

En undersøgelse af
Udvidede Dataindsamlingsformer
i
ONLINEBASEREDE SPØRGESKEMAER

THEIS THOMPSON
AALBORG UNIVERSITET
2015

Titelblad

**Titel: En Undersøgelse af Udvidede Dataindsamlingsformer
i Onlinebaserede Spørgeskemaer.**

Emneområde: Kvantitativ Metode.

Antal Ord: 33.764

Projektperiode: 10. Semester, Efterår/Vinter 2014-15

Theis Thompson

1 Indhold

2	Forord	6
3	Læsevejledning.....	7
4	Indledning.....	8
5	Problemstilling	11
6	Problemfelt	11
6.1	Spørgeskemaets komplikationer.....	12
6.2	Indsamlingsmodes.....	12
6.3	Designkomplikationer	14
6.3.1	Sampling og selektionsbias	14
6.3.2	Som man spørger får man svar.....	16
6.3.3	Formulering af spørgsmål.....	16
6.3.4	Formulering af svarkategorier.....	17
6.4	Respondentadfærd.....	18
6.4.1	Social desirability	18
6.4.2	Svarmønstre og satisficing	19
7	Undersøgelsens fokus	21
7.1	Det webbaserede spørgeskema	21
7.1.1	Fordele ved webbaserede spørgeskemaer.....	21
7.1.2	Komplikationer med webbaserede spørgeskemaer	23
7.1.3	Sampling og selektionsbias i onlinebaserede spørgeskemaer	25
7.1.4	Metodiske beskrivelser af webbaserede spørgeskemaer	26
7.1.5	Respondenttyper i onlinebaserede spørgeskemaer	26
8	Thompson Metoden	27
8.1	Datastruktur og udbytte	28
8.2	Rangering af dataindsamlingsmodes.....	29
8.3	Tertiære data.....	30
8.3.1	Tid.....	31
8.3.2	Svarskift	31
8.3.3	Progression	32
8.3.4	Samspil.....	32

8.3.5	Forskerens evalueringsevne	33
9	Formål.....	34
10	Problemformulering.....	35
10.1	Underspørgsmål	35
11	Spørgeskemaet som observationsstudie.....	38
11.1	Naturalistisk observation	39
11.2	Eventsampling som kvantitativ metode.....	40
11.3	Metodens begrænsninger	40
12	Etik.....	41
12.1	Anonymitet	42
12.2	Følelsen af anonymitet.....	42
12.3	Informeret samtykke.....	43
12.4	Hvad ved vi egentlig.....	44
12.5	Hvad er vi i stand til at vide	45
13	Operationalisering	47
13.1	Operationalisering af metoden	48
13.1.1	Teknisk afklaring	50
14	Empiri	51
14.1	Todelt spørgeskema.....	51
14.2	Designmæssige beslutninger	52
14.3	Spørgeskemaets struktur.....	54
14.4	Sample	56
14.4.1	Kontakt til respondenter.....	56
14.4.2	Komplikationer i respondentindsamling	57
15	Analyse.....	58
15.1	Resultater	58
15.1.1	Respondentoversigt	58
15.1.2	Tertiære data i besvarelserne	60
15.2	Respondentforløb.....	62
15.2.1	Enkeltstående forløb	62
15.3	Designforløb.....	69
15.3.1	Svarskift som henledende faktor	70

15.3.2	Events som henledende faktor	77
15.3.3	Tid som henledende faktor	80
15.3.4	Sammenligning af datasæt	85
15.4	Respondenttyper.....	86
15.4.1	Speeders, cheaters og repeaters.....	86
16	Diskussion.....	97
16.1	Andre respondenttyper	98
16.2	Metodens anvendelighed.....	99
17	Videreudvikling af metoden	101
17.1	Eventsampling.....	101
17.2	Scrolltracking	102
17.3	Out of focus-tracking	103
17.4	Tid, sted og genoptagelse	103
18	Konklusion	104
18.1	Spørgsmål 1	104
18.2	Spørgsmål 2	105
18.3	Spørgsmål 3	105
18.4	Spørgsmål 4	106
18.5	Opsummering	107
19	Litteraturliste	109

2 Forord

Specialet er udarbejdet af Theis Thompson, specialestuderende på Sociologi, Aalborg Universitet i perioden medio 2014 til primo 2015. Specialets indholdsmæssige længde er 37.737 ord.

Projektet har sit afsæt i et tidligere metodisk fundament, der blev etableret i forbindelse med udarbejdelse af forfatterens bacheloropgave i 2012. Målsætningen er således at undersøge og analysere metodens praktiske anvendelse.

I forbindelse med og til anvendelse i specialet er der blevet udviklet følgende af forfatteren;

- Webbaseret spørgeskemasoftware – med mulighed for både opsætning af spørgeskemaer, såvel som indsamling af besvarelser.
- Webbaseret analysesoftware, som evner at bearbejde de indsamlede data.

Jeg vil derudover gerne rette en personlig tak til Ruth Emerek, som gennem fyldestgørende feedback, engagement og interesse for emnet, har spillet en væsentlig rolle i forhold til udarbejdelsen af specialet.

3 Læsevejledning

Specialet indledes med en *indledning*, som har til formål at skabe et historisk indblik i spørgeskemaet som metode. Der vil i dette afsnit kort opridses de forskellige faser eller æraer som spørgeskemaet som metode har gennemgået.

I afsnittet *problemstilling* etableres et begrænset indblik i de problemer som projektet har til formål at belyse og afklare, ved at indsnævre det emnemæssige omfang. I afsnittet introduceres den metode der danner omdrejningspunkt for specialet.

Gennem *problemfeltet* søges der at etablere indsigt i spørgeskemaet som metode, for på den måde at belyse de problemer som er med metoden på tværs af de forskellige indsamlingsformer. I afsnittet belyses der de forskellige komplikationer, i forhold til design samt respondentadfærd.

I *undersøgelsen fokus* indsnævres projektets omfang til det webbaserede spørgeskema. Der gennemgås i afsnittet de fordele, ulemper og komplikationer som er unikke for den indsamlingsform.

Afsnittet *Thompson Metoden* er en detaljeret gennemgang af den anvendte metode i projektet, samt de afgrænsninger der er foretaget i forhold til operationaliseringen af metoden. I dette afsnit introduceres ligeledes begrebet og konceptet om rangering af data og datatyper, her især med fokus på de interaktionsrelaterede *tertiære data*.

Afsnittene *Formål* og *Problemformulering* uddyber projektets formål, samt problemformulering og tillægsspørgsmål.

I afsnittet *Spørgeskemaet Som Observationsstudie*, udforskes og sammenlignes de muligheder der findes i *Thompson Metoden*, med det traditionelle observationsstudie. Det diskuteres ligeledes hvilke forskellige samplingsformer er mulige, samt argumentation for den anvendte metode.

Afsnittet *Etik* adresserer de etiske dilemmaer og problemstillinger, der er i forbindelse med anvendelse af *Thompson Metoden*. Ved at tegne et billede af hvad vi med metoden er i stand til at vide, såvel som hvor meget vi er i stand til at indsamle diskuteres de etiske komplikationer der måtte være i forbindelse hermed.

Afsnittene *Operationalisering* og *Empiri* uddybes hvordan metoden er operationaliseret i forhold til både dataindsamling, såvel som de spørgeskemaer der har dannet fundament for indsamlingen af empirien der er anvendt i projektet.

Dernæst analyseres den indsamlede empiri i afsnittet *Analyse*. Afsnittet har til formål at danne en illustration af hvordan de indsamlede *tertiære* data kan anvendes og analyseres. Der er således tale om en metodisk analyse, hvorved at det indholdsmæssige i respondenternes besvarelser forbigås.

I afsnittet *Diskussion*, diskuteres nogle af de problemstillinger der har været oplevet i forbindelse med udarbejdelse af specialet, såvel som at reflektere, kritisere og diskutere de problemstillinger som metoden tilvejebringer.

Dernæst diskuteres nogle af de videre udviklingsmuligheder, i afsnittet *Videreudvikling af metoden*. Der skabes herigennem indsigt i nogle af de potentielle muligheder, som ville kunne implementeres for samtidigt at søge at forbedre det resultat der er produceret i forbindelse med projektet.

Afslutningsvis i projektet konklusion konkluderes og opsummeres der på specialets problemformulering, såvel som de tillægsspørgsmål, der er stillet i forbindelse med projektet.

4 Indledning

Historisk set, har spørgeskemaet været et udbredt redskab i den sociologiske værktøjskasse, som flittigt er blevet anvendt på tværs af populationsundersøgelser. Med sin spæde begyndelse i det sociologiske felt i midt 1800, så det først sin reelle anerkendelse og udvikling fra 1930 til 1960 (Groves 2011: 861 – 871). Spørgeskemaet som dataindsamlingsform blev allerede anvendt i det samfundsvidenskabelige felt i midten af 1800tallet men blev imidlertid først anerkendt som sociologisk metode op i det 20. århundrede. Fra anerkendelsen af spørgeskemaet i 1930 kan betragtes som tredelt; opfindelsens æra fra 1930 til 1960; ekspansionens æra fra 1960 til 1990; æraen for *designed* og *organic* data fra 1990 til og med i dag.

I den første epoke er der tale om et felt, som kun blot var blevet kortlagt. I denne periode udvikledes især dét, der i dag betragtes som fundamentale konventioner og metoder i

spørgeskemaundersøgelser såsom anvendelsen af et bredere udvalg af skalaer i undersøgelsesregi. Undersøgelsesformen bestod primært af ansigt-til-ansigt interviews, der ofte var centreret omkring geografisk sampling. Der har selvfølgelig været undtagelser heri, som belyser netop hvorfor at ansigt-til-ansigt blev valgt over andre alternativer, hvilket gennemgås i et senere afsnit.

I anden epoke fra 1960erne blev det mere udbredt blandt befolkningen at have telefon i hjemmet, hvilket muliggjorde, at spørgeskemaer kunne besvares på trods af, at interviewer og respondent ikke befandt sig samme sted. Det blev i forlængelse heraf mere alment og relevant for staten og offentlige instanser at opnå indsigt i de tiltag der blev lavet. Dette betød, at der i stigende grad blev dedikeret offentlige midler til undersøgelse af alt fra sundhed til levevilkår.

Sampling blev en essentiel del af den faglige debat, og der arbejdedes fortsat med mål om at overskride de hurdles, der var i undersøgelsesformen. Telefonens udbredelse gjorde geografiske placering til en mindre udfordring end tidligere om end der samtidig hermed oplevedes tendenser til faldende svarrate på skemaerne samt uafsluttede spørgeskemaundersøgelser, idet respondenterne i højere grad afsluttede telefoniske spørgeskemaer før tid (Groves 2011). Denne udvikling var incitament for at foretage de telefoniske undersøgelser kortere end ved et ansigt-til-ansigt scenarie. I forlængelse heraf opstod de første større debatter og forsøg på løsninger på flere forskellige bias-typer. Eksempelvis at benytte vægtning og post-stratificering til at modarbejde den non-response bias, der opstod i de nye undersøgelsesformer såvel som i øvrige teoretiske metoder.

Denne epoke gav ligeledes anledning til at anvende spørgeskemaet i private virksomheder, der bestræbte sig på kortere og ofte mere specifikke skemaer. Dette medførte en konkurrence om respondenterne, hvilket hidtil havde været en noget nær ukendt problemstilling. Denne stigende interesse betød ligeledes, at forskningsinstitutter og offentlige instanser måtte afsætte flere ressourcer i forsøget på at overbevise mulige respondenter til deltagelse (Groves 2011).

Den sidste og aktuelle epoke gav i 1990erne anledning til en række nye muligheder og problemstillinger, hvor især det webbaserede spørgeskema har været i centrum. Frekvensen i det traditionelle ansigt-til-ansigt survey aftog især grundet de implicitte høje

omkostninger og geografiske begrænsninger. Ligeledes vanskeliggjorde mobiltelefonens stigende popularitet den geografiske stratifikation på baggrund af telefonnummeret, der ligeledes oftere end før blev udskiftet. Dette er dog kun en problemstilling, hvis der er tale om undersøgelser, som er egnet til en bestemt region eller område – såvel som at man ønsker at sikre kontakt til bestemte respondenter, hvor man muligvis har adgang til registerdata eller lignende fra. Dernæst blev det lettere for respondenter at screene deres opkald og afvise opkaldet allerede inden en potentiel påbegyndt samtale (Grover 2011).

I forbindelse med internettet og hjemmecomputerens udbredelse op gennem 1990erne blev det muligt at opsætte og dele spørgeskemaer elektronisk (Couper 2001). Dette muliggjorde en hurtigere og mindre ressourcekrævende omdeling af spørgeskemaer såvel som indsamling af besvarelser fra respondenterne. Hvor det tidligere var nødvendigt at distribuere spørgeskemaet enten via post, personligt eller dataindsamling ved assisteret besvarelse – telefonisk eller ansigt-til-ansigt – blev det med den teknologiske udvikling muligt at overlade hele besvarelsesprocessen til respondenterne (Roster et al 2004).

Kort efter introduktionen af onlinebaserede spørgeskemaer så de første systemer til udvikling af spørgeskemaer rettet mod den almene bruger dagens lys. Dette lagde muligheden for at udforme spørgeskemaer til besvarelse online i hænderne på helt almindelige mennesker, i hvad nogen har refereret til som *demokratisering af undersøgelsesværktøjer* (Couper et al 2008).

Tilgængeligheden af disse nye værktøjer har imidlertid vist sig at være en blandet fornøjelse, idet det dels er blevet lettere, billigere og hurtigere at foretage spørgeskemaundersøgelser og dels har medført, at antallet af spørgeskemaundersøgelser er steget markant, hvilket skaber større konkurrence om respondenterne. Dette især kan synes problematisk, i det at mange undersøgelser er afhængige af respondenternes selvselektion. I forbindelse med udbredelsen og stigningen i antallet af spørgeskemaer er der ligeledes opstået en forestilling om forringelse af spørgeskemaets kvalitet samt en antagelse af, at metoden frembringer ringere resultater (Couper et al 2008). Denne antagelse har især bund i at udformningen af spørgeskemaer tidligere var centreret om større undersøgelsesinstitutter og forbundet med større omkostninger. Disse spørgeskemaer var ofte mere genarbejdede og overholdte en længere række standarder,

som havde til formål at sikre kvaliteten heraf. Netop disse antagelser om forringelse i datakvalitet og pålidelig er blandt årsagerne til, at denne dataindsamlingsform er omdrejningspunkt for nærværende undersøgelse.

5 Problemstilling

Gennem de senere år, har især det webbaserede spørgeskema været blandt de foretrukne dataindsamlingsmetoder. Metoden tilbyder et langt mere økonomisk alternativ til de gængse undersøgelser, hvori det er nødvendigt at have langt flere ressourcer til rådighed til interviewere, porto eller telefonopkald. I forbindelse hermed har metoden ligeledes fået ry for at producere en ringere kvalitet af data og samtidigt generere et mere upålideligt resultat, end hvis undersøgelsen havde været foretaget med en anden metode. En række af de problemstillinger der med det online spørgeskema, vil blive gennemgået og diskuteret gennem opgaven – og der søges at etablere indsigt i om det er muligt, at *forbedre* det afkast der fås ved online spørgeskemaer.

