

2020

Device-effekter i web surveys

Har variationer i brugen af PC, smartphone og tablet betydning for surveykvaliteten?

Majbritt C. Severin

Ekstern lektor, Aalborg Universitet

Sanne L. Clement

Lektor, Aalborg Universitet

Ditte Shamshiri-Petersen

Lektor, Aalborg Universitet

Resumé:

Tidene, hvor surveys blev besvaret enten ved hjælp af en interviewer eller på papir hjemme ved spisebordet, er under forandring. Fremvæksten af nye webbaserede indsamlingsmetoder og besvarelses-devices betyder, at respondenter i dag kan udfylde surveyen på mange forskellige måder og i mange forskellige kontekster. Hvor det muligvis er positivt for svarprocenten, rejser det imidlertid spørgsmålet om, hvorvidt det er befordrende for respondenternes engagement i besvarelsen og i sidste ende påvirker kvaliteten af data negativt. På baggrund af en repræsentativ befolkningsurvey finder denne artikel dog, at der ikke er forskel på web survey-besvarelser på henholdsvis PC, tablet og smartphone. Når der tages højde for, at nogle grupper foretrækker bestemte devices frem for andre, såkaldte selektionseffekter, så er der ingen forskel på, hvor meget respondenterne engagerer sig, hverken når de selv skal vurdere det, eller når vi måler det som tiden brugt på besvarelsen. Derfor er det anbefalingen til afsendere af web surveys, at de ikke skal afstå fra at udbyde web surveys på forskellige devices. Tværtimod tyder det på, at fordelene ved at tillade mobile besvarelser overstiger ulemperne, hvis blot der tages højde for, at web survey-layoutet er optimeret til det.

1: Introduktion

Betydningen af indsamlingsmetode for både svarprocenter, bias i bortfald samt for kvaliteten af de indsamlede data har været genstand for stigende opmærksomhed blandt surveyforskere de seneste årtier. Særligt fokus har der været på webbaseret dataindsamling, som ofte er både hurtigere og billigere end andre indsamlingsmetoder (Roberts, 2007; de Leeuw, 2018; Couper, 2000; Dillman, 2000), og ofte sammenlignes web, der er en selvadministreret indsamlingsmetode, med interviewer-administrerede indsamlingsmetoder som telefon- og besøgsinterviewet (for en oversigt, se de Leeuw, 2018).

Imidlertid adskiller web sig fra andre modes på anden vis end ved fraværet af en interviewer, eftersom det her endvidere ikke er muligt at bestemme besvarelses-mediet, altså typen af device. Vi har som afsender af en web survey således ikke kontrol over, om respondenter besvarer spørgeskemaet på en computer, på en tablet eller på en smartphone (medmindre der anvendes surveyprogrammer, der kan udelukke bestemte besvarelses-devices, se eksempelvis Peterson et al., 2017: 206). Baggrunden for dette er en nyere mediemæssig udvikling, hvor der – samtidig med at stadig flere medier udvikles – sker en øget sammensmeltning af eksisterende medier. Inden for medieforskningen begrebsættes denne udvikling ofte som en øget 'mediekonvergens' og er konkret kendetegnet ved, at medieindhold som f.eks. en web survey, der tidligere var knyttet til et bestemt device, nu kan tilgås fra flere forskellige devices (Jenkins, 2006: 2-3; Helles, 2016: 58-60).

For respondenter i web surveys betyder denne udvikling en øget tilgængelighed og fleksibilitet i besvarelsesmetoderne, som kan have positive effekter på svarprocenten. På andre parametre kan udvikling dog også være problematisk og skabe nye udfordringer i forhold til at sikre høj datakvalitet (Couper, Antoun & Mavletova, 2017: 133-134; Antoun, 2015: 100). Fra medieforskningen ved vi, at der er forskel på, hvordan brugerne benytter forskellige devices (se eksempelvis Deng et al. 2019; Helles 2016), og det er således oplagt at antage, at disse forskelligartede brugsmønstre kan have betydning for kvaliteten i besvarelser i web surveys.

Sammenholdt med, at flere og flere respondenter i disse år anvender smartphonen som besvarelses-device i web surveys, finder vi det relevant at undersøge, om besvarelses-device bør være et fremtidigt fokuspunkt i forhold til at sikre høj datakvalitet i web surveys. I dette studie er formålet, via sammenligninger af besvarelser på PC, smartphone og tablet fra det danske ISSP-modul fra 2018:

1. At undersøge, om der er forskelle i datakvaliteten for besvarelser på forskellige besvarelses-devices.
2. At undersøge, om eventuelle forskelle kan tilskrives selektionseffekter eller er deviceeffekter.

3. På baggrund af resultaterne opstille anbefalinger til, hvad afsendere af web surveys kan gøre for at minimere eventuelle nye biases knyttet til besvarelses-devicet.

Studiet supplerer den eksisterende forskningsbaserede viden, der typisk er baseret på web-paneler, populations-subsamples som f.eks. studerende samt korte pseudo-spørgeskemaer (se f.eks. Keusch & Yan, 2017; Schlosser & Mays, 2017; Toepoel & Lugtig 2014), ved at efterteste internationale undersøgelser om device-effekter i et repræsentativt befolknings-sample og på en større videnskabelig survey. Da dette er mere krævende og forudsætter en større kognitiv indsats af respondenten, kan studiet betragtes som en kritisk case, hvor logikken er: ”Hvis det ikke har gyldighed i dette tilfælde, gælder det i ingen (eller kun nogle få) tilfælde” (Flyvbjerg, 2015: 508). Studiet bidrager også til den eksisterende forskningsmæssige viden, der fortrinsvis er baseret på eksperimenter, ved at generere viden om device-effekter, når respondenten selv vælger og dermed formodes at bruge et device, som vedkommende er fortrolig med. Endelig bidrager studiet, gennem sin innovative kobling af surveymetode-litteratur med indsigter fra medievidenskaben, til udviklingen af en mere nuanceret forståelse af mulige device-effekter.

Datakvalitet kan defineres på forskellige måder, men bliver i denne artikel defineret med udgangspunkt i Total Survey Error (TSE), der måler kvalitet i surveyundersøgelser ud fra to overordnede parametre: repræsentation og måling. På repræsentationssiden måles kvalitet ud fra, hvor repræsentativ stikprøven er for studiets målpopulation, samt hvor komplette data er. På målingssiden opgøres kvalitet ud fra, hvor godt data indfanger det underliggende teoretiske begreb, eller empiriske fænomen, vi ønsker at afdække, og her er respondentens grundighed i besvarelsen af stor betydning (se eksempelvis Groves et al., 2009; Frederiksen, Gundelach & Nielsen, 2017). En mere præcis gennemgang og operationalisering af surveykvalitet vil blive foretaget i artiklens metodesektion.

2: State-of-the-art: hvad ved vi om device-effekter?

Gennem 2010'erne er et nyt forskningsfelt vokset frem inden for forskningen i surveymetode, nemlig forskningen i device-effekter. Denne litteratur beskæftiger sig med surveyrespondenters brug af forskellige medier, eller devices, til at besvare web surveys, og hvorvidt brugen af mobile devices som for eksempel en smartphone og tablet påvirker datakvaliteten af besvarelserne negativt. I medieforskningen fremhæves ofte en række centrale forskelle i brugsmønstrene på PC, tablet og smart-phone, som kan være af betydning, når respondenten skal besvare en web survey. Disse forskelle, og deres mulige betydning for surveykvaliteten, præsenteres i det følgende afsnit som afsæt for gennemgangen af state-of-the-art om device-effekter.

2.1. BRUGSFORSKELLE PÅ PC, TABLET OG SMARTPHONE

Der er en række væsentlige forskelle på brugsmønstrene på PC, tablet og smartphone, som både kan tilskrives forskellige brugskontekster samt variationer i tekniske features. Forskellene er lidt stereotypisk sammenfattet i Tabel 1.

Tabel 1. Brugskontekst og brugsmønstre på PC, tablet og smartphone

	<i>PC</i>	<i>Tablet</i>	<i>Smartphone</i>
<i>Brugskontekst</i>	Hjemme eller på arbejdet	Variierende	”På farten”
<i>Brugssessioner</i>	Lange, sammenhængende	Variierende	Korte, fragmenterede
<i>Motivation</i>	Fordybelse	Variierende	Mobilitet og aktualitet
<i>Skærmstørrelse</i>	Stor	Mellem	Lille
<i>Navigation og præcision</i>	Mus og tastatur, stor præcision	Touchscreen, lille præcision	Touchscreen, lille præcision
<i>Surveylayout</i>	Standard	Tilpasset	Tilpasset

Kilde: Egen fremstilling på baggrund af Deng et al. (2019), Klastrup (2013), Helles (2016), Couper, Antoun & Mavletova (2017), Antoun, Katz, Argueta & Wang (2018) og Wells, Bailey & Link (2013).

