N. (2017). Om positivisme og objektivisme i samfundsvidenskaberne. *Dansk Sociologi*, 28(3), 9–39.
- Olsen, H. (2004). De plukkes ikke som blomster på en mark: Tendenser i kvantitativ metodelitteratur vedr. konstruktion og kvalitetssikring af surveydata. *Dansk Sociologi*, 15(1), 65–95.
- Olsen, H. (2005a). *Fra spørgsmål til svar. Konstruktion og kvalitetssikring af spørgeskemadata*. Akademisk Forlag.
- Olsen, H. (2005b). Tværvidenskabelig spørgeskemakonstruktion. *Sociologisk Forskning*, 42(1), 45–70.
- Olsen, H. (2006). *Guide til gode spørgeskemaer : Vol. 06:11*. Socialforskningsinstituttet.
- Olsen, H. (2008). *Holder holdninger? Holdningsbegreber og holdningsmålinger*. Akademisk Forlag.
- Olson, K., Smyth, J. D., Dykema, J., Holbrook, A. L., Kreuter, F., & West, B. T. (2020). The past, present, and future of research on interviewer effects. In *Interviewer effects from a total survey error perspective* (pp. 3–16). Chapman and Hall/CRC.
- Rosenmeier, S. L., & Jeppesen, D. S. (2008). Spørgeskemaets spørgsmål. In N.-H. M. Hansen, B. Marckmann, & E. Nørregård-Nielsen (Eds.), *Spørgeskemaer i virkeligheden* (1st ed., pp. 71–93). Forlaget Samfundslitteratur.
- Saris, W. E., & Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. John Wiley & Sons.
- Schuman, H., & Presser, S. (1996). *Questions & Answers in Attitude Surveys*. Sage Publications.
- Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 21(2), 277–287.
- Severin, M. C., Clement, S. L., & Shamshiri-Petersen, D. (2020). Device-effekter i web surveys: Har variationer i brugen af PC, smartphone og tablet betydning for surveykvaliteten? *Metode & Forskningsdesign*, 3(3), 56–82.
- Smith, T. W. (2011). Refining the total survey error perspective. *International Journal of Public Opinion Research*, 23(4), 464–484. <https://academic.oup.com/ijpor/article-abstract/23/4/464/707060>
- Stenger, R., Olson, K., & Smyth, J. D. (2022). Comparing Readability Measures and Computer-assisted Question Evaluation Tools for Self-administered Survey Questions. *Field Methods*, 1–16. <https://doi.org/10.1177/1525822X221124469>
- Sturgis, P., Brunton-Smith, I., & Jackson, J. (2020). Regression-Based Response Probing for Assessing the Validity of Survey Questions. In P. C. Beatty, D. Collins, L. Kaye, J. Padilla, G. B. Willis, & A. Wilmot (Eds.), *Advances in Questionnaire Design, Development, Evaluation and Testing* (1st ed., pp. 573–612). Wiley.
- Sudman, S., & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 68(344), 805–815.
- Thisted, J. (2011). *Forskningsmetoder i praksis: Vol. 2. (1.)*. Munksgaard Danmark.
- Toninelli, D., & Revilla, M. (2016). Smartphones vs PCs: Does the Device Affect the Web Survey Experience and the Measurement Error for Sensitive Topics? - A Replication of the Mavletova & Couper's 2013 Experiment. *Survey Research Methods*, 10(2), 153–169.

- Tourangeau, R. (2003). Cognitive aspects of survey measurement and mismeasurement. *International Journal of Public Opinion Research*, 15(1), 3–7.
- Tourangeau, R., Maitland, A., Rivero, G., Sun, H., Williams, D., & Yan, T. (2017). Web Surveys by Smartphone and Tablets: Effects on Survey Responses. *Public Opinion Quarterly*, 81(4), 896–929.
- Weisberg, H. (2009). *The total survey error approach: A guide to the new science of survey research*. The University of Chicago Press.
- Weisberg, H. F. (2018). Total Survey Error. In L. R. Atkeson & R. M. Alvarez (Eds.), *The Oxford Handbook of Polling and Survey Methods* (pp. 13–27). Oxford University Press.
- Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design* (1.). Oxford University Press.
- Willis, G. B. (2020). Questionnaire Design, Development, Evaluation, and Testing: Where Are We, and Where Are We Headed? In P. C. Beatty, D. Collins, L. Kaye, J. Padilla, G. B. Willis, & A. Wilmot (Eds.), *Advances in Questionnaire Design, Development, Evaluation, and Testing* (1., pp. 1–23). Wiley.