I forbindelse med min bacheloropgave i 2012 udarbejdede jeg det metodiske fundament for en udvidet dataindsamlingsmetode, hvis formål er at tage hånd om nogle af de komplikationer der opleves i forbindelse med webbaserede undersøgelser. I mangel på bedre navn fik denne metode arbejdstitlen *Thompson Metoden*. Afsæt for metoden er dataindsamlingsfasen, hvor idéen er at registrering og indsamling af interaktionsdata kan åbne op for nye muligheder indenfor det webbaserede spørgeskema. Heriblandt muligheden for at kunne fremhæve nogle af de respondenttyper, der forventes at kunne have en uønsket effekt på datasættets validitet – hvilket uddybes i et senere afsnit. Den gennemgående problemstilling for opgaven er således at undersøge og illustrere hvorvidt at metoden evner, hvad den er blevet sat i verden for – for derved at kunne være i stand til at bidrage til den sociologiske undersøgelsessfære.

6 Problemfelt

For bedst muligt at skabe indsigt i de omkringliggende felter, der støder op til det webbaserede spørgeskema, har dette kapitel til formål at belyse nogle af de problemstillinger, der går på tværs af de forskellige indsamlingsmodes. Dernæst uddybes de unikke karaktertræk, som det online spørgeskema besidder – her særligt med fokus på

nogle af de mest omtalte problemstillinger og antagelser, der er knyttet til netop denne indsamlingsmode.

De forskellige indsamlingsmodes har i oprindelsen haft sine formål, som alle må formodes at skulle forbedre de muligheder, som denne undersøgelsesform tilbød. Dette implicerer, at spørgeskemaet som metode har nogle problemstillinger, der er unikke for de enkelte indsamlingsmodes, hvorimod andre problemer er gennemgående for metoden som helhed.

6.1 Spørgeskemaets komplikationer

Afsnittet er struktureret således at der først vil skabes indsigt i de forskellige indsamlingsmodes indenfor spørgeskemametoden, hvorefter at de mest gængse problemstillinger gennemgås.

6.2 Indsamlingsmodes

I spørgeskemaregi kan der differentieres mellem to overordnede former for besvarelser og indsamlingsmodes – henholdsvis *den personassisterede besvarelse* såsom ansigt-til-ansigt og telefoninterview og *den selvudfyldte besvarelse*, som er gennemført uden direkte indblanding af forsker eller interviewer (Thompson BA 2012).

Personassisterede besvarelser	Selvudfyldte besvarelser
Ansigt-til-ansigt interview	Papirskemabesvarelse <ul style="list-style-type: none">• Udleveret• Fremsendt• Selvprint
Telefonassisteret interview	Webbaseret spørgeskema
	SMS-baseret besvarelse ¹
	IVR Telefonbesvarelse ²

Tabel 6.1.1.1 – Oversigt over indsamlingsmodes

¹ Kun set anvendt i forbindelse med meget få spørgsmål – og vil derfor ikke indgå i diskussioner om dataindsamlingsmetoder i forskningsregi.

² Interactive Voice Response – Automatiseret telefonbesvarelse.

De personassisterede former kræver indblanding af en interviewer, som oplæser både spørgsmål og svarmuligheder. Implicit i denne form ligger ligeledes, at intervieweren ofte kan motivere respondenterne til gennemførelse af spørgeskema samt påvirke tempoet for besvarelsen (De Vaus 2001).

Anerledes opleves det ved de selvudfyldte besvarelsesformer, hvori respondenterne oftest selv ansvarliggøres for både tempo og motivation til gennemførelsen af skemaet. Dette betyder, at det er muligt at besvare skemaet på den tid, der føles passende for den enkelte respondent – uanset om de ønsker at være grundige eller blot gennemføre besvarelsesprocessen hurtigst muligt.

Begge forgreninger har sine unikke problemstillinger, hvilket kan imødegås gennem design eller efterbehandling af respondentbesvarelser. Ydermere har de to forgreninger en række nævneværdige forskelle og ligheder.

Der findes en række forskelligheder mellem de auditive og visuelle former, som udmønter sig i forskellige respondenttendenser og besvarelsesmønstre, hvilket vil blive beskrevet senere i projektet. Hvor de selvudfyldte skemaer primært er visuelle, idet respondenterne selv læser indholdet i spørgeskemaet, er de personassisterede besvarelser modsat auditive, idet en interviewer som oftest vil stille både spørgsmål og svarmuligheder – dog med undtagelse af den selvudfyldte, auditive IVR-telefonbesvarelse, idet der her er tale om en båndet optagelse af en interviewer, der oplæser svar.

Fælles for forgreningerne er, at der søges at være på forkant samt imødekomme og omfavne ny teknologi, der på forskellig vis kan bidrage til spørgeskemafeltet (De Vaus 2001).

Spørgeskemafeltet har altid søgt at forny sig, ved at hurtigt at omfavne nye teknologier, hvilket med nye fordele ligeledes tilvejebringer nye problemstillinger. De problemstillinger, der er fælles for de forskellige spørgeskemamodes, er, at der i alle modes opleves problemstillinger i form af henholdsvis *design* og *respondentadfærd*.

I designet fastlægges struktur, indhold og metode for undersøgelsen, hvorfor der her er en risiko for at påvirke resultatet af den indsamlede data. Nogle af disse problemer og påvirkninger bør der fra forskerens side søges at tage højde for og gøre overvejelser

omkring. Anerledes er det hvis vi betragter problemstillinger som udelukkende beror sig på respondentadfærd, da det her kan være yderst vanskeligt – hvis overhovedet muligt – for forskeren at påvirke disse resultater.

6.3 Designkomplikationer

Dette afsnit har til formål at belyse nogle af de designmæssige problemstillinger, der kan opstå i forbindelse med spørgeskemaundersøgelser. Fokus her vil således være på potentielle komplikationer, som kan undgås ved brug af alternative undersøgelsesdesigns, herunder samplings samt overordnede spørgeskema- og spørgsmåls-designs.

6.3.1 Sampling og selektionsbias

Allerede ved valget af dataindsamlingsform er det muligt at påvirke resultatet af en undersøgelse. For eksempel kan valg af en bestemt indsamlingsmode medføre ekskludering af respondenter, hvis mening kunne være interessante og have indflydelse på undersøgelsens resultat (DST 2013). I en historisk kontekst kan nedenstående nævnes som eksempel herpå.

I 1936 foretog det daværende anerkendte blad *Literary Digest* en meningsmåling om udfaldet på det dengang forstående præsidentvalg mellem Alfred Landon og Franklin D. Roosevelt. Bladet forudså en jordskredssejr til Alfred Landon, om end det egentlige resultat udfaldt stik modsat med en samlet sejr på 61 % til Franklin D. Roosevelt.

Det var en af datidens største kvantitative undersøgelser med en proklameret samlet sample på over 10 millioner respondenter, hvoraf 2.4 millioner besvarede per telefon og postomdelte besvarelser (Freedman et al 2007). Respondentgruppen var samlet således, at den primært bestod af bladets abonnenter, indehavere af telefoner samt registrerede bilejere.

Ved *Literary Digests* sampling og valg af indsamlingsmetode var der aktivt skabt en selektionsbias, som bevirkede, at der blot blev etableret kontakt til og indhentet mening fra respondenter med størst økonomisk råderum (Freedman et al 2007). Ved at benytte telefon som indsamlingsmode havde *Literary Digest* ligeledes afskåret arbejderklassen, idet telefonen på daværende tidspunkt primært var forbeholdt de økonomisk velstillede.

Eksemplet illustrerer, at valg af indsamlingsmode og sampling kan medføre selektionsbias og herved et potentielt skævvredet billede på det undersøgte. Der opstår således dét, der kaldes *under-coverage*, hvor forskeren ikke er i stand til at afdække lige dele af den tilsigtede population grundet valgte samplingsmetode (Bethlehem 2010). Eksempelvis forudsættes det ved de selvudfyldte besvarelser, at respondenterne er i stand til at læse og skrive – dette dog med undtagelse af IVR-telefoninterview, hvor respondenterne elektronisk stilles spørgsmål og svarmuligheder og herefter kan afgive svar via tastetryk.

Undersøgelser påpeger, at op mod 16 % af den danske befolkning har særdeles svære læsevanskeligheder i en sådan grad, at disse vil være udfordret ved at læse ordinære tekster (OECD: PIAAC 2013). Til trods for dette betragtes læsevanskeligheder ikke umiddelbart som udbredt problemstilling i Danmark.

En væsentlig forskel mellem de assisterede og selv-udfyldte besvarelser er *self-selection* og *self-motivation* (Bethlehem 2010).

Ved *self-selection* forstås, at respondenterne selv skal vælge at deltage i undersøgelsen, uanset om denne er fundet på nettet, tilsendt via brev eller på anden vis omdelt. Ved det telefoniske interview, er det muligt for respondenterne at afvise deltagelse. Her har interviewererne dog mulighed for at søge at motivere eller på anden vis engagere respondenterne til deltagelse.

Når respondenterne først har påbegyndt spørgeskemaet, fremstår forskellen mellem de to overordnede forgreninger, idet der i en personassisteret besvarelse vil være en interviewer, som søger at motivere og *pace* respondenterne for herved at sikre gennemførelse. Er respondenterne derimod overladt til at besvare egenhændigt, må gennemførelsen af undersøgelsen foregå ved intern motivation – også kaldet *self-motivation*. Det er ikke til at vide hvor mange påbegyndte men ufuldendte besvarelser, der er havnet i papirkurven, eller hvor browservinduet blot er blevet lukket, idet kun et begrænset antal surveysystemer er i stand til at gemme delvise besvarelser, også kaldet *partial-data* (Thompson BA 2012). Dog har nogle af de mest udbredte systemer efterhånden adopteret delvise besvarelser som en mulighed i konstruktionen.

6.3.2 Som man spørger får man svar

I udarbejdelsen af spørgeskemaets spørgsmål er det centralt, at der både tages højde for formuleringer, rækkefølge og opsætning. Hvor nogle formuleringer kan virke decideret ledende, kan andre fremsat i en bestemt kontekst eller rækkefølge bevirke tendenser til bestemte svar. Dette påpeger således vigtigheden i at overveje den indholdsmæssige vinkel i forhold til design, herunder (De Vaus 2001)(Moore & Notz 2006)(Bryman 2008);

- Formulering af spørgsmålene
- Formulering af svarkategorierne
- Rækkefølgen af svarkategorier
- Mulighed for no-opinion svar
- Rækkefølgen af spørgsmål

Dette implicerer således, at det ikke blot er et spørgsmål om, hvad der spørges om men også hvordan, der spørges og i hvilket sammenhæng. Som følge heraf kan nogle af en undersøgelses udfald og resultater opstå allerede i designfasen, idet måden spørgsmålet formuleres og præsenteres kan bevirke tendenser til bestemte svar.

Således opstår en række faldgruber, der fra et designmæssigt perspektiv bør forsøges at tages højde for. I spørgeskemaregi vil størstedelen af spørgsmålene være lukkede, hvorfor de potentielle svarmuligheder på forhånd må overvejes og tage højde for. Det bør ligeledes medtænkes, hvorvidt der ønskes at præsentere respondenten for neutrale eller nulsvar. Sidstnævnte kan imidlertid bevirke, at respondenten i nogle indsamlingsformer vil tendere til *høj satisficing*³ ved at vælge nulsvaret (Krosnick 1991, 1996). Omvendt er der risiko for at fremtvinge en holdning hos respondenten, som de muligvis ikke allerede besidder, såfremt der ikke er muligt at vælge en neutral position.

6.3.3 Formulering af spørgsmål

Der findes et utal af konventioner, der bør overholdes i forbindelse med spørgsmålskonstruktionen og -formuleringen. Hvilket vil have indflydelse på hvordan respondenten opfatter og efterfølgende besvarer undersøgelsens spørgsmål (de Vaus 2001)(Bryman 2008). Dette afsnit har derfor til hensigt kort at gennemgå nogle af de mest almene konventioner

³ Hvor respondenten springer kognitiv genvej til at finde et acceptabelt svar, hvilket uddybes senere.

for formuleringer af spørgsmål – dog uden at dette hverken må betragtes som en guide eller håndbog hertil.

Ledende spørgsmål bør undgås, idet disse direkte kan påvirke respondentens holdning og opfattelse af det adspurgte. Denne spørgeform kan i højere grad end ellers afføde svar, som ikke afspejler respondentens reelle holdning men i stedet udtrykker en umiddelbart reaktion på adspurgte (de Vaus 2001). I forlængelse heraf bør negativt ladede spørgsmål ligeledes undgås, da også disse kan opfattes som ledende eller holdningsmanipulerende.

I den daglige tale er det ikke ualmindeligt at stille to eller flere spørgsmål samtidigt, hvilket i en ordinær dialog kan udredes, et såkaldt double-barrel spørgsmål (Bryman 2008). Anerledes vil dette i et spørgeskemaregi kræve mere komplekse svarmuligheder i forsøget på at indkapsle alle tænkelige svarsammensætninger. *Stil derfor kun ét spørgsmål ad gangen*, da det ellers kan give komplikationer i forhold til forståelse og besvarelse af det adspurgte.

Lange spørgsmål såvel som spørgsmål, der rummer tekniske termer, som forudsætter at respondenterne besidder en form for *ekspertviden*, bør ligeledes undgås.

Ved spørgsmål, hvori der ligger en indlejret antagelse om, at respondenterne besidder ekspertviden, risikeres det at fremtvinge en holdning hos respondenterne, som de ikke besad ved påbegyndelsen af undersøgelsen (Bryman 2008).

Kategorier kan bidrage til at skabe en klar linje gennem undersøgelsen. Disse kategorier bør imidlertid ikke overlappe eller flette ind i hinanden og kan hjælpe til at fastholde respondenterne i det givende *mind-set*, mens spørgeskemaet besvares.

6.3.4 Formulering af svarkategorier

I udarbejdelsen af de mulige svar, er det nødvendigt at man som forsker tager stilling til hvorvidt man er i stand til at opstille udtømmende svarkategorier, således at man vil være i stand til fyldestgørende at kunne anvende lukkede spørgsmål. Alternativt kan man være nødsaget til at benytte åbne spørgsmål, som giver respondenterne mulighed for at svare med fri tekst. Det i sig selv rejser dog problemstillingen, at respondenter i langt højere grad vil

have tendens til ikke at udfylde disse fyldestgørende, hvis overhovedet (Bryman 2008)(De Vaus 2001).