Hvad angår brugskonteksten, så har litteraturen, som det også fremgår af Tabel 1, peget på, at hvor enen PC ofte benyttes i hjemmet eller på arbejdet, så bruges smartphonen ofte på offentlige steder eller på anden vis ”på farten”, hvor der kan være mange distraktioner. Dette øger risikoen for bortfald undervejs i besvarelsen og kan ligeledes reducere grundigheden i respondentens svar på spørgsmålene. Brugskonteksten for tabletten er mere variierende men anvendes som oftest i hjemmet eller på arbejdet som PC'en.

I forlængelse af disse forskelle i brugskonteksten peger litteraturen også på, at brugssessionerne og motivationen for at anvende devicene er forskellige. I kraft af at PC'en ofte anvendes i rolige omgivelser, er længden af brugssessionerne typisk lange og sammenhængende, og motivationen for at bruge mediet er ofte en form for fordybelse. På smartphonen, hvis anvendelse typisk er i mere urolige omgivelser, er brugssessionerne typisk kortere og mere fragmenterede, og motivationen er i højere grad hurtigt at blive opdateret, mens man er ”på farten”. På baggrund af tablettens variierende brugskontekster er længden af brugssessionerne samt motivationen for at benytte deviceet ligeledes divergerende men minder formentlig også her mere om PC'en end smartphonen. Disse forskelle kan have betydning for den tilgang, og ikke mindst de forventninger, som respondenter har til, hvad der kræves og dermed grundigheden af besvarelser i web surveys.

Endelig har litteraturen, som det også fremgår af tabellen, peget på, at tekniske features som skærmstørrelse og præcisionen af de integrerede navigationssystemer ligeledes er en betydningsfuld forskel på de tre devices. Smartphonen har som oftest en lille skærm og anvender touchscreen, der kan gøre det vanskeligt at navigere og afgive præcise svar på spørgsmål i web surveys samt danne sig overblik over eksempelvis forskellige svarmuligheder. Tabletten anvender ligeledes touchscreen men har til gengæld en større skærm, der kan gøre det lettere at danne et overblik i besvarelsen. PC'en er typisk det device, der har den største skærm, og anvender som regel mus og tastatur, der gør navigationen og svarafgivelsen mere præcis. Navigationen på smartphone og tablet kan yderligere besværliggøres, såfremt surveylayoutet, der ofte som standard er designet til PC'en, ikke er optimeret til mobile devices.

Forskellene i besvarelses-devicets brugskontekst og tekniske features underbygger antagelsen om devicets betydning for datakvaliteten i web surveys. Spørgsmålet er, om det også reelt er tilfældet.

2.2. BESVARELSES-DEVICETS BETYDNING FOR DATAKVALITETEN I WEB SURVEYS

Forskningen i device-effekter er endnu i sin spæde start, med de fleste udgivelser inden for de allerseneste år (se f.eks. Couper, Antoun & Mavletova, 2017; Antoun, Couper & Conrad 2017; Keusch & Yan 2017; Schlosser & Mays 2018; Tourangeau et al. 2018). Dette til trods, er der nogle forhold, der er empirisk veldokumenterede:

- Når besvarelses-devicet er valgfrit, optræder der selvselektionseffekter, og sociodemografiske variabler som for eksempel køn, alder, uddannelse og etnisk baggrund har betydning for surveyrespondenters valg af besvarelses-device. Antoun (2015) taler ligefrem om et "device divide" i valget af besvarelses-device (ibid.: 103).
- Web surveys, der kun udbydes til smartphone eksempelvis via en survey-applikation (i litteraturen omtalt f.eks. "mobile-only surveys" eller "smartphone-based surveys"), er forbundet med dækningsfejl. Dette skyldes, at det ikke er alle i målpopulationen, der har adgang til en smartphone (Antoun, 2015: 100; Couper, Antoun & Mavletova, 2017: 135-136; Antoun et al., 2019: 93-94).

Hvad angår selvselektionseffekter, så er respondenter, der besvarer web surveys på en smartphone, typisk yngre, mere veluddannede, oftere kvinder og tilhører oftere etniske minoritetsgrupper end PC-respøndenter (Antoun, 2015: 114; Keusch & Yan, 2017: 751; Sommer, Diedenhofen & Musch, 2017: 378).

Med hensyn til dækningsfejl så er denne fejltipe formentlig mindre betydningsfuld i et land som Danmark, hvor tæt på 90 % af den danske befolkning i aldersgruppen 15-89 år har adgang til en smartphone. Det er dog fortsat væsentligt at have for øje, særligt når ældre borgere, der er mindre fortrolige med de nye mobile devices, indgår i

målpopulationen (Nielsen & Tassy, 2018). Som Couper, Antoun og Mavletova (2017) påpeger, knytter respondentens fortrolighed med device sig i denne sammenhæng til spørgsmålet om dækning: "A distinction should be made between possession of a device and ability to reach a respondent using that device. (...) To be covered, sample members must also receive the invitation that is sent to them" (ibid.: 136-137).

Den eksisterende forskningsmæssige viden om egentlige device-effekter på datakvaliteten i web surveys, altså effekter der ikke skyldes skævheder pga. selv-selektion eller en mangelfuld dækning, er derimod mere begrænset. Tabel 2 indeholder en oversigt over væsentlige studier af besvarelses-devicets betydning for datakvaliteten, deres metode og datagrundlag, anvendte indikatorer for surveykvalitet samt studierne hovedresultater. Fælles for studierne er det, at de, ligesom dette studie, måler datakvalitet med udgangspunkt i TSE-rammearbejdet og dets sondring mellem repræsentation og måling.

Tabel 2. Oversigt over væsentlige studier om devices betydning for datakvaliteten

Studie	Design, stikprøve og surveyinstrument	Land	Besvarelses-devices	Indikatorer for surveykvalitet	Hovedresultater
Andreadis (2015)	<ul style="list-style-type: none"> Tværsnitstudie: Selvsselektion af device. Stikprøve: Den græske stemmerådgivnings-app 'HelpMeVote' (ikke probability-baseret). Stikprøve: Ca. 80.000. Antal items: 31. 	Grækenland	Smartphone, PC	<ul style="list-style-type: none"> Tendensen til at vælge ikke-svar ('Hverken/eller' og 'Ønsker ikke at svare') Tendensen til at vælge den første svarkategori (primacy effekter) 	<ul style="list-style-type: none"> Ingen signifikante forskelle.
Antoun, Couper & Conrad (2017)	<ul style="list-style-type: none"> Eksperimentelt crossover studie: Respondenterne besvarede den samme survey på både PC og smartphone. Stikprøve: Det probability-baserede onlinepanel LISS. Stikprøve: 1.390. Antal items: 46. 	Holland	Smartphone, PC	<ul style="list-style-type: none"> Tendensen til at... Svare det samme på en serie af items med samme svarkategorier Give afrundede numeriske svar Give korte svar på åbne spørgsmål Vælge lukkede svarkategorier frem for åbne Lavt kognitivt refleksionsniveau (CRT) Give socialt uønskede svar 	<ul style="list-style-type: none"> Færre smartphone- end PC- respondenter giver meget korte svar på åbne spørgsmål. Derudover ingen signifikante forskelle.
de Bruijne & Wijnant (2013)	<ul style="list-style-type: none"> Eksperimentelt studie: Respondenterne opdelt i tre treatmentgrupper: (1) Besvarelse på PC, (2) Besvarelse på mobile devices og (3) Besvarelse på PC, men med smartphone-layout. Stikprøve: Det probabilitybaserede panel 'CentERpanel' Stikprøve: 2.029. Antal items: Ukendt. 	Holland	PC, mobile devices	<ul style="list-style-type: none"> Nonresponse Tendensen til at vælge den første svarkategori (primacy effekter) Besvarelsestid Positiv eller negativ evaluering af spørgeskema 	<ul style="list-style-type: none"> Markant større nonresponse i mobile devices-gruppen Længere besvarelsestid på mobile devices Derudover ingen signifikante forskelle