På samme måde som når der stilles spørgsmål er det nødvendigt at fremsætte svarmuligheder som er letforståelige, ikke rummer flere svar samtidigt og ikke mindst er relevante og knyttede til det stillede spørgsmål. Sidstnævnte kan lyde som en selvfølge, men det opleves oftere end man skulle forvente at spørgsmål og svarmuligheder ikke er direkte sammenhængende, og i stedet ændrer enten karakter eller betydning af spørgsmålet – som illustreret i senere afsnit.

6.4 Respondentadfærd

I besvarelsen af spørgeskemaer, er der risiko for at respondenter ubevidst besvarer eller har en bestemt adfærd. Eksempelvis kan der være tale om adfærd som nødvendigvis ikke er en nøjagtig parallel med hvad respondenterne i virkeligheden mener eller har erfaret, men i stedet en påtaget indstilling eller korrigeret erfaring som vil være mere social acceptabel eller ønskværdig. Respondenternes interaktion medfører således problemstillinger som man som forsker bør være opmærksom på.

6.4.1 Social desirability

I besvarelsen, ses det ofte at nogle respondenter vælger positioner og svar som stemmer overens med antagelsen om at vælge det korrekte, fra et normmæssigt standpunkt, fremfor at præsentere deres egentlige holdning – såkaldt *social desirability* (Kreuter et al 2008: 848).

Ligeledes vil respondenterne afgive svar som får dem til at fremstå mere konsistente end de egentlig er, for på den måde at imødekomme og underlægge sig de sociale normer der måtte være omkring emnet (Blasius & Thiessen 2012). Overordnet set betyder det at respondenterne ofte tegner et billede af sig selv og deres tilværelse, som er nærmere hvad der kan betragtes som værende *normalt* eller *socialt acceptabelt* end virkeligheden.

Dette betyder ligeledes at der i tale om outlier adfærd⁴, holdninger eller stigmatiserende⁵ emner, ofte vil opleves at respondenter vælger svar der lægger sig tættere på hvad der i et

⁴ Adfærd der adskiller sig markant fra hvad der kan betragtes som normen.

samfundsmæssigt perspektiv vil være betragtet som *normen*. De færreste respondenter vil anerkende adfærd som er betragtet som direkte socialt uacceptabelt, især hvis spørgsmålet er formuleret efter den direkte handling, uden kongruent sammenhænge. Det vil altså sige, at respondenter ofte vil være modvillige til at svare ærlige på et spørgsmål hvori de bliver spurgt om de har deltaget i decideret socialt uacceptabel adfærd – såsom hustruvold eller lignende, især hvis spørgsmålet er formuleret direkte eller synes inkriminerende.

Det er dog værd at pointere at der er i onlinebaserede spørgeskemaer generelt ses mindre tendens til social desirability (Fricker et al 2005), hvorved at det forventes at respondenterne overladt til sig selv, vil være mere åbne og ærlige om meninger og erfaringer. Man kan selvfølgelig sætte spørgsmålstejn ved, om dette er en generel udvikling der er knyttet til onlineadfærd som helhed, hvor der i den offentlige debat ofte har været draget fokus omkring en tendens til et *mere ærligt budskab* og hårdere tone, end man tidligere har set, som har udviklet sig på nettet.

6.4.2 Svarmønstre og satisficing

Der eksisterer flere antagelser forbundet med dataindsamling via spørgeskema. En af antagelserne er, at data indsamlet via spørgeskemaer varierer mellem forskellige modes og i nogle indsamlingsformer producerer data i *ringere* kvalitet i forhold til øvrige modes (Fricker et Al 2005)(Roster et al 2005). En anden antagelse er, at respondenter har større tendens til satisficing i besvarelsen på selvudfyldte spørgeskemaer. Herved menes, at respondenter i højere grad stiller sig tilfredse med svar, som ikke nødvendigvis er tilstrækkeligt dækkende (Krosnick 1991). I forlængelse heraf ses en mere kontroversiel antagelse om, at respondenter skyder kognitiv genvej for hurtigst muligt at finde et tilfredsstillende svar, hvilket kaldes satisficing. En tendens som der er svingende meninger om, hvorvidt det er udbredt i højere eller lavere grad i onlineundersøgelser (Fricker et al 2005)(Roster et al 2005).

Satisficing handler således om de kognitive processer, som respondenter gennemløber i forbindelse med besvarelse af spørgeskemaer. Et normalt kognitivt forløb i besvarelsen af spørgsmål, sker ved følgende fire trin, hvor respondenter (Krosnick 1991)(Krosnick et al 1996);

⁵ (Goffman 2001)

1. Overvejer den præcise mening ved spørgsmålet;
2. Søger internt i egen hukommelse efter relevant viden og information;
3. Foretager en vurdering på baggrund af internt fundne viden;
4. Afgiver svar på baggrund af foregående vurdering.

I tilfælde af satisficing er der i stedet tale om, at respondenterne springer henover nogle af de kognitive trin i processen, for i stedet blot at slå sig til tåls med en svarmulighed eller blot vælge en svarmulighed, for derefter at fabrikere enighed heri. Forstået på den måde at respondenterne siger at de er enige, uden nødvendigvis at være det. Der er i den forstand tale om, at respondenterne kan gennemgå svag eller stærk satisficing, hvor graden af hvor meget de kognitivt springer over varierer. I undersøgelser på tværs af dataindsamlingsmodes er der ligeledes påvist tendens til, at respondenterne i forbindelse med online spørgeskema er mere tilbøjelige til at vælge svarmuligheden '*ved ikke*', samt har en mindre spredning på spørgsmål med samme skala. Ligeledes opleves der tendens til et større antal manglende besvarelse ved de online spørgeskemaer i forhold til øvrige metoder (Herweegh og Loosveldt 2008).

Det må således forventes, at satisficing i højere grad forekommer i forbindelse med dataindsamlings-scenarier, hvor respondenterne egenrådigt er ansvarlig for besvarelsesprocessen, og derved i mindre grad i forbindelse med personassisterede spørgeundersøgelser. Hvilket ligeledes illustreres i sammenligningen af web og telefonassisterede spørgeskemaer (Roster et al 2005), hvoraf det fremgik, at spredningen via internettet var mindre end per telefon. Modsat argumenteres der ligeledes for, at respondenterne tenderer til at udvise lavere satisficing i kraft af muligheden for at besvare undersøgelsen i vante rammer og eget tempo, hvorfor den kognitive byrde lettes (Fricker et al 2005). Dog tages der ved denne antagelse ikke højde for de potentielle udefrakommende elementer, som respondenterne kan interagere med – både på og udenom skærmen. Her kan alt fra familie til anden aktivitet på sociale medier, tv eller lignende være forstyrrende for den kognitive proces og kan derfor bevirke, at der opstår en højere grad af satisficing hos respondenterne.

7 Undersøgelsens fokus

Vi befinder os i skrivende stund i den sidste af de tidligere nævnte epoker, hvor internettet og det webbaserede spørgeskema så dagens lys og udbredtes. Denne udvikling tilvejebragte en demokratisering af udviklingsværktøjerne til spørgeskemaproduktion og gjorde det derudover muligt for alle med internetadgang hurtigt at konstruere og omdele et spørgeskema. Ligesom ved introduktionen af telefonen som gangbar indsamlingsmode opstod der i forbindelse med det webbaserede spørgeskema nye spekulation og problemstillinger.

Én af disse spekulationer er gået på, at den webbaserede dataindsamling producerer ringere data eller forøger selektions- og samplingsbias, idet respondenterne udsættes (eller eksponeres for) for et væld af spørgeskemaer.

7.1 Det webbaserede spørgeskema

Af de forskellige spørgeskemametoder har den onlinebaserede overhalet de fleste andre modes indenom, hvad angår kvantitet (Fricker et al 2005). Udviklingen i indsamlingsmåden har fra et teknisk perspektiv dog været fanget i noget lignende et standstill. Det webbaserede spørgeskema ligner i høj grad de traditionelle metoder og lider ligeledes under de samme problemstillinger. De senere års udvikling på internettet, har vist hvad der er muligt at indsamle af data, her især med henblik på den interaktion brugerne har med deres computer, tablet eller mobiltelefon. Kigger man i stedet på den webbaserede spørgeskema, ses den udvikling der har oplevedes i andre sfærer på nettet, ikke på samme vis har påvirket dette. På den baggrund kan det konkluderes at der er tale om en dataindsamlingsform med udviklingspotentialer. Men hvad indkapsler egentlig det webbaserede spørgeskema, og hvilke potentialer og udfordringer ligger herved? Det er nogle af de spørgsmål, der søges afklaret med følgende afsnit.

7.1.1 Fordele ved webbaserede spørgeskemaer

Hvor det førhen kunne være både bekosteligt og tidskrævende at foretage selve dataindsamlingen i en kvalitativ undersøgelse, er det i de senere år blevet langt mere almindeligt at se undersøgelser som helt eller delvist foretager dataindsamlingen online.

De onlinebaserede undersøgelsers popularitet skyldes både økonomisk fordelagtighed såvel som nye potentialer såsom:

- **Hurtigt data turnaround:** For forskeren betyder dette, at det er muligt hurtigt at få både distribueret spørgeskema og indsamlet data (Kenett & Salini 2011: 94 – 96). Spørgeskemaet kræver ingen interaktion mellem interviewer og respondent, fremstår lettere tilgængeligt og vil ved undersøgelsesscenarier med etablerede respondentbaser ofte vare mellem 48 til 72 timer (Harris 1997). Ligeledes er det muligt for respondenten selv at tilrettelægge, hvornår spørgeskemaet besvares, idet der ikke længere forudsættes interaktion med en interviewer.
- **Svarraten:** Online spørgeskemaer med fastlagte respondentbaser vil oftest have højere svarrate end i de fleste andre dataindsamlingsmodes (Kennet & Salini 2011: 94 – 96). Anerledes gør det sig gældende ved undersøgelser, der ikke på forhånd har en etableret respondentbase. Disse lider imidlertid af kritisk lave svarrater, hvilket ligeledes diskuteres senere.
- **Multimedieintegration:** Forskeren har mulighed for at vedhæfte både billeder, lyd og video til spørgeskemaet, hvilket bevirker, at respondenten mere ubesværet kan besvare spørgeskemaet uden samtidigt at anvende andre redskaber (Kennet & Salini 2011: 95).
- **Conditionals:** Spørgeskemaerne kan struktureres således, at opsætning og spørgsmål der stilles, opsættes i såkaldte *conditionals* eller betingelser, hvor bestemte spørgsmål kun stilles eller svarmuligheder vises/skjules såfremt at nogle bestemte kriterier er mødt (Kennet & Salini 2011: 96). Disse udvidede kapaciteter giver ligeledes mulighed for at strukturere mere komplekse undersøgelser, som benytter sig af såkaldt *piping* – forstået ved direkte omdirigering af respondenter, hvorved respondenten ledes udenom spørgsmål, der er irrelevante på baggrund af de tidligere besvarelser. Dette betyder, at respondenten undgår unødigt læsning eller kognitiv kobling af emner⁶, idet der helt kan undlades at vise spørgsmålsbatterier som ikke her forekommer relevante for respondenten.
- **Geografi:** Udfordringer forbundet med fysisk afstand reduceres, idet internettet muliggør en mere ubesværet distribution af spørgeskemaer til respondenter, som for eksempel bor i yderområder eller udlandet. Såvel som de geografiske barriere

⁶ Når respondenten forsøger at danne et helhedsbillede af spørgeskemaet kan det være forstyrrende at læse spørgsmål som allerede har syntes besvaret eller som synes uvedkommende eller irrelevante.

mindskes, er det samtidig muligt at omdele spørgeskemaet i flere lande og på flere sprog uden at dette nødvendigvis kræver flere ressourcer for at gennemføre. Det må dog pointeres, at der i flersprogede undersøgelser ofte vil skulle afsættes mere tid og flere ressourcer på at sikre ensartede forståelser og fortolkninger af spørgsmål på tværs af sprog og lande.

- **Anonymitet:** Én af de væsentligste fordele ved det onlinebaserede spørgeskema er muligheden for personlig *anonymitet*. Respondenter er vanskelige identificere på baggrund af den indsamlede data og vil ofte føle sig mindre *udstillet*, da det er muligt at besvare spørgeskemaet i vante rammer og uden indblanding af forsker, personlige brevforsendelser eller andet identitetsorienteret data. Den oplevede anonymiteten kan bevirke, at respondenter er mere villig til at besvare personlige og følsomme spørgsmål (Kenett & Salini 2011: 95). Det må derfor antages, at følelsen af anonymitet er vigtig i forhold til at fremme ærlige svar fra respondenteren – især i mere kompromitterende emner.

7.1.2 Komplikationer med webbaserede spørgeskemaer

Der er opstået en række problemstillinger i forbindelse med, at det online spørgeskema blev tilgængeligt og let anvendeligt for alle med adgang til en computer.

Undersøgelser, der ikke i forvejen har en etableret respondentbase, lider, som tidligere omtalt, af kritisk lave svarrater. Derudover kan forekomme vanskeligt at akkumulere et tilstrækkeligt antal respondenter for at skabe et repræsentativt grundlag for undersøgelsen.

Dernæst kommer, at årsagerne til de lave svarrater kan skyldes flere parametre. Alene det stigende antal spørgeskemaer har bevirket, at der er opstået en art *survey fatigue* eller spørgeskematræthed – her forstået ved at kvantiteten af spørgeskemaer medfører en manglende villighed til at deltage (Couper et al 2008). Som nævnt sker det i kraft af demokratiseringen af redskaberne til spørgeskemaudviklingen, som ikke længere blot redskaber for forskningsinstitutter, men nu i stedet er hvermandseje. Dette betyder, at der i højere grad konkurreres om respondenternes deltagelse – både blandt virksomheder der foretager kundeundersøgelser, blandt foreninger der ønsker indsigt i medlemmernes ønsker og vaner samt i forbindelse med studieopgaver og lignende.

Ved respondentindsamlingen ses en række komplikationer, som vanskeliggør akkumuleringen af besvarelser særligt i forhold til *sted*, *tid* og *interesse*. Herved forstås nødvendigheden i, at respondenten er til stede i det forum, hvor undersøgelsen uddeles. Hvorvidt at der er tale om et forum, sociale medier såsom Facebook og Twitter er det nødvendigt at respondenten anvender mediet i et tidsrum hvor det er sandsynligt at det delte opslag bliver vist for dem. Slutteligt skal respondenten være aktivt interesseret i at deltage i undersøgelsen, idet det påhviler dem at *vælge sig selv/ selvseleksion* til at deltage i undersøgelsen (Kennet & Salini 2012).

I forsøg på at rekruttere respondenter anvendes der i forbindelse med nogle spørgeskemaundersøgelser belønningsmetoder, hvor respondenten søges motiveret ved konkurrencer, gevinster eller bonusordninger. Heri ligger ligeledes muligheden for, at nogle respondenter deltager udelukkende på baggrund af den udlovede gevinst eller lod til deltagelse i konkurrencen samt risikoen for, at nogle af disse respondenter vil søge at besvare skemaet flere gange for derved at øge egen adgang til belønningen. Dette kan både resultere i tendens til stærk satisficing samt situationer, hvor respondenterne ikke tager stilling til adspurgte men blot besvarer hurtigst muligt (Kennet & Salini 2012).

Hvorvidt dette skyldes mængden af spørgeskemaer eller varierende kvalitet vides ikke. Ikke desto mindre har det onlinebaserede spørgeskema fået ry for at producere dårligere og mindre pålideligt data (COPAN 2014). Som følge heraf er der opstået et undersøgelsesfelt for sig selv med det formål dels at undersøge kvaliteten af data, der indsamles i online spørgeskemaer, og dels at sammenligne disse med andre indsamlingsmodes med henblik på at belyse forskelle og eventuelle problemstillinger metoderne imellem. Herved er det muligt at opnå indsigt i, hvilke betydninger indsamlingsformen har for undersøgelsens resultater. Der kan i forbindelse hermed nævnes en række undersøgelser, som har til formål at undersøge datakvaliteten og respondentadfærd i onlinebaserede spørgeskemaer sammenlignet med det personassisterede telefoninterview (Fricker et al 2005)(Roster et al 2004). Såvel som undersøgelser der har til formål at belyse respondentadfærd på tværs af onlinebaserede spørgeskemaer og nogle der er udleveret i papirform, ved personlig overlevering (Fricker et al 2005).

7.1.3 Sampling og selektionsbias i onlinebaserede spørgeskemaer

Ved at anvende online spørgeskema som indsamlingsmetode accepteres således også en hvis forudsat bias.

Biasen opstår især som følge af, at populationen på internettet er skævvredet, da det ikke er alle, der har adgang til computer, internet eller opholder sig steder, hvor spørgeskemaet omdeles. Der ses ligeledes en skæv alders- og uddannelsesfordeling, idet nogle grupperinger er overrepræsenterede, mens andre stærkt underrepræsenteres (Bethlehem 2010). Herved skabes basis for *under-coverage*, hvorved undersøgelsens population ikke afspejler den reelle fordeling på den population, der egentlig søges undersøgt.

Iagtages den potentielle danske population på internettet ses det, at jo ældre populationen bliver, desto færre har adgang og/eller erfaring med at benytte internettet. Dertil kommer, at under 50% af danskere over 70 år har stiftet bekendtskab med internettet og kun over halvdelen heraf har været på internettet indenfor den seneste uge (Ældre Sagen 2014)(DST 2013). Dog ses en stigende tendens til, at flere ældre får både adgang og erfaring med internettet – også på mere hyppig basis (Ældre Sagen 2014). Dette betyder imidlertid, at nogle af de selektions- og bias-problemstillinger må forventes mindsket over tid i takt med, at respondentgrupperne bliver ældre, hvorved de daglige it-vaner og -erfaringer forandres (Bethlehem 2010)(Couper 2011).

I skrivende stund er der dog fortsat risiko for skævvridninger. Ønskes der for eksempel en analyse af digitalisering af det offentlige udelukkende via online spørgeskemaer, vil de grupperinger med lidt eller ingen interneterfaring potentielt ekskluderes fra undersøgelsen.

Ligeledes kan der opstå bias ved påhvile respondenterne at *vælge sig selv*, til at deltage i undersøgelsen – hvad der også kaldes selvselektion. Forstået på den måde at respondenterne skal både kunne få adgang til en computer, internet og samtidigt se og vælge at deltage i undersøgelsen. Man forværrer således sandsynligheden for at opnå en repræsentativ sample.

7.1.4 Metodiske beskrivelser af webbaserede spørgeskemaer

Websurveys eller onlinespørgeskemaer er i sig selv en bred beskrivelse, som ingelunde afdækker eller fortæller de anvendte metoder. Normalvis vil man i undersøgelser referere til den anvendte metode ud fra en beskrivende devise, som kan være sigende for selve undersøgelsesformen.

I de fleste studier er det muligt at kortlægge en beskrivende metode såsom RDD⁷ telefonsurvey, der både er beskrivende for måden, hvorpå samplingen er foretaget såvel som mediet, der anvendes til dataindsamling (Couper & Miller 2008: 831 – 835).

Der findes ingen konventioner for beskrivelse af webbaserede surveys. Dette kan både vanskeliggøre læserens indsigtsmuligheder i forbindelse med undersøgelsen og dels gør det mere besværligt at vurdere kvaliteten af undersøgelsesprocessen (Couper & Miller 2008: 832). Det er derfor ikke muligt ud fra metodebeskrivelsen at dechifrere, hvilken sampling der er anvendt eller hvilke, data der er indsamlet i processen. Dette er dog en mindre og sideløbende diskussion, hvis introduktion i dette afsnit blot har til formål at skabe indblik i et felt, som både for de som sidder med det, såvel som udefra, kan fremstå som en smule broget.

7.1.5 Respondenttyper i onlinebaserede spørgeskemaer

Da det elektroniske spørgeskema ikke involverer en interviewer, kan der opstå nye situationer og mønstre blandt respondenter, hvilket bevirker, at der kan være tale om nye respondenttyper. I forlængelse heraf er der her risiko for skævvridninger af data. Dette kan eksempelvis komme til udtryk ved at de afgivne svar foregår i hast og uden nøje overvejelser ved at respondenter søger at fremstille sig selv mere lukrativt og påtage sig holdninger og adfærdsmønstre, eller ved at samme respondent deltager flere gange med henblik på at erhverve sig diverse frynsegoder.

Disse beskrevne respondenttyper refereres ofte til som *cheaters*, *speeders* og *repeaters* (Kenett & Salini 2012: 96 – 100).

⁷ Random digit dialing – hvor telefonnumre genereres tilfældigt. Er ikke lovligt i alle lande.

Cheaters er respondenter, der oftest deltager af andet incitament end interesse i undersøgelsen eller i at give sin holdning til kende samt respondenter, der ikke tager stilling eller forholder sig til adspurgte. Herved kan det være diskutabelt, hvorvidt der er tale om satisficing, idet respondenteren ikke stiller sig tilfreds med et givent svar men blot vælger tilfældigt.

Speeders er respondenter hos hvem, der formodes at forekomme en høj grad af satisficing, da disse blot overfladisk forholder sig til spørgsmålene med henblik på en gennemførelse af besvarelsen på kortest mulige tid (Kennet & Salini 2012). Mange af disse ville kunne identificeres på baggrund af gennemførelsestid, som vil blive registreret i de fleste surveysystemer registreres. Dog er det væsentligt at have i mente, at det kan forekomme vanskeligt at kortlægge præcist, hvor længe det vil tage respondenter at gennemføre spørgeskemaet.

Repeaters er respondenter, der besvarer spørgsmål på samme eller lignende måde, idet disse ofte vil vælge ensartede svarkategorier. Eksempelvis kan der her være tale om respondenter, som generelt placerer sit svar omkring midten på en likert skala.

8 Thompson Metoden

Med afsæt i ovenstående udarbejdede jeg i forbindelse med min bacheloropgave i sociologi et metodisk fundament med henblik på at forbedre og afhjælpe nogle af disse udfordringer, der kan forekomme i anvendelsen af online spørgeskemaer.

Metoden opstod som en reaktion på en længere række af de førnævnte problemstillinger og med et mål om at forbedre validitet, reproducerbarhed og indsigt i anvendelsen af webbaserede spørgeskemaer. Kortlagt i min bacheloropgave fra 2012, som vil fungere som den primære reference i dette afsnit.

Metoden kan betragtes som et samspil mellem sociologi og programmering, hvorigennem det søges at forbedre og udbygge det datamæssige udbytte ved at indsamle information som er tilgængelig i respondenterens interaktion med det webbaserede spørgeskema. Herved er der tale om data, der rækker udover dét, der indsamles under normale omstændigheder, for

på den måde at søge at bringe metoden tættere på det potentielle udbytte, der eksisterer i de personassisterede besvarelsesformer.

Ved de selvudfyldte spørgeskemaer er det på nuværende tidspunkt ikke muligt for afsenderen at vurdere respondentens reaktion såsom, hvorvidt der hastes eller hersker usikkerhed i forbindelse med besvarelsen. Den grundlæggende idé bag metoden er at indsamle og anvende data, der er direkte relateret til respondentens interaktion med spørgeskemaet. De indsamlede data formodes således at kunne anvendes til at belyse forskellige aspekter af både respondentbesvarelsen, viden om respondenterne såvel som indsigt i selve spørgeskemaets struktur og design.

Formålet med projektet er at etablere indsigt i den grundlæggende metodes anvendelighed, men i metoden ligger ligeledes et potentiale i indsamling af data. Omdrejningspunktet for undersøgelsen er her afgrænset til at have fokus på metodens centrale elementer og kan således reduceres til følgende tre *tertiære variable*; Tid, svarskeft og progression.

8.1 Datastruktur og udbytte

I forbindelse med kvantitative metoder er det muligt at rangere efter indsamlingsformens datamæssige potentiale, ved eksempelvis at sammenligne et onlinebaseret survey med en personassisteret besvarelse via telefon, via brev – hvilket følger i næste afsnit.

Data der indsamles kan kategoriseres som primære, sekundære og terciære.

Primære data er svar, som afgives i spørgeskemaet. Disse kan imidlertid variere på tværs af de forskellige indsamlingsmodes alt efter, om der opstår satisficing, social desirability eller holdningsneutralitet⁸ (Couper et al 2008)(Fricker et al 2005).

Sekundære data er svar, som respondenterne aktivt videregiver, og information som kan inkluderes i det endelige datasæt. Her kan der være tale om besvarelser, som ikke er at finde i det etablerede spørgeskema men derimod er blevet tilføjet af respondenterne for eksempel som en note i margin på et postomdelt spørgeskema eller oplyst/oplevet af en interviewer via telefon. Disse data kan således medvirke til at belyse fejl og mangler i svarmuligheder, spørgsmål eller design. Om det er sandsynligt at forvente at disse data vil

⁸ Når respondenterne besvarelse er centreret omkring et givent punkt med en lavere spredning.

blive anvendt eller noteret af en eventuel interviewer, er selvfølgelig svært at vide. Det kan blot påpeges at der eksisterer mulighed for at indsamle disse data i nogle modes.

Tertiære data er denne metodes omdrejningspunkt og er data, som respondenterne ubevidst videregiver. Der er her tale om en del af den interaktion, respondenterne har med spørgeskemaet, hvorved information som rækker ud over selve besvarelsen latent videregives via parametre såsom den tid, det tager respondenterne at besvare et givent spørgsmål, skiftende svarvalg og andre ikke direkte kontekstrelaterede data. Omdrejningspunktet er således de variable, som beretter om besvarelsesforløbet og herved giver mulighed for at gengive respondenternes interaktion med skemaet.

8.2 Rangering af dataindsamlingsmodes

Da de mest almindelige dataindsamlingsformer kan rangeres efter potentielle dataudbytte, er det muligt at skabe et billede af og et overblik over de forskellige modes mulige afkast og potentialer.

Mode	Primære data	Sekundære data	Tertiære data
Ansigt-til-ansigt	<i>x</i>	<i>x</i>	<i>x</i>
Telefonassisteret	<i>x</i>	<i>x</i>	<i>x</i>
Papirform (Selvudfyldt)	<i>x</i>	<i>x</i>	
Web / IVR (Selvudfyldt)	<i>x</i>		

Tabel 10.2.1 – Dataafkast i indsamlingsmodes

Øverst i ovenstående tabel findes de personassisterede dataindsamlingsformer såsom besvarelser foretaget ansigt til ansigt eller telefonisk. I disse undersøgelser er det muligt for *intervieweren* både at indsamle primære, sekundære og tertiære data.

Dog kan der i de personassisterede modes opstå komplikationer, som kan være vanskelige at kompensere for – med mindre der er tale om et *mixed-modes* studie. Sådanne komplikationer kan være påvirkninger såsom reaktivitet og social desirability, hvor respondenterne bestræber sig på at fremstå mere kongruent og sammenhængende end de reelt er (Roster et al 2005). Ligeledes kan inklusionen af interviewerens i dataindsamlingen både fungere som motivation og som en *stressfaktor*. Da interviewerens som oftest er trænet

til at guide respondenterne videre og derved medvirke til at sætte gennemførelsens tempo, er der således risiko for, at respondenterne udviser tendens til blot at vælge svar, som muligvis ikke repræsenterer deres egentlige holdning (Fricker et al 2005).

De nederste to modes af ovenstående tabel er de selvudfyldte besvarelser, hvor respondenterne gennemfører spørgeskemaet uden assistance fra en interviewer såsom omdelte spørgeskemaer samt IVR- og onlinebaserede surveys. Hvoraf skemaet i papirformen tillader sekundære data, som kommentarer i margin, vil der i både IVR telefonundersøgelser og online surveys normalvis kun indsamles primær data, med undtagelse af muligheden for at registrere den samlede varighed af besvarelsen, hvilket betyder at indsamlingsformen kun har kapaciteten til at indsamle primær og, minimal, tertiær data

På baggrund af denne rangering af dataindsamlingsmetoder kan det således konkluderes, at webbaserede surveys er den dataindsamlingsmetode, der tilvejebringer mindst data. Dette fremstår paradoksalt og rejser en interessant problemstilling, da én af de mest udbredte indsamlingsformer samtidigt er det mode, der giver mindst indsigt i selve besvarelsen. Det betyder at der i de fleste tilfælde hvor selvudfyldte besvarelser anvendes, kun er tale om at indsamle respondenterens direkte besvarelse. Selvom det er muligt for forskere at støtte sig til såkaldte *mixed modes*, der verificerer anvendte data ved hjælp af blandede dataindsamlingsformer, ses det ofte, at undersøgelser kun anvender det online spørgeskema som indsamlingsmetode.

8.3 Tertiære data

Da formålet med dette speciale er at undersøge Thompson Metodens anvendelighed afgrænses der her til at fokusere på metodens mest centrale elementer, herunder tid, svarskifte og progression. Således bliver spørgsmålet om anvendeligheden i metodens grundform omdrejningspunktet, hvorfor der her ikke vil skabes blik for perspektiver såsom de øvrige dataindsamlingsmuligheder og disse implicite potentialer i at skabe indflydelse og indblik i eventuelle besvarelsesforløb.

Ved at tilvælge netop disse tre variable som omdrejningspunkt for projektet opstår en form for *projicering*⁹. Dette kommer til udtryk ved antagelsen om, at netop disse variable rummer sigende information om respondenternes interaktion med spørgeskemaet, og som potentielt kan skabe resultater - der kunne være tale om at de variable der egentlig vil have værdi for undersøgelsen ville være nogle helt andre. Dette kan sammenlignes med når projicering forekommer i observationsstudier, hvorved at man som observatør søger at iagttage en bestemt adfærd, hvilket er årsagen til at netop den adfærd observeres.