Studie	Design, stikprøve og surveyinstrument	Land	Besvarelses-devices	Indikatorer for surveykvalitet	Hovedresultater
Keusch & Yan (2017)	<ul style="list-style-type: none"> • Eksperimentelt design: Respondenterne opdelt i to treatmentgrupper: (1) PC-gruppen og (2) iPhone-gruppen. • Stikprøveramme: MTurk. • Stikprøve: 1.085. Survey kun gennemført blandt iPhone-ejere. • Antal items: 27. 	USA	iPhone, PC	<ul style="list-style-type: none"> • Breakoffrate • Gennemsnitlig besvarelsesetid • Item nonresponse • Tendensen til at svare det samme på en serie af items med samme svarkategorier • Tendensen til at være enig, uanset spørgsmålet • Tendensen til at vælge den midterste svarkategori • Tendensen til at vælge ikke-svar 	<ul style="list-style-type: none"> • Højere break-offrate, længere gennemsnitlig besvarelsesetid og højere item nonresponse i iPhone-gruppen men forskellene er ikke statistisk signifikante. • Signifikant større tendens til at vælge samme svarkategori på en serie af items med samme svarkategorier i PC-gruppen. • Derudover ingen signifikante forskelle.
Lugtig & Toepoel (2016)	<ul style="list-style-type: none"> • Longitudinelt studie: Gennemført over 6 runder, med selvselektion af besvarelses-device. • Stikprøveramme: Det probabilitybaserede onlinepanel LISS. • Stikprøve: 6.229. • Antal items: Ukendt. 	Holland	Smartphone, PC, tablet	<ul style="list-style-type: none"> • Item nonresponse • Længden af svar på åbne spørgsmål • Tendensen til at svare det samme på en serie af items med samme svarkategorier • Tendensen til at vælge den første svarkategori (primacy effekter) • Gennemsnitligt antal valgte items på spørgsmål med flere svarmuligheder (CATA) • Gennemsnitlig besvarelsesetid • Positiv eller negativ evaluering af spørgeskemaet 	<ul style="list-style-type: none"> • PC-respondenter producerer data af højere kvalitet på alle parametre med undtagelse af, at de er mere tilbøjelige til at svare det samme på en serie af spørgsmål med samme svarkategorier. • Forskellene skyldes dog selvselektion: "Device"-skiftende respondenter ensartede data over tid.
Mavletova (2013)	<ul style="list-style-type: none"> • Eksperimentelt studie: Respondenterne opdelt i to treatmentgrupper: (1) PC-gruppen og (2) Smartphone-gruppen. • Stikprøveramme: Russisk ikke-probabilitybaseret onlinepanel. • Stikprøve: 513 i survey 1 og 500 i survey 2. 	Rusland	Smartphone, PC	<ul style="list-style-type: none"> • Breakoffrate • Tendensen til at vælge den første svarkategori (primacy effekter) • Tendensen til at vælge ikke-svar • Villigheden til at give socialt uønskede svar • Tendensen til at give korte svar på åbne spørgsmål 	<ul style="list-style-type: none"> • Signifikant højere breakoffrate i smartphone-gruppen • Markant højere gennemsnitlig besvarelsesetid i smartphone-gruppen. • Kortere svar på åbne spørgsmål i smartphone-gruppen. • Derudover ingen signifikante forskelle

Studie	Design, stikprøve og surveyinstrument	Land	Besvarelses-devices	Indikatorer for surveykvalitet	Hovedresultater
Schlusser & Mays (2018)	<ul style="list-style-type: none"> • Antal items: 21 i survey 1 og 54 i survey 2. • Eksperimentelt studie: Respondenterne opdelt i to treatmentgrupper: (1) PC-gruppen og (2) Mobile devices-gruppen. • Stikprøveramme: Studentsample. • Stikprøve: 519. • Antal items: 32. 	Tyskland	PC, mobile devices	<ul style="list-style-type: none"> • Reaktionsid på invitation til deltagelse • Breakoffrate • Item nonresponse • Tendensen til at give korte svar på åbne spørgsmål • Besvarelsesid 	<ul style="list-style-type: none"> • Højere gennemsnitlig besvarelsesid på mobile devices, men dette kan tilskrives en højere surveyloadingtid. • Derudover ingen signifikante forskelle.
Sommer, Diedenhoff & Musch (2017)	<ul style="list-style-type: none"> • Tværsnitstudie: Selvspektion af device. • Stikprøveramme: Panel ved Dusseldorf Universitet. • Stikprøve: 1.859. • Antal items: Ukendt. 	Tyskland	PC, mobile devices	<ul style="list-style-type: none"> • Besvarelsesid • Breakoffrate • Svarenes reliabilitet og validitet 	<ul style="list-style-type: none"> • Højere breakoffrate i mobile devices-gruppen. • Længere gennemsnitlig besvarelsesid på mobile devices. • Derudover ingen signifikante forskelle.
Struminskaya, Weyandt & Bosnjak (2015)	<ul style="list-style-type: none"> • Longitudinelt studie: Gennemført over 6 runder, med selvsektion af device. • Stikprøveramme: Det probability-baserede GESIS-panel. • Stikprøve: 2.913. • Antal items: Ukendt. 	Tyskland	PC, tablet, smartphone	<ul style="list-style-type: none"> • Item nonresponse • Tendensen til at give korte svar på åbne spørgsmål • Tendensen til at svare det samme på en serie af items med samme svarkategorier • Tendensen til at vælge den første svarkategori (primacy effekter) • Besvarelsesid 	<ul style="list-style-type: none"> • Højere item nonresponse for smartphone-besvarelser. • Større tendens til at vælge den samme svarkategori på en serie af items med samme svarkategorier på smartphone. • Smartphone- og tablet-respondenter giver kortere svar på åbne spørgsmål. • Signifikant højere gennemsnitlig besvarelsesid på tablet og smartphone. • Finder fortsat deviceeffekt når tiden inddrages som variabel modsat Toepoel & Lugtig (2016).

Studie	Design, stikprøve og surveyinstrument	Land	Besvarelses-devices	Indikatorer for surveykvalitet	Hovedresultater
Toepoel & Lugtig (2014)	<ul style="list-style-type: none"> • Tværsnitstudie: Selvslekction af device. • Stikprøveramme: Det probabilitybaserede onlinpanel 'Market Research'. • Stikprøve: 443. • Antal items: 12. 	Holland	Smartphone, PC	<ul style="list-style-type: none"> • Længden af svar på åbne spørgsmål • Item nonresponse • Gennemsnitligt antal valgte items på spørgsmål med flere svarmuligheder (CATI) • Tendensen til at vælge den første svarkategori (primacy effekter) 	<ul style="list-style-type: none"> • Ingen signifikante forskelle.
Tourangeau et al. (2018)	<ul style="list-style-type: none"> • Eksperimentelt studie: Respondenterne blev tilfældigt opdelt i én af tre treatmentgrupper: (1) Tablet, (2) PC og (3) Smartphone. • Stikprøveramme: Adresselister i 8 U.S. countries. • Stikprøve: 512. • Antal items: 69. 	USA	Smartphone, PC, tablet	<ul style="list-style-type: none"> • Item nonresponse • Tendensen til at svare det samme på en serie af items med samme svarkategorier • Besvarelsesetid • Reliabilitet i besvarelser 	<ul style="list-style-type: none"> • Højere gennemsnitlig besvarelsesetid hos smartphone-respondenterne. • Derudover ingen signifikante forskelle.

Note: I oversigten fokuseres der udelukkende på studier, der sammenligner datakvalitet på tværs af forskellige besvarelses-devices, og med fokus på direkte device-effekter. Sammenligning af forskellige indsamlingsmetoder internt på ét device samt studier af device-interaktionseffekter er derfor ikke blevet inddraget.