8.3.1 Tid

Tid er et af de mest grundlæggende elementer i spørgeskemaprocessen. Tiden der bruges på hvert enkelt spørgsmål, på hver kategori og på hele spørgeskemaet kan berette om både den enkelte respondent, hele grupper såvel som spørgeskemaet. Anvendelse af tidsparameteret kan danne et billede af, hvorvidt respondenterne har besvaret spørgsmålene hurtigst muligt, eller om respondenterne som følge af enkelte spørgsmål er stoppet op eller har brugt tænkepause.

Tid fungerer i Thompson Metoden som kronologisk markør såvel som kronometrisk måleenhed for events og hændelser i respondentens besvarelse. Dette medvirker således til at skabe en holistisk indsigt i respondentens besvarelsesproces. Dog beretter tidsfaktoren blot om én del af undersøgelsen, hvorfor dets potentiale først ses i samspil med andre variabler.

8.3.2 Svarskift

Svarskift forekommer, når en respondent vælger at ændre sit svar på et givent spørgsmål. Eksempelvis når en respondent skiftevis væger mellem og markerer forskellige svarmuligheder. Dette kan således skabe indblik i, hvordan respondenterne har valgt sit svar, og om der er incitament for at undersøge selve spørgsmålet eller de givne svarmuligheder nærmere.

Ydermere kan svarskift skabe blik for emner, der forekommer kontroversielle eller stigmatiserende – emner hvor respondenterne kan opleve usikkerhed omkring sit ønske til at fremsætte visse svar. Dette kan blandt andet komme til udtryk ved, at respondenterne i et

⁹ Mangler ordentlig kilde – kommer.

givent spørgsmål skifter mellem svarmuligheder og en neutral værdi, såsom 'Ved ikke' eller 'Ønskes ikke oplyst'. Dette scenarie er dog ikke nødvendigvis forbeholdt kontroversielle eller personlige emner men kan ligeledes vedrøre emner, som respondenterne ikke besidder tilstrækkelig viden om eller holdning til (Krosnick 1991).

8.3.3 Progression

Progression udtrykker respondentens bevægelser mellem forskellige sider af spørgeskemaet såsom registrering af kronologisk eller tilbagegående sideskifte. Herved bliver det muligt at registrere, hvor respondenterne bevæger sig fra og til såvel som hvor længe, de opholder sig på de mellemliggende sider. Ligeledes kan det eksempelvis iagttages, hvis en respondent bevæger sig tilbage fra side otte til side fire, har kort og mellemliggende ophold på side seks for dernæst at bladde hurtigt over de resterende sider.

Progressionsvariabel forudsætter således, at spørgeskemaet er struktureret i sider bestående af spørgsmål. Der kan således ikke være tale om en decideret progression der i en enkeltsides-undersøgelse – udover start og sluttidspunkt.

Denne information giver således ikke en endegyldig indsigt i processen men kan henlede forskerens opmærksomhed på givende sider og spørgsmål for på den måde at forsøge at analysere mulige årsagsforklaringer på progressionen.

8.3.4 Samspil

Samspillet mellem disse tre nøglevariabler gør det muligt for forskeren at reducere respondentens interaktion med spørgeskemaet til *events*, bestående af svarskifte og progression. Disse events beretter om, hvilket spørgsmål respondenterne har været i færd med at besvare, hvilke svarmuligheder de har valgt og hvornår, samt hvorvidt respondenterne bevæger sig frem eller tilbage i undersøgelsen – hvis de eksempelvis på en given side, bevæger sig flere sider tilbage og hvor længe de opholder sig på hver side efterfølgende.

De tertiære variabler skaber i dette samspil grundlag for at opnå indsigt i respondentens besvarelsesforløb og interaktion med spørgeskemaet. Der kan på baggrund heraf tegnes en tidslinje af respondentens forløb, som kan analyseres og fortolkes individuelt, gruppevis

eller på tværs af hele samplen, hvorved det er muligt for forskeren at blive opmærksom på gennemgående tendenser og adfærd i respondenternes besvarelser. På baggrund af disse tidslinjer vil det ligeledes være muligt at belyse respondenters meantiden for besvarelser af både de enkelte spørgsmål og hele spørgsmålsbatterier/sider samt svarskeftendenser.

Dette skaber for muligheden for at undersøge en række forskellige sammenhænge og adfærdsmønstre i besvarelsen – såsom en antagelse af, at yngre populationsgrupper bruger kortere tid på besvarelsen af selve skemaet, imens ældre vil udvise større usikkerhed eller være mere grundige og derved både anvende længere tid og flere svarskeft i forløbet.

Da forskeren herved kan blive opmærksom på adfærdstendenser i spørgeskemaets besvarelser, er det ligeledes muligt at undersøge selve spørgeskemaet i forhold til formuleringer og design. Dette kan foregå som en undersøgelse i sig selv eller sideløbende med en ordinær undersøgelse med henblik på at forbedre validitet og reproducerbarhed af spørgeskemaer, hvorved der åbnes op for en række af information, som ikke tidligere har været tilgængelig.

Hvor det tidligere kun var muligt at undersøge respondentens endelige svar, åbnes der nu op for at undersøge forløbet op til. Såfremt en undersøgelse gentages med afsæt i *Thompson Metoden*, vil det være muligt at analysere forløbene nærmere.

Det kan antages, at en respondent besvarer et spørgsmål med *Ved ikke*. I nuværende scenarier vil der være tale om stærk satisficing men med indsigt i forløbet, er det muligt at iagttage, hvorvidt respondenter har valgt andre svarmuligheder før det endelige valg af *Ved ikke*. Dette kan berette om elementer, som forskeren bør være opmærksom på eller bekræfte tendenser til satisficing igennem illustration af respondentens besvarelsesforløb.

Dette rummer både mange muligheder for kombinationer af de givende data samt fortolkninger af de events, der forekommer i respondenternes forløb. Nogle af hvilke projekter har til formål at kortlægge gennem analysen.

8.3.5 Forskerens evalueringsevne

Det er vigtigt at understrege at intentionen med metoden er at fremhæve besvarelser samt fejl og mangler i spørgeskemaet, således at forskeren lettere kan blive gjort opmærksom på

de forskellige facetter i datasættet. Det vil altså sige, at det ikke er intentionen at omgå forskerens evalueringsevne gennem automatisering – men i stedet at fremme muligheden for at blive gjort opmærksom på hændelser i spørgeskemabesvarelsen som ville kunne have effekt eller interesse i forhold til det undersøgte.

9 Formål

Formålet med projektet er at kaste lys over, hvorvidt at *Thompson Metoden* evner at modarbejde eller løse nogle af de problemstillinger som opleves i anvendelsen af webbaserede spørgeskemaer. Som nævnt, har denne indsamlingsmode haft ry for at producere ringere og mindre pålidelig data. Hvilket især har bund i vanskelighederne ved at akkumulere respondentbaser, i kraft af respondenternes selvseleksion samt den reaktion der blev affødt heraf, hvor respondenter forsøges at blive lokket til deltagelse ved udlovning af præmier.

Antagelsen bag metoden beror sig på, at et større datamæssigt afkast, i form af tertiære data, der indsamles i respondentens interaktion med spørgeskemaet, vil kunne medvirke til at forbedre pålideligheden af de indsamlede besvarelser. Det overordnede formål er således at teste, hvorvidt at der er tale om et plausibelt scenarie og hvorvidt at de data der indsamles giver anledning til at kunne drage sådanne konklusioner.

Projektet fungerer som en illustration af metodens potentiale, hvorved at der ikke lægges fokus på indholdsmæssig analyse af respondenternes besvarelser, men i stedet på de tertiære data der indsamles. Der er derfor tale om et projekt, hvor intentionen er at gribe fat i et af sociologiens redskaber, for på den vis at søge at forbedre det og de muligheder de giver.

10 Problemformulering

For at kunne undersøge hvorvidt at *Thompson Metoden* evner at belyse eller løse de problemstillinger der er i det webbaserede spørgeskema, er det nødvendigt at opstille en overordnet problemstilling. Med en udvidet dataindsamlingsmetode, vil der selvsagt akkumuleres en større mængde data – hvorved at det også er nødvendigt at undersøge, hvilke implikationer dette kunne have for forskeren der anvender metoden. Formuleringen af problemformuleringen er derfor;

Hvilken effekt, hvis nogen, vil det have at anvende Thompson Metoden?

10.1 Underspørgsmål

I forbindelse med Thompson Metoden og anvendelsen af de tertiære data er der nogle underspørgsmål om hvad metoden egentlig evner, som vil blive testet og forsøgt belyst igennem analysen af de indsamlede data. Disse formodninger vil kort skitseres her – og vil fungere som senere referenceramme for både problemformulering samt analysen.

Metoden evner at fremhæve mulige fejl og mangler i spørgsmålene i spørgeskemaet.

Spørgsmål 1

På baggrund af den indsamlede data, antages det at det er muligt at iagttage såfremt at respondenterne har vanskeligheder ved et eller flere spørgsmål. Dette formodes at kunne ske på en række forskellige måder, såsom at der i et givent spørgsmål er en langt højere tendens til svarsift eller besvarelsen tager længere tid end spørgsmålet kompleksitet og formulering giver anledning til. Ligeledes vil det give anledning til at forskeren kan blive gjort opmærksom på spørgsmål der overlapper eller på anden vis påvirker hinanden –

såfremt at der ved et givent spørgsmål ses tendens til at respondenterne eksempelvis bevæger sig tilbage i skemaet til et tidligere spørgsmål, for enten at genlæse eller ændre svar på et spørgsmål på en tidligere side. Det må på den vis formodes at der er tale om emner som forskeren bør blive opmærksom på.

Metoden evner at fremhæve mulige designfejl i spørgeskemaet.

Spørgsmål 2

I det at det bliver muligt for forskeren at se når respondenterne skifter svar i et spørgsmål, vil det ligeledes være muligt for forskeren at blive gjort opmærksom på deciderede designfejl i skemaet. Der kan eksempelvis være tale om forkert valg svarmulighedsdesign, i den forstand at der kunne være tale om et spørgsmål, hvor det kun er muligt at vælge en enkelt svarmulighed selvom spørgsmålet og svarmulighederne lige så godt kunne give anledning til at afgive flere svar – hvilket der vil være eksempel på i analysen.

Metoden skaber en dybere indsigt i respondenternes besvarelsesforløb.

Spørgsmål 3

I det at det nu bliver muligt at opsætte både timelines for respondenternes forløb, bliver det på den måde muligt at iagttage deres interaktion med spørgeskemaet, hvorved at det bliver muligt at synliggøre ting som ikke tidligere har været mulige. Dette giver mulighed for at iagttage og undersøge spørgeskemaet på en måde, som førhen ville have krævet at respondenterne var deltagere i et observationsstudie.

Metoden evner at fremhæve mulige speeders, cheaters og repeaters.

Spørgsmål 4

Som tidligere nævnt, så er en af målsætninger med metoden at forsøge at forbedre det webbaserede spørgeskema, ved at gøre det lettere for forskeren at identificere – og eventuelt udelade – de forskellige respondenttyper. Dette foregår især ved at sammenholde respondenternes besvarelsesforløb på tværs af respondenterne, samt med afsæt i forskerens formodninger og eventuelle pilottests. Det er på den måde muligt at belyse respondenter som skiller sig ud fra de gennemsnitlige forløb – hvilket kan iagttages og udregnes på en række forskellige måder.

11 Spørgeskemaet som observationsstudie

Når respondenterne overlades alene med spørgeskemaet, har forskerne ikke tidligere haft mulighed for at få indsigt i forløbet. Dét, der var muligt at iagttage, var blot det endelige resultat, mens selve processen i store træk forblev ubekendt. Det betyder således, at indsigten i respondentadfærden næsten udelukkende baseredes på slutresultatet. Dog forekommer der mulighed for at bede respondenter udfylde spørgeskemaet på en overvåget computer under kontrollerede forhold, hvilket rummer en risiko for, at dette vil påvirke respondentens adfærd og mønster grundet bevidstheden om, at de bliver observeret.

Observationsstudier kan ligesom spørgeskemametoder opdeles i to hovedkategorier – henholdsvis observationsstudie med og uden interaktion.

Den primære forskel er således, hvorvidt forskeren aktivt interagerer med respondenterne eller tilkendegiver og tydeliggør det egentlige motiv ved observationen, hvilket kan tilvejebringe en såkaldt Hawthorne-effekt (Homans 1965). Det betyder, at resultaterne vil variere, da respondenterne er/bliver klar over, at der er en underliggende intention med undersøgelsen. Respondenterne vil herefter forsøge på at imødekomme forskeren i deres mål og bestræbe sig på at tilvejebringe de undersøgelsesresultater, som respondenterne formoder forskeren ønsker.

Som eksempel herpå ville det kunne formodes at nogle respondenter er bevidste om, at events og svarskift registreres, hvorefter de ubevidst besværliggør spørgsmål og derved vælger flere svarmuligheder eller bruge længere tid end forventeligt. Dette kaldes *reaktivitet (reactivity)* og belyser en respondents afvigende ageren og reageren end den normale grundet bevidstheden om, at de bliver observerede (Homans 1965). Dette ses eksempelvis i kritikken af observationsmetoden *marthe meo* (videoselvobservation).

En af problemstillingerne med observationsstudier er, at de oftest kræver involvering af enten forskeren/repræsentant eller udstyr såsom videokamera, båndoptager eller lignende. Det er derfor forventeligt, at der kan opstå reaktivitet – i det at respondenterne er bevidste om observationen. Dog er det i kraft af stigende forbrugerelektronik blevet muligt at indsamle mere data via brug af respondenternes eget udstyr såsom mobiltelefoner, computere og tablets uden, at der er tale om reel overvågning.

11.1 Naturalistisk observation

For at undgå reaktivitet fremstår naturalistisk observation oplagt, idet respondenterne ikke er klar over at de bliver observeret. Dette rejser imidlertid nogle etiske problemstillinger, hvilke adresseres senere i opgaven. Da respondenterne ikke er bevidst om observationen sikres det så vidt muligt, at respondenterne agerer naturligt ved at besvare spørgeskemaet lig det normale.

I forbindelse med naturalistisk observation fastlægges sampling normaltvis, som tre forskellige metoder (Reis & Gabel 2000);

- **Timesampling:** Hvor respondenternes handlinger noteres med faste intervaller
- **Eventsampling:** Hvor man i stedet observerer bestemte hændelser og kun noterer disse.
- **Situationssampling:** Hvor respondenternes adfærd observeres i multiple situationer, hvilket gør det lettere bekræfte adfærden.

Når data indsamles via nettet, er det muligt at anvende samtlige former for sampling – dog med det i mente, at ikke alle samplingsmetoder tilvejebringer et godt resultat. Man kan argumentere for at timesampling ikke nødvendigvis vil tilvejebringe noget som eventsampling ikke kan, når der er tale om data indsamlet via en computer. Hvor forskeren i et *normalt* observationsstudie vil være i stand til at kortlægge og identificere adfærd med faste mellemrum, vil det for software være langt mere effektivt blot at vide hvilken viden og handlinger vi er interesserede i. Dette både med afsæt i at softwaren kun ser, hvad den bliver bedt om at se og at der ligeledes er begrænsninger i hvad den vil være i stand til at observere. Det mest naturlige at gøre, fra et dataindsamlingsperspektiv, vil derfor være at anvende *eventsampling*, hvor igennem at adfærd kan bekræftes på tværs af respondenterne eller undergrupperinger heraf. Forstået på den måde, at den verifikation der findes i situationssampling, hvor events registreres i multiple situationer i stedet bredes ud over større respondentgrupper – hvor adfærden så kan valideres på tværs af gruppen¹⁰.