På repræsentationssiden af TSE-frameworket finder flere af studierne, som det også fremgår af tabellen, at respondentbortfaldet undervejs i surveyen er større på smart-phone sammenlignet med PC. Respondentbortfald henviser i denne sammenhæng både til en højere breakoffrate hos smart-phone-respondenterne samt et højere bortfald på item-niveau ('item nonresponse'), fordi respondenterne springer et eller flere enkeltspørgsmål over. Et nyere eksperimentelt studie af Keusch og Yan (2017) finder bl.a., at tæt på 10 % af de respondenter, der havde besvaret spørgeskemaet på en iPhone, havde sprunget mindst ét spørgsmål over, mens det tilsvarende tal for PC-respondenter kun var 3,6 % (ibid.: 758-759). Ifølge forskningen er mulige forklaringer på disse markante forskelle en mangelfuld tilpasning af survey-layoutet til mobile devices, da disse ofte som standard er designet til PC'en, samt forskellige brugs-kontekster for devices. Flere studier opfordrer dog fremtidig forskning til at generere mere dybdegående empirisk viden om de bagvedliggende årsager til disse device-forskelle (Antoun et al., 2018: 557-558; Couper, Antoun & Mavletova, 2017: 139-140).

På målingsiden af TSE-frameworket er resultaterne derimod mere tvetydige. Når forskelle i målefejl, der omhandler, hvor grundigt respondenterne har besvaret spørgeskemaet, undersøges, peger de fleste studier på enten marginale eller ingen forskelle mellem smart-phone- og PC-besvarelser. Et nyere eksperimentelt studie af Tourangeau et al. (2018), der primært sammenligner datakvaliteten i PC- og smart-phone-besvarelser, konkluderer om datakvaliteten: "data from the smartphone group seem, by most standards, to be just as good as those obtained on tablet or laptop computers." (ibid.: 550) Der er dog enkelte resultater, der peger i den modsatte retning. For eksempel finder et andet eksperimentelt studie af Keusch og Yan (2017) – imod den umiddelbare forventning – at smart-phone-respondenter er mindre tilbøjelige til at "straightline", dvs. at give det samme svar på en serie af spørgsmål med samme svarkategorier for at minimere den kognitive indsats under besvarelsen. På alle andre måleparametre for datakvalitet finder Keusch og Yan, ligesom Tourangeau et al. (2017), sammenlignelige resultater for PC og smart-phone.

Som det også fremgår af det foregående, sammenligner langt hovedparten af den eksisterende forskningslitteratur PC- og smart-phone-besvarelser. Det skyldes, at disse devices, som bekendt, er meget forskellige med hensyn til bl.a. brugskontekst, skærmstørrelse og andre tekniske features. Andre mobile devices og herunder særligt tablettens indplacering i dette spektrum er omdiskuteret, men state-of-the-art er i øjeblikket at placere disse i samme gruppe som PC'en grundet ligheder med hensyn til skærmstørrelse og brugskontekst (se f.eks. Couper, Antoun & Mavletova, 2017: 134; Peterson et al., 2017: 204; Wells, Bailey & Link, 2013: 2 & 10).

Mod den eksisterende forskning må gøres den indvending, at langt hovedparten af studierne som f.eks. Antoun, Couper & Conrad (2017) og Lugtig & Toepoel (2016) er gennemført på webpaneler frem for repræsentative befolknings-samples. Foruden mulige dækningsfejl, fordi panelet måske ikke er repræsentativt for målpopulationen

på alle parametre, er dette problematisk, fordi webpanelister er mere trænedede i at besvare surveys end den almene befolkning. En anden problemstilling er, at mange studier er gennemført på 'pseudo-spørgeskemaer' med få items konstrueret med metodeeksperimenterne for øje. Der mangler derimod undersøgelser med afsæt i større, rigtige videnskabelige surveys, der kræver en større kognitiv indsats af respondenterne og øger deres tilbøjelighed til at 'slække' på grundigheden i besvarelsenerne (Peterson et al., 2017: 219). En tredje problemstilling knytter sig til den udbredte anvendelse af det eksperimentelle design. Designet er et populært valg, fordi det eliminerer selvselektionseffekter. Det kan dog samtidig være med til at skabe en unaturlig situation, hvor respondenter kan blive tvunget til at anvende besvarelses-devices, de måske ikke er fortrolige med, med en mulig overestimering af device-effekterne til følge (se f.eks. diskussion i Lugtig & Toepoel, 2016: 80). Endelig peger den eksisterende forskning selv på et behov for en mere nuanceret og kontekstualiseret deviceforståelse, der blandt andet tager højde for, at der internt i device-kategorierne kan være meget store forskelle på de enkelte devices, eksempelvis i deres størrelse, layout, brugervenlighed og lignende (Antoun, Couper & Mavletova, 2017: 149-150).

3: Metode og data

3.1. DATA OG DELTAGERE

Det empiriske grundlag for artiklen er data fra den danske del af ISSP-undersøgelsen fra 2018, som er gennemført blandt et repræsentativt udtræk af danskere i alderen 18-79 år, samlet fra CPR-registret. Invitationsbrevet blev sendt digitalt via e-Boks og indeholdt et link til en selvadministreret web survey i SurveyXact. Data blev indsamlet i perioden ultimo november til ultimo december 2018, og nettostikprøven bestod af hele eller delvise besvarelser fra i alt 1.865 ud af de 5.000 udtrukne deltagere (svarprocent: 37 %).

Tabel 3 nedenfor viser stikprøvens køns- og aldersfordeling i forhold til populationens sammensætning. Som det ses, indeholder stikprøven en anelse flere mænd end andelen i populationen som helhed, mens aldersfordelingen er tæt på identisk med fordelingen i populationen. Eftersom mænd samtidig er en smule mere tilbøjelige til at anvende deres PC til besvarelse af web surveys end kvinder, er analyserne gennemført både med og uden vægte. Disse analyser viser ikke signifikante forskelle, og artiklens resultater er derfor gennemgående vist med uvejede data.

Tabel 3. Køns- og aldersfordeling i stikprøve og population (pct.)

	<i>Stikprøve</i>	<i>Population</i>
<i>Køn</i>		
Mand	56 %	50 %
Kvinde	44 %	50 %
<i>Alder</i>		
18-25	13 %	14 %
26-35	19 %	17 %
36-45	16 %	16 %
46-55	18 %	19 %
56-65	17 %	16 %
66-79	17 %	18 %

Kilde: Populationstallene er hentet fra Danmarks Statistik, FOLK1A, 4. kvartal 2018, der er sampling-tidspunktet, og populationen udgør 4.381.807. Oplysningerne om kønsfordelingen i stikprøven er registerbaserede, mens oplysningerne om aldersfordelingen i stikprøven er svarbaserede og indeholder 1.652 valide svar.

Temaet for undersøgelsen var danskernes holdninger og forhold til religion og religiøse praksisser og indgår som en del af den årlige internationale spørgeskemaundersøgelse ISSP i 2018 . Foruden 73 holdnings- og faktaspørgsmål om religion og religiøs praksis indeholdt spørgeskemaet et stort baggrundsmodule på 46 items, der bl.a. omhandlede respondentens uddannelse, job og civilstatus samt baggrundsspørgsmål om en eventuel ægtefælle eller partner.

3.2. OPERATIONALISERING

Device: Device-variablen, der er studiets primære uafhængige variabel, er blevet beregnet ud fra paradata om besvarelses-devicets skærmstørrelse. Dette logges automatisk i SurveyXact, når respondenter åbner spørgeskemaet i en webbrowser, og efterfølgende er paradataene blevet rekodet til en kategorisk variabel med tre kategorier (henholdsvis smartphone, tablet og PC) ud fra oplysninger om standardmål for devicestørrelser .

Vi er bevidste om, at denne fremgangsmåde kan være forbundet med reliabilitetsproblemer, da skærmstørrelsen ikke i alle sammenhænge er en god indikator for device-type. Derfor har vi forud for analyserne gennemført en 'reliabilitetskontrol' for målet ved at krydse det med et fakta-spørgsmål fra spørgeskemaet, hvor respondenter selv er blevet bedt om at angive besvarelses-devicet. Kontrollen viser, at reliabilitetsproblemer kan opstå i grænsefladerne mellem henholdsvis PC og tablet samt tablet og smartphone, hvor skærmstørrelserne i praksis kan være meget lig hinanden. Det overordnede resultat af kontrollen er dog, at reliabilitetsproblemerne er meget minimale.

Når paradata-målet er valgt i stedet for selvrapporert device skyldes det, at dette giver os mulighed for at koble besvarelses-device med andre paradata, der logges forud for besvarelsen – også når respondenter er bortfaldet undervejs i besvarelsen. Derudover er der ikke forskningsmæssig evidens for, at selvrapporerede data er mere troværdige end loggede paradata (se f.eks. Deng et al., 2019).

Surveykvalitet: Surveykvalitet defineres og måles med udgangspunkt i TSE-frameworket, der som tidligere nævnt opdeler fejlkilder i surveyundersøgelser i to overordnede typer: repræsentation og måling (se eksempelvis Groves et al., 2009; Frederiksen, Gundelach & Nielsen, 2017). På repræsentationssiden vil vi i denne artikel udelukkende fokusere på forskelle i nonresponse samt selektionseffekter i valg af device, idet variationer i brug af device først introduceres i surveyundersøgelsen efter samplingprocessen og har dermed ikke har betydning for eventuelle dækningsfejl. I logisk forlængelse af dette er det ikke muligt at undersøge samlede svarprocenter for forskellige besvarelses-devices, og nonresponse vil derfor i artiklen blive operationaliseret som breakoffs undervejs i besvarelsen. Selektionseffekter er relevante at undersøge, da det har været valgfrit for respondenterne, hvilket device, de ønsker at tilgå skemaet på, eftersom der ikke i denne undersøgelse har været fastsat begrænsninger i den henseende. Derfor undersøges det nærmere, om der er bestemte typer af respondenter, der systematisk foretrækker et bestemt device, samt om dette kan forklare eventuelle forskelle på besvarelses-devices.

På målingssiden fokuseres der i denne artikel udelukkende på fejl, der opstår i forbindelse med respondenterens besvarelse af surveyen, de såkaldte målefejl. I denne kontekst vil vi undersøge, om respondenterne gør sig større kognitiv anstrengelse med deres besvarelser på nogle devices end andre. Graden af kognitivt engagement i besvarelsen af et spørgeskema er i surveylitteraturen blevet målt både direkte og indirekte og ved hjælp af både konstruerede og ikke-konstruerede mål (Baumgartner & Steenkamp 2001). I analyserne i denne artikel inddrager vi to direkte mål:

- Et konstrueret mål i form af et afsluttende spørgsmål i spørgeskemaet, som måler respondenterens selvvaluerede indsats.
- Et ikke-konstrueret mål i form af loggede paradata om respondenterens besvarelsestid.

For sidstnævnte mål er antagelsen i artiklen, at længere besvarelsestid er et udtryk for større kognitivt engagement i besvarelsen.

3.3. ANALYSEMETODE

Da forskningsfeltet endnu er nyt, og den eksisterende viden om device-effekter er relativt begrænset, benytter vi i dette studie en kombination af deskriptive analysemetoder til at efterteste den eksisterende viden på feltet. Analysen er gennemført i tre trin.

I første trin beregnes frekvensfordelinger for besvarelses-device (smartphone, tablet og pc) krydstabuleret med sociodemografiske nøglevariable (køn, alder, uddannelse og etnisk baggrund). Der bliver i denne forbindelse foretaget χ^2 -test for at undersøge, om der er signifikante selvselektionseffekter i stikprøven. Disse analyser danner baggrund for kontrolanalyser i andet og tredje trin, hvor device-effekterne analyseres nærmere.

I andet trin krydstabuleres besvarelses-device med breakoffraten i stikprøven, og der beregnes χ^2 -test for at teste, om besvarelses-device har signifikant betydning for tendensen til at bortfalde undervejs i surveyen.

I tredje og sidste trin, hvor variable er kontinuerte, beregnes gennemsnitsværdier og standardafvigelse (S.D.) for besvarelses-tid og selv vurderet kognitiv indsats i besvarelsen af surveyen for hvert besvarelses-devices. Begge mål anvendes som indikatorer for størrelsen af målefejl i besvarelseserne.

I alle tre trin beregnes 95 % sikkerhedsinterval (herefter "95 % SI") for andele og gennemsnit.

4: PC-, smartphone- og tablet-responderer – er der en forskel?

Da besvarelses-devicet er valgfrit for respondenterne i den danske ISSP-undersøgelse, sigter denne første del af analysen mod at afdække, hvorvidt der er signifikante selvselektionseffekter i undersøgelsen, som vi skal tage højde for i analysen af device-effekter.

Som det fremgår af Tabel 4, er der statistisk signifikante sammenhænge mellem alle de inddragede sociodemografiske variable og det valgte besvarelses-device ($p < 0,05$). For begge køn gælder det, at PC'en er det foretrukne besvarelses-device, men i overensstemmelse med eksisterende forskning svarer en større procentandel af mændene i stikprøven på PC, hvorimod lidt flere kvinder end mænd anvender smartphonen som besvarelses-device. Forskellen er dog marginal og, som 95 % sikkerhedsintervaller viser, ligger den inden for den statistiske usikkerhed. Køn udgør dermed ikke den mest betydningsfulde faktor i valget af besvarelses-device.

Hvor tidligere forskning viser, at smartphone-responderer typisk er yngre end PC-responderer, er vores resultater imidlertid mindre entydige. I den yngste respondent-gruppe (18-25 år) gælder det nemlig, at PC'en er det foretrukne besvarelses-device, hvorimod den største procentandel af smartphone-respondererne findes i aldersgruppen 36-45 år. I overensstemmelse med tidligere resultater er PC'en dog klart det mest udbredte besvarelses-device blandt de to ældste respondentgrupper (56-65 og 66-75 år).

Endvidere er smartphone-respondenterne generelt mere veluddannede, idet omkring 40 % af disse har en gymnasial uddannelse (STX, HF eller HHX), mens det samme tal for respondenterne med en folkeskoleeksamen kun er godt 30 % .

Endelig viser vores resultater, ligesom tidligere forskning, at en større procentandel af respondenterne med mindst én forælder af anden etnisk oprindelse har besvaret surveyen via smartphone, dog med forbehold for et relativt lille n og stor usikkerhedsmargin i indvandregrupperne (n=194).

De mest interessante fund i forhold til tabletten, der fortsat udgør en udefinerbar midterkategori i spektret mellem PC og smartphone, er, at markant flere ældre end yngre, samt respondenter med etnisk danske forældre, anvender denne som besvarelses-device, og noget tyder ligeledes på, at tablet-respondenter har et lavere uddannelsesniveau.

Tabel 4. Demografiske kendetegn ved PC-, smartphone- og tablet-respondenter. Rækkeprocenter, n, 95 % SI og χ^2

	PC	95 % SI	Tablet	95 % SI	Smart- phone	95 % SI	Total	χ^2
<i>Køn</i>								
Mand	52,0 (478)	±3,2	13,9 (128)	±2,2	34,1 (313)	±3,1	100 (919)	$\chi^2=9,16$ 8, df=2, p=0,010
Kvinde	44,6 (334)	±3,6	15,6 (117)	±2,6	39,8 (298)	±3,5	100 (749)	
<i>Alder</i>								
18-25	52,8 (114)	±6,7	1,9 (4)	~	45,4 (98)	±6,7	100 (216)	$\chi^2=251,$ 927, df=10, p<0,001
26-35	41,0 (129)	±5,5	7,3 (23)	±2,9	51,7 (163)	±5,5	100 (315)	
36-45	34,8 (94)	±5,7	9,6 (26)	±3,5	55,6 (150)	±6,0	100 (270)	
46-55	44,1 (128)	±5,7	16,9 (49)	±4,3	39,0 (113)	±5,6	100 (290)	
56-65	54,9 (158)	±5,8	27,4 (79)	±5,1	17,7 (51)	±4,4	100 (288)	
66-79	67,4 (184)	±5,6	22,0 (60)	±4,9	10,6 (29)	±3,6	100 (273)	
<i>Skoleuddannelse</i>								
Folkeskole	49,4 (322)	±3,9	19,5 (127)	±3,1	31,1 (203)	±3,5	100 (652)	$\chi^2=33,6$ 69, df=4, p<0,001
STX/HF	45,7 (295)	±3,9	14,1 (91)	±2,7	40,2 (260)	±3,7	100 (646)	
HHX	53,9 (159)	±5,7	6,8 (20)	±2,9	39,3 (116)	±5,6	100 (295)	
<i>Forældres oprindelse</i>								
Begge født i Danmark	48,7 (694)	±2,6	15,4 (220)	±1,8	35,9 (511)	±2,5	100 (1.425)	$\chi^2=10,6$ 59, df=4, p=0,031
Én forælder født i Danmark, én i udlandet	55,3 (21)	±15,8	2,6 (1)	~	42,1 (16)	±15,7	100 (38)	
Begge født i udlandet	50,0 (78)	±7,8	8,3 (13)	~	41,7% (65)	±7,8	100 (156)	

~: For få respondenter til estimering af sikkerhedsinterval

Note: Oplysninger om respondenternes køn, alder, skoleuddannelse og forældres oprindelse er svarbaserede og stammer fra ISSP 2018-undersøgelsen.