Med Thompson Metoden åbnes der i stedet op for at selve spørgeskemaprocesen ligeledes kan iagttages som et observationsstudie, med modifikationer. Hvor man tidligere har villet anvende personlige iagttagelse eller videooptagelser til at kortlægge og indsamle

¹⁰ Selvsagt med mulighed for undergrupperinger og ikke blot samplen som helhed.

respondenternes handlinger, er det i stedet muligt automatisk at generere en slags logbog over deres besvarelsesforløb, uden at der nødvendigvis opstår reaktivitet med respondenterne.

11.2 Eventsampling som kvantitativ metode

Observationsstudier er kvalitative undersøgelser, hvor et fåtal af respondenter observeres eller interageres med. Dette formodes at skyldes det faktum, at det har været nødvendigt direkte at indlemme forskere og observatører samt deres udstyr i observationsprocessen.

Ved introduktionen af respondentens egen computer som registreringsredskab var det nu muligt at overgå til scenarier, hvori flere respondenters adfærd kunne observeres uden samtidigt at være mere tidskrævende for forskeren. I takt med at potentialet for at mængden af den indsamlede data herved stiger, er det dog nødvendigt at holde fast i hvad der for undersøgelsen vil være relevant. For at undgå at forskeren betynges i sin efterbehandling, kan det være fordelagtigt at have alle interaktioner lagret i de før omtalte events, da forskeren således kan nøjes med at fokusere på relevante data. Dette kan dog medføre bias eller projiceringsproblemstilling, da der er tale om foruddefinerede variabler, hvorved data og viden potentielt forbigås.

I anvendelse af den *autonome*¹¹ eventsampling er det muligt for forskeren at observere hændelser, der har interesse for dem, som allerede ved dataindsamlingen lagres som variable. Dette betyder ligeledes, at forskeren blot opnår viden om de events, der på forhånd menes at have interesse, hvorfor potentiel viden aktivt fravælges. Det er derfor centralt, at forskeren dels er bevidst om at undersøgelsen blot delvist giver et billede af forløbet og dels om betydning eventuelle fravalg.

11.3 Metodens begrænsninger

Det har tidligere været forsøgt at opnå indsigt i respondenternes besvarelsesforløb, hvilket ofte er foregået i et organiseret observationsstudie, hvor respondenterne møder op og besvarer spørgeskemaet i kontrollerede rammer (Groves et al 2009). Således er den grundlæggende idé om at opnå indsigt i respondenternes besvarelsesmønster og vaner

¹¹ Selvkørende – I det at det ikke kræver indblanding af en aktiv observatør

ingenlunde ny. Derimod er det nye ved Thompson Metoden måden, hvorpå data indsamles, hvilket giver forskeren mulighed for at analysere indsamlede resultater.

I skrivende stund er det ikke muligt at opnå indsigt i, hvad der foregår omkring respondenterne såsom interaktion med venner, familie og mobiltelefon eller, hvorvidt respondentens opmærksomhed afledes væk fra spørgeskemaet over på et andet emne. Denne mulighed findes derimod i det traditionelle observationsstudie i begrænset form, idet der til trods for at forskerne er i stand til at observere, hvad respondenterne interagerer med - befinder respondenterne sig ikke i et vanligt miljø og er i et kontrolleret scenarie ikke omgivet af de ting man normalvis vil forvente ville kunne have effekt på respondenterne, såsom familie.

Både i Thompson Metoden såvel som det traditionelle observationsstudie ligger der en begrænsning i, at det er ikke muligt at undersøge, hvad respondenterne tænker. I observationsstudiet er det muligt at bede respondenterne om at tænke højt – dog med fare for at det kan skabe et mere kunstigt scenarie, da respondenterne i gængse rammer ikke formodes at skulle remse tanker op. Det betyder således, at forskeren ikke er i stand til at opnå indsigt i respondenterne tanker men blot deres handlinger. Det vil altså sige, at sidder en respondent og overvejer hvilket svar de ønsker at afgive, er det os ikke muligt at opnå indsigt i den proces – medmindre at de i stedet bevæger sig med et bestemt mønster eller har tendens til eksempelvis at have en delvis satisficing, hvorved at de vælger det svar der passer dem bedst som de læser dem. Det vil altså sige at de vælger det svar på listen først, som de føler repræsenterer deres holdning nogenlunde – hvorefter de korrigerer til et svar som de synes passer bedre som forekommer senere på listen.

12 Etik

Thompson Metodens observations- og indsamlingsmuligheder rejser en række etiske dilemmaer om, hvorvidt det er forsvarligt at indsamle data, som respondenterne ikke er klar over bliver indsamlet, samt hvor grænsen for hvad man bør indsamle går?

I den naturalistiske observation opnår forskeren adgang til viden, som ikke er dem bevidst videregivet og de indlemmes i en lang række information som respondenterne ikke er bevidste om at de videregiver. Disse informationer er ubevidste i den forstand, at det ikke

nødvendigvis er handlinger, som respondenten er bevidste om at de foretager - og ikke mindst fordi at vi er i stand til at observere de handlinger. Ligeledes åbnes der op for at forskeren kan iagttage svars-kifte, som respondenten søger at skjule ved vælge 'Ved ikke' efter at have markeret et *reelt* svar. På den måde er det muligt for forskeren at opnå indsigt i noget, som respondenten enten ubevidst eller ikke har ønsket at videregive.

12.1 Anonymitet

En af de mest essentielle opgaver, som forskeren varetager i sin forskning, er at opretholde respondentens anonymitet og sikkerhed. Det betyder altså at den indsamlede data på ingen måde må kunne lede til en direkte identifikation af den enkelte respondent eller gruppe som de tilhører. Der har gennem historien været en længere række tilfælde, hvori at hele respondentgrupper har vist sig let-identificerbare – hvilket i værste tilfælde har udsat respondenterne for decideret fare (Hviid-Jakobsen & Kristiansen 2001: 13 – 17).

Fare er ikke nødvendigvis et forstå som fysisk skade, men udsætte dem for situationer der på baggrund af hvad de har oplyst kan kompromittere deres nuværende livssituation, såsom ansættelsesforhold, anseelse med videre. Der kunne eksempelvis være tale om en arbejdsmarkedsundersøgelse, hvori forhold både professionelle og personlige forhold undersøges. Såfremt at information omkring holdning eller hændelser i forhold til arbejdet, ville kunne tilbageføres til den enkelte respondenten, kan det kompromittere respondentens ansættelsesforhold (Hviid-Jakobsen & Kristiansen: 75 – 76).

12.2 Følelsen af anonymitet

Opretholdelsen af anonymitet er ikke kun for respondentens sikkerhed men ligeså vel for at sikre, at de afgiver korrekt information. Det er derfor centralt, at respondenten oplever sin anonymitet opretholdt, da det ellers kan have effekt på besvarelser – særligt i forbindelse med kompromitterende spørgsmål, som eventuelt ville kunne udsætte respondenten for fare (Kvale 2009).

Spørgsmålet er således, hvilke ting kan bevirke, at respondenten føler sin anonymitet kompromitteres, og hvor stor del af dette kan bevirke reaktivitet? Det ses, at respondenter, der deltager i undersøgelser med præmiechancer, har tendens til at afgive svar, som de formoder forskeren eller virksomheden ønsker (Couper et al 2011). Dette kan skyldes

respondentens tro på, at et positivt svar vil medføre større vinderchancer, samt det faktum at de nødsages til at afgive personlig information for at kunne modtage en eventuel præmie. Derfor må det formodes, at følelsen af anonymitet i et sådan scenarie er ringe i forhold til undersøgelser uden tilknytning af *direkte* personlig data.

Det er således interessant at undersøge, hvorvidt respondentens manglende bevidsthed om at de bliver digitalt observeret har en effekt på den oplevede anonymitet.

På baggrund af observationsstudier er der skabt blik for, at respondenter agerer anderledes og med en anden bevidsthed, men hvordan påvirker det respondenternes oplevelse af anonymitet? Det må formodes, at det i høj grad afhænger af respondentens forståelse af de data, der bliver indsamlede, og hvad dataene er i stand til at fortælle forskeren.

Videre må det antages, at respondenterne vil reagere på, hvis de med teknisk snilde, er i stand til at aflure at der bliver videregivet data som de ikke er informerede om. Dette kan pludselig sætte spørgsmålstegn ved den *fortrolighed* der bør eksistere respondent og forsker imellem. Men vil respondenterne på anden vis føle at deres anonymitet kompromitteres, hvis de på forhånd oplyses eller har mulighed for at blive oplyst hvad der videregives og med hvilken intention.

12.3 Informeret samtykke

Det er normal konvention at respondenterne ved deltagelse i en undersøgelse afgiver et informeret samtykke, hvilket betyder at de på velinformeret og saglig baggrund giver accept for deres deltagelse (Hviid-Jakobsen & Kristiansen: 74 – 76). Hvilket betyder at respondenterne er velvidende om præcist hvad de deltager i og er bekendt med konsekvenserne heraf. Dette er oftest refereret til som *fuldt informeret samtykke*, som igennem de seneste år er begyndt at blive betragtet som værende en utopi, mere end et realistisk opnåeligt mål (Hviid-Jakobsen & Kristiansen: 74).

Man er i sociologien, antropologien og psykologien i stedet begyndt at anvende og acceptere *tilstrækkelig informeret samtykke* og afarter heraf, såsom *stiltiende samtykke*, *efterfølgende samtykke* og *ubevidst samtykke*. Nogle af hvilke kan betragtes som værende uetiske (Hviid-

Jakobsen & Kristiansen: 74 – 75). Dog synes ingen af disse accepter at dække over, den accept der forekommer ved deltagelse i en spørgeskemaundersøgelse.

Det er almen praksis at oplyse hvem undersøgelsen foretages af, emnet der undersøges, såvel som oplysning om at al data indsamles anonymt og at den indsamlede data ikke vil kunne anvendes i reklame- eller virksomhedsmæssige henseende, såfremt at dette er tilfældet. Spørgsmålet er således om det er tilstrækkelig viden til at respondenterne vil være i stand til at afgive et informeret samtykke – eller om man, som forsker, i stedet bør acceptere at det delvise samtykke nu også rummer viden som betragtes som værende implicit.

12.4 Hvad ved vi egentlig

Med et etisk udgangspunkt i Thompson Metoden, er spørgsmålet således, hvad vi er i stand til at vide og om respondenterne på baggrund af den tertiære data der indsamles. Om disse data på nogen måde vil være i stand til at kompromittere respondenterens anonymitet. Der er i projektet kun anvendt et begrænset antal af de mulige tertiære data der kan indsamles, hvilket derfor vil være omdrejningspunktet for denne diskussion.

De indsamlede tertiære data er, som nævnt, *tid*, *svarskift* og *progression*. Data som har til formål at tegne et sigende billede på respondenterens interaktion med spørgeskemaet. Der kan på baggrund af de indsamlede data opnås indsigt der rækker ud over den sædvanlige dataindsamling, i det at vi eksempelvis er i stand til at opnå indsigt i svar som respondenterne måske har valgt at ændre til i stedet at være neutral- eller nul-svar. Det betyder selvfølgelig, at hvis netop dét spørgsmål omhandler noget der ville kunne kompromittere respondenterens anonymitet, vil der kunne opnås viden derom. Dette er dog kun et tænkt eksempel, som kun vil kunne opstå hvis der i forvejen er en designfejl i spørgeskemaet som tillader dette.

Det synes ikke sandsynligt at de indsamlede data vil kunne sammensættes på en sådan vis at det vil kunne være direkte identificerende for respondenterne. Den indsigt der tilbydes på baggrund af de tertiære data synes ikke at medvirke til at kunne skabe et billede direkte billede på respondenterne. Det betyder selvsagt ikke, at det ikke vil være muligt at udbygge de tertiære data der indsamles til også at rumme informationer, som bringer os tættere på

respondentens egentlige identitet. Dog vil der, som oftest, eksistere en barriere som beror sig på at respondentens computer er lettere identificerbar end respondenten bag skærmen, da det er svært at stille nogen som helst garanti for hvem der sidder ved tastaturet.

12.5 Hvad er vi i stand til at vide

Det kan være en vanskelig balancegang, at balancere respondenternes anonymitet – og i jagten på et mere uddybet billede og indsigt i respondentbesvarelsen, er det muligt at man på et givent tidspunkt møder en grænse. Hvis grænsen overskrides, vil det kompromittere respondentens anonymitet – sågar muligvis uden at kompromittere deres følelser heraf. Men hvor går grænsen – og hvad er vi i det hele taget i stand til at vide om respondenten?

Afsættet for hvad vi er i stand til at vide er udelukkende baseret på den data vi er i stand til at indsamle og bearbejde i forbindelse med respondentens interaktion med spørgeskemaet. Der vil i afsnittet betragtes muligheder der findes i både de nuværende datasæt, samt de som tilvejebringes ved anvendelse af Thompson Metoden.

I arbejdet med større kvantitative datasæt, vil man ofte opleve at man ser nogle helt unikke mønstre i besvarelsene. Ind imellem findes nogle mønstre som kun dukker op en enkelt gang og som kun karakteriserer den enkelte respondent. I et utopisk scenarie, ville besvarelsen og mønstret gentage sig, såfremt at respondenten skulle gennemføre spørgeskemaet atter engang. Men sandsynligheden for at det sker yderst lille – da respondenternes besvarelse selvfølgelig også er afhængig af deres hverdag – alt sammen påvirket af et hav af forskellige faktorer. Selv i de tilfælde hvor det skulle ske at mønstret gentager sig, er der tale om et exceptionelt lille antal – og der er fortsat grænser for hvor identificerbar denne data derfor er.

Alternativt kan man i stedet kigge på en udbredelse og viderebygning af baggrundsvARIABLE. Hvor alder og køn i de fleste tilfælde vil være forudsigelige, vil alle andre baggrundsvARIABLE kunne variere fra gang til gang i en uforudsigelig grad. Det betyder altså, at vi på ingen måde kan være sikre på at respondenten der i et givent arbejde også vil være det næste gang vi spørger. Enkelte situationer vil dog kunne betragtes som værende mere permanente, såsom forskellige former for pension. Men det er atter engang

ikke noget der nødvendigvis vil kunne give os indsigt og identificere den enkelte respondent, mellem undersøgelserne.