Således viser vores resultater, i overensstemmelse med tidligere forskning, at der er signifikante selvselektionseffekter i valget af besvarelses-device i web surveys. Derfor vil de sociodemografiske variabler blive inddraget som kontrolvariabler i analysen af, hvorvidt der eksisterer egentlige device-effekter, i de kommende afsnit.

5: Resultater

5.1. REPRÆSENTATION: HELE OG DELVISE BESVARELSER

På repræsentationssiden af TSE-frameworket har breakoff undervejs i besvarelsen som bekendt betydning for den samlede surveykvalitet. I undersøgelsen har vi paradata for tre respondentgrupper, henholdsvis respondenter der 1) gennemførte hele surveyen, 2) delvist gennemførte surveyen og 3) åbnede surveyen uden at afgive nogle svar. Som Tabel 5 viser, har valget af besvarelses-device tilsyneladende betydning for respondenters tilbøjelighed til at falde bort undervejs i besvarelsen. En større procentandel af smartphone-respondenterne gennemførte kun delvist surveyen end PC- og tablet-respondenterne, der har sammenlignelige resultater med hensyn til helt og delvist gennemførte besvarelser. Derfor er det også iøjnefaldende, at vi finder den største andel af respondenter, der har åbnet surveyen uden at afgive nogle svar, blandt PC-respondenterne, og sammenlignet med smart-phone-respondenterne er forskellen statistisk signifikant ($p < 0,05$). Dette fortæller os, at sammenhængen mellem besvarelses-device og bortfald er mere kompleks end først antaget, og at smartphone ikke per se genererer større bortfald på item-niveau.

Ved kontrol for alder og uddannelse, der var de variabler, der havde størst betydning for valget af besvarelses-device, forsvinder sammenhængen imidlertid. Forskellene kan derfor snarere tilskrives selvselektionseffekter, og deraf udledte brugerforskelle, end devicene i sig selv (for kontrol-analyser, se appendiks 2-4). Her må der dog gøres det forbehold, at variablerne er svarbaserede, og derfor er andelen af delvise besvarelser en anelse underestimeret i kontrolanalyserne.

**Tabel 5. Hele og delvise besvarelser efter besvarelses-device.
Rækkeprocenter, n og 95 % SI.**

	PC	95 % SI	Tablet	95 % SI	Smart- phone	95 % SI
<i>Gennemførte survey</i>	89,1 (799)	±2,1	91,2 (238)	±3,5	82,8 (594)	±2,7
<i>Delvis besvarelse</i>	10,0 (90)	±1,9	8,4 (22)	±3,3	17,0 (122)	±2,7
<i>Åbnede survey uden at give svar</i>	0,9 (8)	~	0,4 (1)	~	0,1 (1)	~
Total	100 (897)		100 (261)		100 (717)	

Note: $\chi^2=26,302$, $df=4$, $p<0,001$. Kontrolleret for alder og uddannelse er sammenhængen ikke længere statistisk signifikant. Det har ikke været muligt at kontrollere for etnisk baggrund grundet manglende variation på den afhængige variabel (gennemførelsesstatus).
n=1.875.

5.2. MÅLING: FORSKELLE I DATAKVALITETEN

På målingssiden af TSE-frameworket har respondenternes grundighed i besvarelsen af spørgeskemaet betydning for surveykvaliteten. I undersøgelsen inddrager vi som tidligere nævnt to direkte mål herfor, ét konstrueret mål (respondentens selvvalgte kognitive indsats) og ét ikke-konstrueret mål (besvarelestid).

Som Tabel 6 viser, er PC-, tablet- og smartphone-respondenters gennemsnitlige selvvalgte kognitive indsats tilnærmelsesvis den samme, og statistisk test for forskel i gennemsnit er insignifikant på 0,05-niveau. Således er der ikke, ud fra respondenternes egen vurdering, forskel på grundigheden i besvarelsene afhængig af besvarelses-devicet.

Tabel 6. Selvvalgt kognitiv indsats efter besvarelses-device. Gennemsnit, S.D. og 95 % SI

	<i>Gennemsnit</i>	<i>S.D.</i>	<i>95 % SI</i>
<i>PC</i>	3,06	0,83	±0,06
<i>Tablet</i>	2,99	0,94	±0,11
<i>Smartphone</i>	3,04	0,88	±0,07

Note: Spørgsmål: "Enig eller uenig. Jeg brugte meget tid på at tænke mine svar på spørgsmålene i spørgeskemaet igennem for at besvare så præcist som muligt". Svarkategorier: Meget enig, enig, uenig og meget uenig. Skalaen går fra 1-4, hvor 1: Meget lille enighed og 4: Meget stor enighed. n=1588.

Tilsvarende resultater finder vi, når besvarelestid i minutter inddrages som indikator på grundighed i besvarelsene. Besvarelestiden er beregnet som differencen mellem start- og sluttidspunkt for besvarelsen, og meget høje besvarelestider må antages at være et udtryk for, at surveyen er besvaret ad flere omgange. Derfor afrapporteres den gennemsnitlige besvarelestid med forskellige øvre "cutpoints" med henblik på at teste målets robusthed. Skæringspunktet "60 minutter eller mindre" er valgt ud fra pilottesten af skemaet, der viste en forventet gennemsnitlig svartid på ca. 30 minutter, og ikke højere end 60 minutter, mens skæringspunktet "75 minutter eller mindre" er valgt ud fra statistiske analyser af outliers i data. Som Tabel 7 nedenfor viser, er det fælles for målene, at smart-phone-respondenterne er den gruppe, der gennemsnitligt bruger kortest tid på besvarelsen.

Tabel 7. Besvarelsestid i minutter efter device. Gennemsnit, S.D. og 95 % SI

	<i>Gennemsnit</i>	<i>S.D.</i>	<i>95 % SI</i>	<i>% outliers</i>
Besvarelsestid for besvarelser <= 60 minutter				
<i>PC</i>	28,0	11,4	±0,95	16,6 %
<i>Tablet</i>	28,8	10,1	±1,5	15,5 %
<i>Smartphone</i>	26,6	10,4	±0,82	16,3 %
Besvarelsestid for besvarelser <= 75 minutter				
<i>PC</i>	29,1	12,5	±0,95	14,0 %
<i>Tablet</i>	30,0	12,1	±1,7	12,6 %
<i>Smartphone</i>	27,8	12,5	±1,1	13,8 %

Note: Paradata blev benyttet til at beregne besvarelsestiden (forskel på start- og sluttidspunkt, dvs. at respondenterne ikke behøver at have interageret med spørgeskemaet hele tiden).

n=1.374-1.642.

Forskellene er dog marginale, og test for forskel mellem grupperne er ikke statistisk signifikante på 0,05-niveau. Heller ikke når den procentvise andel af outliers sammenlignes, finder vi nogle signifikante forskelle på tværs af besvarelses-devicene. Der er således ikke empirisk belæg for, at smartphone-respondenter er mere tilbøjelige til at besvare surveyen ad flere omgange, hvilket er imod den umiddelbare forventning ud fra dette device' mere urolige brugskontekst og de kortere og mere fragmenterede brugssessioner.

Dette indikerer således, at der heller ikke er signifikante device-forskelle, når det gælder den tid, respondenterne investerer i besvarelsen.

6: Konklusion og diskussion

Web mode, der bliver en stadig mere udbredt indsamlingsmetode, adskiller sig fra andre modes; ikke bare ved fraværet af en interviewer, men også det forhold, at vi som udgangspunkt ikke har kontrol over respondentens besvarelses-device. Brugen af mobile devices til at besvare web surveys bliver stadig mere udbredt i disse år, og i og med at brugsmønstrene for devices er forskellige, var formålet med dette studie at undersøge, om valget af besvarelses-device har betydning for den samlede surveykvalitet, når respondenterne anvender et device, som vedkommende er fortrolig med. Da surveyforskningen i device-effekter endnu er på et spædt stadie og ikke er uden metodiske problemer, bidrager studiet til at efterteste hidtidige resultater i en repræsentativ befolkningssample og på en større videnskabelig survey.

I overensstemmelse med hidtidig forskning finder vi minimale device-effekter på surveykvaliteten. På repræsentationssiden finder vi ganske vist, at smartphone-respondenter, sammenlignet med PC- og tablet-respondenter, har en større

breakoffrate. Kontrolleret for de sociodemografiske nøgle-variabler alder og uddannelse forsvinder sammenhængen imidlertid, og forskellene kan derfor snarere tilskrives selvselektionseffekter end egentlige device-effekter. På målingssiden finder vi, på to direkte mål for respondenternes kognitive engagement i besvarelsen, ingen forskelle på besvarelses-device, når der tages højde for den statistiske usikkerhed. Alt i alt peger vores resultater således på, at forskelle i surveykvalitet skyldes forskelle i brugerne og ikke er effekter afledt af besvarelses-devicet i sig selv. Givet den større kognitive belastning, som respondenterne har været udsat for i dette studie, finder vi det usandsynligt, at device har selvstændig indflydelse på surveykvaliteten i andre kontekster – i hvert fald så længe surveylayoutet er tilpasset til mobile devices. Dog må der tages det forbehold, at undersøgelsen er foretaget i Danmark, hvor størstedelen af befolkningen både har adgang til og er relativt fortrolige med mobile devices som f.eks. en smartphone. Samtidig kan det ikke udelukkes, at der er særlige forhold, der har gjort sig gældende i denne undersøgelse som f.eks., at surveylayoutet har været optimeret til smartphones. Derfor kan resultaterne ikke ukritisk overføres til andre kontekster, hvor færre har adgang til smartphones, og det er derfor både påkrævet og nødvendigt, at undersøgelsens resultater eftertestes på andre surveys og i andre sammenhænge i fremtiden.

Mod studiets konklusioner må naturligvis gøres nogle metodiske indvendinger. Studiet anvender et tværsnitsdesign med selvselektion af besvarelses-device, og derfor kan der forekomme selvselektionseffekter, som det ikke er muligt for os at tage højde for inden for rammerne af studiet. Dette er dog ikke udelukkende en metodisk svaghed, da det giver et indblik i deviceeffekter, når respondenterne får lov til at vælge et device, som han/hun er fortrolig med. Derudover skal det nævnes, at studiets målpopulation (danske borgere i alderen 18-79 år) udelukker de ældste borgere, der forventeligt er mindst fortrolige med de nye medier, hvilket kan have implikationer for estimeringen af device-effekter. Da besvarelses-devicet er selvselektet, forventer vi dog ikke, at dette øver stor indflydelse på resultaterne, da respondenterne formodentlig vælger devices, de i forvejen er bekendte og fortrolige med. Endelig må det indvendes, at studiet har bibeholdt en kategorisk inddeling, hvor devicetyper bliver afgørende for respondentens indplacering. Da der er store forskelle internt i de enkelte devicetyper (Couper, Antoun & Mavletova 2017: 149-150; Wells, Bailey & Link 2013: 7), kan fremtidig forskning på feltet med fordel afprøve nye kategoriske inddelinger, hvor eksempelvis navigations- eller styresystemet afgør respondentens indplacering frem for devicetyper, eller alternativt kontinuerte mål med for eksempel skærmstørrelse som uafhængig variabel.

I forlængelse heraf vil vi tilskynde fremtidige studier til, i højere grad end det har været muligt i dette studie, at inddrage mediekonteksten i et forsøg på at forklare og forstå, hvorfor bestemte surveyrespondenter vælger bestemte devices, og hvorvidt brugskonteksten for mediet har betydning for surveykvaliteten (Couper, Antoun & Mavletova 2017). Her kan med fordel anlægges mere kvalitative tilgange, der giver

indsigter i respondentens oplevelser af surveys, og variationer i besvarelseskontekst, på tværs af forskellige devices.

For designere af web surveys er studiets resultater alt i alt godt nyt. Besvarelses-devicet har ikke afgørende indflydelse på surveykvaliteten, og derfor skal afsendere af web surveys ikke afstå fra at udbyde web surveys på forskellige devices. Tværtimod mener vi, at mulige fejlkilder knyttet til udelukkelsen af besvarelser fra mobile devices som for eksempel dækningsproblemer så langt overstiger ulemperne ved at tillade mobile besvarelser. Dog er det her en helt afgørende forudsætning, at web survey-layoutet er optimeret til mobile devices. Da det er forskningsmæssigt veldokumenteret, at smartphone-respondenter producerer større item nonresponse – uanset at det reelt skyldes selvselektion – kan viden om besvarelses-device endvidere bruges i dataredigeringsfasen til at identificere mulige 'problem-respondenter'.

Litteratur

- Andreadis, I. (2015). Web Surveys Optimized for Smartphones: Are there Differences Between Computer and Smartphone Users? *Methods, data, analyses*, 9(2), s. 213-228.
- Antoun, C. (2015). Who Are the Internet Users, Mobile Internet Users, and Mobile-Mostly Internet Users?: Demographic Differences across Internet-Use Subgroups in the U.S. I: Toninelli, D., Pinter, R. & de Pedraza, P. (red.). *Mobile Research Methods. Opportunities and Challenges of Mobile Research Methodologies* (kapitel 7, s. 99-118). London: Ubiquity Press Ltd.
- Antoun, C., Conrad, F. G., Couper, M. P. & West, B. T. (2019). Simultaneous estimation of multiple sources of error in a smartphone-based survey. *Journal of Survey Statistics & Methodology*, 7(2019), s. 93-117.
- Antoun, C., Couper, M. P. & Conrad, F. G. (2017). Effects of Mobile versus PC Web on Survey Response Quality: A Crossover Experiment in a Probability Web Panel. *Public Opinion Quarterly*, 81(S1), s. 280-306.
- Antoun, C., Katz, J., Argueta, J. & Wang, L. (2018). Design Heuristics for Effective Smartphone Questionnaires. *Social Science Computer Review*, 36(5), 557-574.
- Baumgartner, Hans & Jan-Benedict E.M. Steenkamp (2001): Response Styles in Marketing Research: A Cross-National Investigation. *Journal of Marketing Research*, vol. XXXVIII (May 2001), s. 143-156.
- Couper, M. P. (2000). Web surveys. A review of issues and approaches. *Public Opinion Quarterly*, 64(4), s. 464-494.
- Couper, M. P., Antoun, C. & Mavletova, A. (2017). Mobile Web Surveys: A Total Survey Error Perspective. I: Biemer, P. P. et al. (red.). *Total Survey Error in Practice* (kapitel 7, s. 133-154). New York: Wiley.
- Couper, M. P. & Peterson, G. J. (2017). Why Do Web Surveys Take Longer on Smartphones? *Social Science Computer Review*, 35(3), s. 357-377.
- de Bruijne, M. & Wijnant, A. (2013). Comparing Survey Results Obtained via Mobile Devices and Computers: An Experiment With a Mobile Web Survey on a Heterogenous Group of Mobile Devices Versus a Computer-Assisted Web Survey. *Social Science Computer Review*, 31(4), s. 482-502.
- de Leeuw, E. D. (2018) Mixed-Mode: Past, Present, and Future. *Survey Research Methods*, 12(2), s. 75-89.
- Deng, T., Kanthawala, S., Meng, J., Peng, W., Kononova, A., Hao, Q., Zhang, Q. & David, P. (2019). Measuring smartphone usage and task switching with log tracking and self-reports. *Mobile Media & Communication*, 7(1), s. 3-23.
- Dillman, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method* (2. udgave). New York: Wiley.
- Flyvbjerg, B. (2015): Fem misforståelser om casestudiet. I: Brinkmann, S. & Tanggaard, L. *Kvalitative metoder. En grundbog* (kapitel 24, s. 497-520). København: Hans Reitzels Forlag (2. udgave).
- Frederiksen, M., Gundelach, P. & Nielsen, R. S. (2017). *Survey. Design, stikprøve, spørgeskema, analyse*. København: Hans Reitzels Forlag.
- Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey Methodology*. Hoboken: Wiley (2. udgave).
- Helles, R. (2016). Mobile medier. I: Jensen, K. B. (red.). *Danske mediehistorie 4: 1995-2015* (s. 57-74). Frederiksberg: Samfundslitteratur (2. udgave).