Breder man derfor baggrundsvariable ud over noget der mere statisk, sjældnere skifter og som samtidigt har rum for variation, kan man med udgangspunkt i nogle af de data der videregives af respondentens computer, danne en slags digitalt fingeraftryk (Eckersly 2014). Ved at sammenholde informationer som automatisk videregives af respondentens *browser*, skabes der et unikt id som respondenten kan identificeres ved. Der anvendes data såsom version af styresystem, indstillinger i deres browser, skrifttyper der er installeret og meget andet til at danne et samlet billede. Det betyder selvfølgelig at hvis respondenten installerer en ny tilføjelse til sin *browser*, en ny skrifttype eller lignende vil fingeraftrykket ikke være nøjagtigt det samme. Dette kan dog betragtes som værende en forudsigelig forandring, hvor det accepteres at der med en given variation indenfor bestemte parametre stadig vil kunne identificere individet, på samme måde som med *delvise* fingeraftryk. Men i sig selv siger det intet om individet, men om den enkelte computer, smartphone eller tablet. Vælger man da, at kombinere det med baggrundsdata fra en spørgeskemaundersøgelse, vil man dog med større sikkerhed kunne sige hvorvidt at sandsynligheden for at der er tale om den samme respondent. Men er det ønskeligt og ikke mindst, i hvor høj grad kompromitterer det respondentens identitet?

De data der afgives, kan opsnappes uden respondentens vidende og er baseret på ganske almindelige informationer som videregives i en normal transaktion, hvor man besøger en hjemmeside via nettet. Der deles et hav af informationer, som sikrer kompatibilitet og kommunikationsvej – i en sådan grad, at det altså gør det muligt at skelne mellem brugerne. Det kan nævnes, at det først i 2014 blev ændret på de data der videregives af de største danske teleoperatører. Det har indtil da været muligt for alle de hjemmesider personer med hemmeligt nummer har besøgt, at aflæse brugerens telefonnummer blandt de informationer som blev videregivet, blot ved at de gik ind på hjemmesiden¹² (Politiken 11/2014).

¹² Såkaldt header-data, der altid forekommer ved et normalt *handshake* med en hjemmeside, har som levn fra sen-90erne indeholdt brugerens hemmelige nummer, når hjemmesiden blev tilgået via mobil/smartphone.

I essensen betyder egentlig at det er muligt, med relativt få midler at skabe mulighed for at kunne identificere respondenter på tværs af spørgeskemaer. Der er som sådan ikke tale om en personlig identifikation, hvor vi med udgangspunkt i den indsamlede data vil være i stand til at deducere respondentens bopælsadresse, udseende eller lignende. Der er i stedet tale om at man på tværs af disse data, kan gøre respondenten identificerbar – og genkende dem på tværs af spørgeskemaer og nettet som helhed. Det vil altså sige, at vi kan genkende respondenten, når vi *ser* dem – og altså ikke decideret være i stand til at opsøge dem.

Men hvor går den etiske grænse? Er det i orden at søge at gøre respondenten elektronisk identificerbar – eller bør vi respektere den anonymitet som nettet tilbyder nu? Det må være en diskussion den enkelte forsker bør tage op med sig selv.

13 Operationalisering

Den overordnede målsætning med projektet, er at belyse hvorvidt at Thompson Metoden evner at fremhæve designfejl i spørgeskemaet, identificere forskellige respondenttyper såvel som forbedre sikkerheden af de indsamlede data, samt til hvilken grad. Således er målet at belyse, hvorvidt denne metode evner:

- At fremhæve respondenttyper som kan være problematiske for undersøgelsesresultaternes validitet.
- At henlede forskerens opmærksomhed på mulige designfejl i spørgeskemaet.
- At skabe en reel indsigt i respondenternes forløb for herved at kunne anvendes i undersøgelser, der har til formål at belyse eksempelvis satisficing eller social desirability.
- At kunne anvendes til at underbygge argumenter for respondentadfærd, såsom den lave spredning der opleves i webbaserede undersøgelser.

For at kunne undersøge disse forskellige funktioner og aspekter, har det derfor været nødvendigt at konstruere en undersøgelse som ville kunne belyse og indsamle data fra forskellige vinkler. I det at der ikke tidligere har fandtes systemer som evner at indsamle tertiære data i det omfang der kræves af metoden, har det været nødvendigt at udvikle et nyt surveysystem, som jeg sideløbende med projektet har programmeret fra bunden. Hvilket betyder at der har skullet tages højde for alt fra spørgeskemateknisk design, til en decideret implementering af metoden. Grundet datasættets sammensætning og arkitektur, har det ligeledes været nødvendigt at programmere samtlige af analyse- og

databehandlingsredskaber som er anvendt i analysen. Hvilket betyder, at der ikke er taget udgangspunkt i et i forvejen eksisterende statistikprogram, men at der i stedet er udviklet et skræddersyet analyseværktøj som evner at håndtere datakonstruktionen.

I dette afsnit vil jeg således både søge at redegøre for operationaliseringen samt hensigten med metoden og undersøgelsen. Introduktionen til operationaliseringen af metoden inkluderes for herigennem at skabe indsigt i det tekniske perspektiv og en forståelse af, hvordan dataindsamlingen er forløbet og formodet oplevet.

13.1 Operationalisering af metoden

Dette afsnit har til formål at give en indsigt i systemets funktionalitet for herved at illustrere og skabe indblik i teknikaliteterne i indsamlingsprocessen såvel som de data, der lagres til endelig anvendelse. Dette skyldes imidlertid, at det har været nødvendigt at programmere et specialindrettet system, som evner at registrere og opbevare relevante informationer. Det betyder, at jeg sideløbende med udviklingen af projektet har været programmeret et spørgeskemasystem, som rummer de muligheder, der har været behov for.

Systemet er struktureret i *slides* lig Microsofts *Power Point*. Hvert slide rummer en række spørgsmål og/eller informationer samt navigationsfunktionalitet, hvorved et slide både kan være sammensat af spørgsmål vedrørende baggrundsdata samt mulighed for at gå frem eller tilbage i spørgeskemaet. Ved hver svarmulighed er der tilknyttet en såkaldt *event listener* – en lytter – som har til opgave identificere, hvornår der sker noget i det gældende spørgsmål og rapportere til databasen. Respondentens valg af eksempelvis årstal i fødselsår vil således blive lagret midlertidigt, indtil at de bevæger sig videre i spørgeskemaet.

På den måde er systemet konstant på vågeblus og venter på ændringer, om det er at respondenter vælger et svar – hvor ændringen altså er at de går fra intet svar til at have et – eller til at de ændrer en given svarmulighed, til at de bevæger sig videre til næste side. Hver gang en ændring forekommer, bliver eventet lagret lokalt af respondentens browser, som opbevarer disse data indtil at de bevæger sig frem eller tilbage mellem siderne i spørgeskemaet. Hver ændring får automatisk tilknyttet en række værdier, som gør det muligt at identificere præcist, hvad der foregår.

Disse værdier er opdelt i fire variable;

- Tidsmarkør
- Type
- Navn
- Værdi

Tidsmarkøren er et såkaldt *unix*-tidspunkt, som registreres i millisekunder med henblik på at tage højde for respondenters eventuelle og pludselige skift. *Typen* beretter om hvilken type handling, der er foretaget såsom besvarelse af et spørgsmål eller progression i skemaet. *Navn* identificerer det spørgsmål eller element, der arbejdes med, mens *værdien* beretter om respondentens valg. Et eksempel herpå kan se ud som følgende;

- Tid: 1415992019179
- Type: Svar
- Navn: Fødselsår
- Værdi: 1989

Disse handlinger og ændringer kædes sammen, så at de fremstår i en sammenhængende kæde af events, som er lagret i den rækkefølge de er forekommet. Disse lagres i en database når respondenten skifter side, ved at bevæge sig frem eller tilbage, såvel som når de afslutter skemaet. Det betyder, at for hver gang at respondenten bevæger sig i spørgeskemaet, vil alle forgående events gemmes, hvilket betyder at delvise besvarelser også lagres. Det vil altså sige, at hvis en respondent i løbet af processen har valgt at afslutte skemaet før det er gennemført, vil forskeren være i stand til at se hvad der er forløbet op til – samt hvilken side de befandt sig på, da de afsluttede skemaet. Det er dog vigtigt at være opmærksom på at en ufuldendt besvarelse ikke vil have en decideret slutmarkør, da det ikke er muligt at indsamle data fra en browser i det øjeblik den lukkes. Hvilket betyder at en besvarelse som er afsluttet før tid, blot vil have sidste progressionsevent som afsluttende hændelse, modsat det sædvanlige *end*-event der tilføjes til færdiggjorte besvarelser.

Alt om alt betyder det at der i databasen bliver lagret en samlet tidslinje for hver enkelt respondent, der nøje beretter om deres interaktion med spørgeskemaet. Al dette foregår uden synlige spor, hvorfor at det vil være uden respondentens vidende at det foregår.

13.1.1 Teknisk afklaring

I forbindelse med udarbejdelse af undersøgelsen har der rejst sig spørgsmål vedrørende dataopbevaring og lovgivning på området. Disse tekniske overvejelser og diskussioner er både relevante fra et teknisk samt et juridisk perspektiv og vil i det følgende kort berøres med henblik på at skabe afklaring for nogle af de spørgsmål, der måtte være i forbindelse med operationalisering af metoden. Et har fokus især været om, hvorvidt der kan være nogen former for implikationer i forhold til *cookies*¹³, som ved lov kræver samtykke fra brugeren (Retsinformation)¹⁴. Systemet anvender ikke cookies men er i stedet bygget på JavaScript og lagrer derfor hverken permanente sporingsparametre eller anden data på respondentens computer. Det er derfor ikke nødvendigt at kræve accept til cookies fra respondenterne før deltagelse i undersøgelsen.

Ligeledes har der været usikkerhed omkring metodens implementering i forbindelse med JavaScript, som ofte har været forvekslet med Java fra blandt andet NemID. JavaScript og Java er imidlertid kun historisk relaterede. Java fordrer, at brugeren har installeret en applikation, som kommunikerer og kræver accept fra respondenterne blandt andet til adgang i respondentens filsystem. Modsat er JavaScript allerede implementeret i respondentens *browser*¹⁵ og kræver ikke accept, da den blot er i stand til at håndtere dét, der foregår i det enkelte browservindue. Dette betyder, at det ikke er nødvendigt at spørge respondenterne om hverken cookies og at de ligeledes heller ikke skal acceptere eksekveringen af den kode som registrerer de forskellige *events*. Dette har hele tiden været i trit med målsætningen for spørgeskemasystemet, da det fra begyndelsen var intentionen, at spørgeskemaet **ikke** unødigt skulle interagere eller påvirke respondenterne.

Slutteligt har der hersket tvivl om udbredelsen af JavaScript, som er anvendt i surveysystemet, og om hvorvidt der er tale om en udbredt standard. Hertil kan det påpeges, at der vil være tale om op imod 1 % af de som bruger internettet, har deaktiveret JavaScript eller benytter en browser der ikke understøtter det. Dette gør sig gældende for både computer, tablet og mobil (GDS UK GOV 2013).

¹³ Statisk lagret brugerdata – som lagres af brugerens browser.

¹⁴ <https://www.retsinformation.dk/forms/R0710.aspx?id=139279>

¹⁵ Program der anvendes til at gå på nettet.

14 Empiri

Rent empirisk støtter dette speciale sig til en tidligere undersøgelse *Digitalisering af det offentlige* foretaget af KMD i 2011. Inddragelsen af denne undersøgelse belyser den udvikling, som det offentlige fortsat befinder sig i, og som blandt andet vedrører digitaliseringen af offentlige ydelser såsom post i form af e-boks. KMD er én af de primære leverandører af digitale løsninger til det offentlige, hvorved en undersøgelse der påviser at danskerne forholder sig positivt til digitalisering, kan åbne op for yderligere opgaver for KMD. Dertil kommer, at KMDs undersøgelse var tilstrækkeligt behæftet med problemstillinger, der under normale omstændigheder kan betragtes som utilstrækkelig både i forhold til design, formuleringer og sammensætninger. Der ville her kunne argumentere for et brist i undersøgelsens integritet, da der opstår interessekonflikt.

Undersøgelsen forekom således som en oplagt mulighed for at opdele undersøgelsen i to spørgeskemaer; Det ene bestående af størstedelen af det oprindelige KMD-skema; Det andet i en sociologisk kongruent pendant hertil. Det skal forstås på den måde at spørgsmål er udredt, således at de ikke rummer flere spørgsmål i et – at der nu er et tydeligt tematisk sammenhæng, såvel som at mulige tvetydige formuleringer er blevet omformuleret for at give tydeligere og lettere forståelige spørgsmål.

14.1 Todelt spørgeskema

Projektet er designet således at den består af to tematisk ensartede spørgeskema, hvoraf det ene er uddrag af det spørgeskema som oprindeligt er udarbejdet af KMD, hvor både spørgsmål, svarmuligheder og tematisk rækkefølge er bibeholdt. Det andet spørgeskema er en omstrukturering, hvor der er søgt at tage hånd om de problemer der var i det oprindelige KMD skema. Dette er gjort ved eksempelvis at sikre at spørgsmål og svar matchede, etablere kategorisk kongruens i forhold til temaer og skabe et mere naturligt *flow*.

I det at respondenterne åbner skemaet, bliver de tilfældigt vist ét af de to spørgeskemaer. Det er respondenterne ubevidst at de er blevet styret ind på et bestemt spørgeskema og de er ligeledes heller ikke bevidst om hvilket, i og med at der ikke er noget der angiver at denne opdeling forekommer. Det visuelle udtryk, såvel som indledning/introduktion til skemaerne er identisk, for så vidt muligt at undgå at påvirke respondenterne.

Formuleringer, rækkefølge og svarmuligheder er i KMD skemaet bevaret som i den oprindelige undersøgelse, imens det i det andet skema er en forbedret *parallel* til KMD undersøgelsen, hvor spørgsmål, svar og struktur er blevet ændret og revurderet.

Det betyder at spørgsmål er blevet omformuleret og udredt for tvetydigheder, multiple spørgsmål og umulige svarsammensætninger – såvel som at de er blevet omstruktureret, med målet om at skabe en klarere tematisk sammensætning. Der ses i det oprindelige spørgeskema fra KMD flere tilfælde, hvor et spørgsmål egentlig dækker over flere spørgsmål – eller svarmulighederne ikke passer til det egentlige spørgsmål. Nogle potentielle meninger/svar er derfor ikke at finde i skemaet, hvorved at der opstod en mangel på repræsentative svar.

Manglen på repræsentative svar eller svarmuligheder, som afdækker flere spørgsmål eller som bevirker at respondenterne ikke vil være i stand til at besvare hvad der egentlig spurgt, formodedes at kunne bevirke at nogle respondenter ville benytte længere tid eller have besvær ved at besvare spørgsmålet. Det synes derfor at være en oplagt mulighed for at sammenligne det forskelle der måtte være på tværs af de to spørgeskemaer. Der er en række tænkelige scenarier, der vil kunne opleves i designet, som kan formodes oplevet på tværs af de to spørgeskemaer.