- Jenkins, H. (2006). *Convergence Culture. Where Old and New Media Collide*. New York: New York University Press.
- Keusch, F. & Yan, T. (2017). Web Versus Mobile Web: An Experimental Study of Device Effects and Self-Selection Effects. *Social Science Computer Review*, 35(6), s. 751-769.
- Klastrup, L. (2013). Mobile medier. I: Agger, G., Jauert, P., Sandvik, K. & Schrøder, K. (red.). *Medie- og kommunikationsleksikon*. Frederiksberg: Samfundslitteratur. Senest tilgået d. 25. august 2019: <http://medieogkommunikationsleksikon.dk.zorac.aub.aau.dk/mobile-medier/>
- Lugtig, P. & Toepoel, V. (2016). The Use of PCs, Smartphones, and Tablets in a Probability-Based Panel Survey: Effects on Survey Measurement Error. *Social Science Computer Review*, 34(1), s. 78-94.
- Mavletova, A. (2013). Data Quality in PC and Mobile Web Surveys. *Social Science Computer Review*, 31(6), s. 725-743.
- Nielsen, M. B. & Tassy, A. (2018). Mindst hver femte smartphone er ubeskyttet, *Nyt fra Danmarks Statistik*, 431(2018).
- Peterson, G., Griffin, J., LaFrance, J. & Li, J. (2017). Smartphone Participation in Web Surveys. Choosing Between the Potential for Coverage, Nonresponse, and Measurement Error. I: Biemer, P. P. et al. (red.). *Total Survey Error in Practice* (kapitel 10, s. 203-234). New York: Wiley.
- Roberts, C. E. (2007) Mixing modes of data collection in surveys: A methodological review. *ESRC National Centre for Research Methods, NCRM Methods Review Papers, NCRM/008*.
- Schlosser, S. & Mays, A. (2018). Mobile and Dirty: Does Using Mobile Devices Affect the Data Quality and the Response Process of Online Surveys? *Social Science Computer Review*, 36(2), s. 212-230.
- Sommer, J., Diederhufen, B. & Musch, J. (2017). Not to Be Considered Harmful: Mobile-Device Users Do Not Spoil Data Quality in Web Surveys. *Social Science Computer Review*, 35(3), s. 378-387.
- Struminskaya, B., Weyandt, K. & Bosnjak, M. (2015). The Effects of Questionnaire Completion Using Mobile Devices on Data Quality. Evidence from a Probability-based General Population Panel. *Methods, data, analyses*, 9(2), s. 261-292.
- Toepoel, V. & Lugtig, P. (2014). What Happens if You Offer a Mobile Option to Your Web Panel? Evidence From a Probability-Based Panel of Internet Users. *Social Science Computer Review*, 32(4), s. 544-560.
- Tourangeau, R., Sun, H., Yan, T., Maitland, A., Rivero, G. & Williams, D. (2018). Web Surveys by Smartphone and Tablets: Effects on Data Quality. *Social Science Computer Review*, 36(5), s. 542-556.
- Wells, T., Bailey, J. & Link, M. (2013). Filling the Void: Gaining a Better Understanding of Tablet-Based Surveys. *Survey Practice*, 6(1), s. 1-13.

Appendiks

APPENDIKS 1. UDVIKLING I BESVARELSES-DEVICE I DEN DANSKE ISSP-UNDERSØGELSE 2015-2018. KOLONNEPROCENTER SAMT ANTAL RESPONDENTER I PARENTES.

	2015-16	2017	2018
<i>PC</i>	69,8 (733)	57,6 (582)	49,8 (813)
<i>Tablet</i>	15,2 (160)	14,9 (150)	13,5 (221)
<i>Smartphone</i>	13,0 (136)	25,8 (261)	36,6 (597)
<i>Papir</i>	0,9 (9)	0 (0)	-
<i>Telefon</i>	-	1,0 (10)	-
<i>Flere besvarelses-devices</i>	1,1 (12)	0,7 (7)	-
Total	100 (1.050)	100 (1.010)	100 (1.631)

Kilde: Den danske ISSP-undersøgelse, 2015-2018. Spørgsmål: "Hvordan besvarede du dette spørgeskema?", svarkategorier: PC, smartphone, tablet, papir og telefon (hvis relevant). Det var muligt at vælge flere svarkategorier. Note: I 2017 overgik den danske ISSP-undersøgelse til hovedsageligt at udsende links til spørgeskemaer gennem digital post, hvor links til spørgeskemaerne tidligere blev udsendt gennem fysisk post. Dette betyder, at udviklingen i 2015-16 og 2017 muligvis skyldes en ændring i dataindsamlingsproceduren.

APPENDIKS 2. HELE OG DELVISE BESVARELSER EFTER BESVARELSES-DEVICE OPDELT PÅ KØN. KOLONNEPROCENTER SAMT ANTAL RESPONDENTER I PARENTES.

	<i>PC</i>	<i>Tablet</i>	<i>Smartphone</i>
Mænd			
<i>Gennemførte survey</i>	88,7 (472)	91,3 (126)	81,6 (305)
<i>Delvis besvarelse</i>	10,7 (57)	8,0 (11)	18,4 (69)
<i>Åbnede survey uden at give svar</i>	0,6 (3)	0,7 (1)	0 (0)
Kvinder			
<i>Gennemførte survey</i>	89,6 (327)	91,1 (112)	84,3 (289)
<i>Delvis besvarelse</i>	9,0 (33)	8,9 (11)	15,5 (53)
<i>Åbnede survey uden at give svar</i>	1,4 (5)	0 (0)	0,3 (1)

Note: $\chi^2=26,302$, $df=4$, $p<0,001$. Oplysninger om respondenternes køn er registerbaserede. $n=1.875$.

**APPENDIKS 3. HELE OG DELVISE BESVARELSER EFTER BESVARELSES-DEVICE
OPDELT PÅ ALDER. KOLONNEPROCENTER SAMT ANT. RESP. I PARENTES.**

	<i>PC</i>	<i>Tablet</i>	<i>Smartphone</i>
18-25			
<i>Gennemførte survey</i>	97,4 (111)	100 (4)	96,9 (95)
<i>Delvis besvarelse</i>	2,6 (3)	0 (0)	3,1 (3)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
26-35			
<i>Gennemførte survey</i>	99,2 (128)	100 (23)	99,4 (162)
<i>Delvis besvarelse</i>	0,8 (1)	0 (0)	0,6 (1)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
36-45			
<i>Gennemførte survey</i>	100 (94)	96,2 (25)	96,7 (145)
<i>Delvis besvarelse</i>	0 (0)	3,8 (1)	3,3 (5)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
46-55			
<i>Gennemførte survey</i>	96,9 (124)	95,9 (47)	96,5 (109)
<i>Delvis besvarelse</i>	3,1 (4)	4,1 (2)	3,5 (4)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
56-65			
<i>Gennemførte survey</i>	99,4 (157)	97,5 (77)	96,1 (49)
<i>Delvis besvarelse</i>	0,6 (4)	2,5 (2)	3,9 (2)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
66-79			
<i>Gennemførte survey</i>	97,8 (180)	96,7 (58)	100 (29)
<i>Delvis besvarelse</i>	2,2 (4)	3,3 (2)	0 (0)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)

Note: $\chi^2=2,109$, $df=2$, $p=0,348$. Da oplysninger om respondenternes alder er svarbaserede, er andelen af ”Delvis besvarelser” en anelse underestimeret i kontrolanalyserne.
n=1.617.

**APPENDIKS 4. HELE OG DELVISE BESVARELSER EFTER BESVARELSES-DEVICE
KONTROLLERET FOR UDDANNELSE**

	<i>PC</i>	<i>Tablet</i>	<i>Smartphone</i>
Folkeskole			
<i>Gennemførte survey</i>	99,1 (319)	95,3 (121)	98,5 (200)
<i>Delvis besvarelse</i>	0,9 (3)	4,7 (6)	1,5 (3)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
STX/HF			
<i>Gennemførte survey</i>	98,6 (291)	98,9 (90)	97,7 (254)
<i>Delvis besvarelse</i>	1,4 (4)	1,1 (1)	2,3 (6)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)
HHX			
<i>Gennemførte survey</i>	99,4 (158)	100 (20)	99,1 (115)
<i>Delvis besvarelse</i>	0,6 (1)	0 (0)	0,9 (1)
<i>Åbnede survey uden at give svar</i>	0 (0)	0 (0)	0 (0)

Note: $\chi^2=4,449$, $df=2$, $p=0,108$. Da oplysninger om respondenternes uddannelse er svarbaserede, er andelen af "Delvise besvarelser" en anelse underestimeret i kontrolanalyserne.
n=1.593.