For at undersøge progression og tidsforbrug i undersøgelsen, er der i det videreudviklede skema ligeledes forsøgt at skabe et mere kongruent og sammenhængende tematisk struktur. Det oprindelige spørgeskemas struktur synes at springe, rent tematisk, hvor der skiftevis omhandler det ene og det andet tema. Dette formodes at kunne give anledning til forundring eller forvirring for respondenterne, hvilket formodes at kunne fremprovokere at nogle respondenter vælger at gå tilbage og sammenholde med tidligere spørgsmål.

14.2 Designmæssige beslutninger

Der er i forbindelse med udarbejdelsen og opsætningen af begge spørgeskemaerne truffet en række designmæssige beslutninger, som kan have indflydelse på selve undersøgelsen. Nogle af disse beslutninger er truffet, for på at fremme eller skabe mulighed bestemte typer adfærd blandt respondenterne.

Der er i spørgeskemaregi forskel på om taler om *fri mobilitet* eller *lineær mobilitet* (Thompson 2012). Hvor den lineære mobilitet, kun gør det muligt for respondenterne at bevæge sig fremad i undersøgelsen, giver den fri mobilitet respondenterne mulighed for at bevæge sig frem og tilbage i undersøgelsen på et hvilket som helst tidspunkt. I denne undersøgelse er det et aktivt valg, at give respondenterne fri mobilitet – i det at der er behov for at undersøge respondenternes progression. Dette kunne selvsagt også gøres i et lineært forhold, mens muligheden for at bevæge sig tilbage i spørgeskemaet giver ligeledes respondenterne mulighed for at beslutte sig for at gå tilbage i undersøgelsen for at læse et spørgsmål igen, skifte svar eller blot sikre sig at de har markeret den ønskede svarmulighed. Afskærer man respondenterne for muligheden for at bevæge sig frem og tilbage i undersøgelsen og ændre svar løbende, risikerer man ligeledes at respondenterne kan blive i tvivl om de har markeret det ønskede svar – eller føler at det svar de har valgt alligevel ikke var repræsentativt for dem. Det kan for nogle respondenter betyde at deres koncentration omkring det resterende af skemaet sænkes – eller i nogle tilfælde at de bliver frustrerede og afslutter undersøgelsen. Til trods for at der lagres delvise besvarelser, er det selvfølgelig langt mere ønskeligt at så mange som muligt gennemfører undersøgelsen.

Det blev i Soc-KMD ligeledes besluttet, at respondenterne i langt de fleste spørgsmål skulle suppleres med et neutral- eller nulsvar, således at de var i stand til at vælge det som en given mulighed. Dog med de få undtagelser vedrørende alder og bopælsregion, hvor det ikke er muligt at vælge et nulsvar.

Sidst, er det aktivt besluttet at respondenterne skal besvare samtlige spørgsmål på en given side før de bevæger sig videre. Det betyder altså, at alle spørgsmål på side et skal besvares før de kan gå videre til side to. Denne beslutning er truffet, for at sikre at respondenter tager stilling på den ene eller anden vis – såvel som for at etablere muligheden for at studere deres reaktion, såfremt det opstår. Det er i systemets progressionstracker muligt at se, hvis en respondent forsøger at bevæge sig videre i spørgeskemaet uden at besvare alle spørgsmål – og derfor bliver promptet med en fejl.

14.3 Spørgeskemaets struktur

Dette afsnit har til hensigt at skabe indsigt i spørgsmålenes udvikling fra det oprindelige KMD-spørgeskema til det sociologisk færdigudviklede. Baggrundsdata såsom fødselsår, køn, jobsituation og bopælsregion er dog ensartede i begge skemaer og vil derfor ikke indgå i oversigten.

Udgangspunktet for spørgsmålenes kronologi er opstillet efter KMDs skema – dog vil der kun inddrages en håndfuld af spørgsmål og deres udvikling som er i skemaet.

	KMD	Soc-KMD
Spørgsmål	Tror du, at den offentlige sektor kan blive mere effektiv og spare penge ved at anvende digitale løsninger og selvbetjening?	Tror du, at den offentlige sektor kan blive mere effektiv ved at anvende digitale løsninger og selvbetjening?
Svarmuligheder	<ul style="list-style-type: none">- Ja, den offentlige sektor bliver meget mere effektiv.- Ja, den offentlige sektor bliver noget mere effektiv.- Den offentlige sektor bliver hverken mere eller mindre effektiv.- Nej, den offentlige sektor bliver noget mindre effektiv.- Nej, den offentlige sektor bliver meget mindre effektiv.- Ved ikke	<ul style="list-style-type: none">- Ja, den offentlige sektor bliver meget mere effektiv.- Ja, den offentlige sektor bliver noget mere effektiv.- Den offentlige sektor bliver hverken mere eller mindre effektiv.- Nej, den offentlige sektor bliver noget mindre effektiv.- Nej, den offentlige sektor bliver meget mindre effektiv.- Ved ikke

I det første eksempel ses indtil flere af de før omtalte problemstillinger i spørgeskemaet, i det at spørgsmålet som oprindeligt stillet af KMD rummer to spørgsmål, men kun svarmulighed der dækker den ene. Svarmulighederne på de to spørgsmål er identiske, men spørgsmålet i det selvudviklede skema rummer i stedet kun et spørgsmål der er direkte knyttet til de svarmuligheder respondenterne præsenteres for. Man kan argumentere for at *effektivitet* og *økonomi* hænger sammen, imens at det samtidigt kan give associationer der er knyttet til hvordan borgerne oplever det offentlige.

	KMD	Soc-KMD
Spørgsmål	Synes du, at den offentlige sektor skal gøre mere for at hjælpe borgerne med at anvende computere og smartphones?	Mener du at der er behov for at hjælpe borgerne med at anvende computere og smartphones?
Svarmuligheder	<ul style="list-style-type: none"> - Ja, det er der helt sikkert behov for. - Ja, det er der behov for. - Måske - Nej, det er der ikke behov for. - Nej, det er der slet ikke behov for. 	<ul style="list-style-type: none"> - Ja, det er der helt sikkert behov for. - Ja, det er der behov for. - Måske - Nej, det er der ikke behov for. - Nej, det er der slet ikke behov for.
Spørgsmål	-	Synes du, at den offentlige sektor skal gøre mere for at hjælpe borgerne med at anvende computere og smartphones?
Svarmuligheder	-	<ul style="list-style-type: none"> - Ja, det er det offentliges ansvar at tilbyde borgerne den fornødne undervisning og information. - Nej, det er ikke det offentliges opgave at ajourføre borgerne med ny teknologi. - Ved ikke

I dette tilfælde er der i stedet tale om et spørgsmål, som ændres til to – i det at spørgsmålet oprindeligt omhandler hvorvidt at det offentlige bør gøre mere for at hjælpe borgerne, imens svarmulighederne i stedet er vægtede på at det behov der måtte være. Det rejser et dilemma, hvori at en respondent kan mene at der vil være behov for at hjælpe og undervise borgerne, men at det ikke er det offentliges opgave. Det betyder således at der i KMD spørgeskemaet ikke vil være muligt at angive som svarmulighed, hvorved at der ikke findes svarmuligheder der for respondenteren er repræsentativ. I retrospekt, vil jeg dog understrege at formuleringen af svarmuligheder i andet spørgsmål i Soc-KMD skemaet muligvis burde være en anden, da det først og fremmest kan fortolkes og rummer fremmedord.

14.4 Sample

I kraft af at spørgeskemaerne udelukkende er onlinebaseret, accepteres ligeledes visse selektionsbias såsom skævvridninger i populationsfordelingen i forhold til alder og uddannelse, som også tidligere illustreret.

Der opstår ligeledes en kraftig bias grundet det emnemæssige valg og undersøgelsesformen, da emnet for undersøgelsen er fokuseret omkring digitalisering af det offentlige. I forlængelse heraf må det antages, at respondenter, der har fundet og deltaget i spørgeskemaet, vil have et vis kendskab til computere og internettet som helhed, hvilket i sig selv betyder, at de selv må forventes bedre rustede til at håndtere en digitalisering af offentlige instanser.

Anerledes ser det ud med de potentielle respondenter, som ikke kunne nås – herunder personer som ikke har adgang eller benytter computere og internet på en regulær basis. Netop disse grupper må også forventes at være dem, der vil have det vanskeligst i forhold til digitaliseringen.

Havde projektets undersøgelse været rettet mod at skabe en reel indsigt i danskernes holdning til digitalisering, ville det derfor have været et mere robust design såfremt at der var anvendt mix-modes til indsamling af data, således at man bedre kunne undgå *under-coverage*.

14.4.1 Kontakt til respondenter

I forbindelse med projektet er kontakten til respondenterne er primært gået gennem online midler, både i form af sociale netværk, forskellige fora og email. Da nettet oftest er opdelt i interessegrupper, er det forsøgt at brede udsendelse af spørgeskemaet ud over flere forskellige grupperinger. Her har der især været fokus på at ramme nogle med både høj og lav computerkompetence, såvel som det har været stræbet efter at opnå en repræsentativ aldersfordeling. Det betyder derfor at skemaet bl.a. har været uddelt via Ældre Sagens Facebookside.

Spørgeskemaerne emne, ”Digitalisering af det offentlige”, har været tydeligt fremhævet i de forskellige oplæg, hvilket på daværende tidspunkt var et højaktuelt emne i diverse medier.

14.4.2 Komplikationer i respondentindsamling

Som tidligere nævnt, lider onlinebaserede spørgeskemaer af en kritisk lav svarrate – hvilket ligeledes har oplevedes i indsamlingen af empiri til undersøgelsen. Til trods for at søge en bredspektret tilgang til indsamling, for på den måde at overkomme problemstillinger såsom tid og sted – har det været vanskeligt at akkumulere det antal respondentbesvarelser der ved undersøgelsens start har været antaget som værende tilstrækkeligt for projektet.

Alternativt kunne man have søgt et samarbejde med allerede eksisterende undersøgelser, som allerede besidder en eksisterende respondentbase – dog med mulighed for en begrænsning af det emnemæssige output og den mulige indflydelse på spørgeskemaet, både i kraft af indhold og struktur. Ligeledes kunne man have søgt en bredere udbredelse, ved omdeling gennem udsendelse via sekretariatene på Aalborg Universitet. Denne omdeling lider dog under en uddannelsesmæssig bias og en lav responsrate – men det kunne potentielt set have været et supplement til de anvendte medier.

15 Analyse

Gennem analysen arbejdes der imod at belyse hvorvidt at de underspørgsmål der er opsat for projektet har nogen merit – eller om de kan afvises. Der vil igennem kapitlet være en nøje gennemgang af den indsamlede empiri – for ligeledes at undersøge hvorvidt at der er muligheder og kapaciteter i Thompson Metoden, som ikke på forhånd var antaget.

15.1 Resultater

For at skabe indsigt i de data der er blevet indsamlet, vil der i dette afsnit være en gennemgang af de indsamlede data – for at give et overblik over det datamæssige omfang, såvel som at illustrere et fundament for videre undersøgelse. Først, vil fokus være på de primære data der er blevet indsamlet i undersøgelsen – hvorefter der vil være en gennemgang af indsamlede tertiære variable.

Som nævnt har det været utroligt vanskeligt at få indsamlet respondentbesvarelser, hvilket oprindeligt også var antaget, dog var omfanget af problemet langt større end oprindeligt antaget. Hvilket betyder derfor at det samlede antal besvarelser er lavere end oprindeligt ønsket.

Dataindsamlingen er foretaget over en tre ugers periode, fra anden uge af november 2014, frem til slutningen af måneden. Det var et bevidst valg at forsøge afslutte dataindsamlingen inden december, for på den måde at søge at undgå et sammenfald med den travlhed og stress der må forventes at kunne opstå i netop den måned.

15.1.1 Respondentoversigt

For kort at skabe indsigt i respondentbasen og fordelingen af respondenter over de to spørgeskemaer, vil der i dette afsnit være en kort gennemgang af respondenternes baggrundsvARIABLE. Formålet er blot at skabe en oversigt over hvem der har besvaret spørgeskemaerne, uden at dette udarter sig til en dybdegående analyse af respondenternes besvarelser.

	KMD	Soc-KMD	Total
N	38	28	66
Fuldendte besvarelser	28	25	53
Ufuldendte besvarelser	10	3	13

Tabel 15.1.1.1 – Besvarelser

I Tabel 15.1.1.1 ses besvarelserne for de to respektive spørgeskemaer, fordelt i den samlede sample såvel som fuldendte og ufuldendte besvarelser. Det kan i tabellen ses at der er en forskel på antallet af samlede antal besvarelser varierer med ti mellem de to skemaer.

	KMD	Soc-KMD	Total
Køn			
Mand	25	15	40
Kvinde	10	11	21
Region			
Nordjylland	21	16	37
Midtjylland	2	6	8
Syddanmark	5	1	6
Sjælland	1	1	2
Hovedstaden	6	2	8
Jobsituation			
Privat ansat	10	9	19
Offentlig ansat	3	4	7
Selvstændig	6	1	7
Arbejdssøgende	4	4	8
Under uddannelse	9	6	15
Udenfor arbejdsmarkedet	-	-	-
Ved ikke	1	-	1

Tabel 15.1.2 – BaggrundsvARIABLE

De overordnede baggrundsspørgsmål er i begge spørgeskemaer formuleret ens – og kan derfor let tabuleres som ovenfor. Der er i den indsamlede data en overvægt af mandlige respondenter, samt en klar overvægt at respondenter fra Nordjylland. Der er i begge spørgeskemaer omtrent lige mange kvindelige respondenter – hvor det i Soc-KMD resulterer i en mere ligelig fordeling mellem antallet af mænd og kvinder.

*Omdelingsmetoden*¹⁶ af spørgeskemaet har været en blanding af sociale medier, i form af både internetfora samt Facebook. Der opstår en selektionsbias, i og med at der i høj grad har været anvendt sociale medier til respondentrekruttering. De interne algoritmer, på Facebook, for sociale relationer og tilknytninger betyder det at en større andel af de respondenter der opnås kontakt til, vil være centreret omkring mit bopælsområde. Det synes derfor logisk at der vil opstå en overvægt af respondenter fra Nordjylland.

Hvad angår respondenternes jobsituation er der således tale om en høj koncentration af respondenter i den private sektor, såvel som respondenter der på nuværende tidspunkt er under uddannelse. Imens er offentligt ansatte og selvstændige sammenlagt repræsenterede med et samlet antal af syv respondenter i hver kategori – dog med den primære forskel at klart flest af respondenter med selvstændigt erhverv ganske tilfældigt er blevet præsenteret for KMD spørgeskemaet.

Med de ovenstående tabeller er baggrundsvariable kortlagt, hvorved der er dannet en oversigt over de indsamlede respondenter. Det er dog værd at bemærke at den data der er illustreret indtil nu er data som under normale omstændigheder også vil blive indsamlet, såfremt at det anvendte spørgeskemasystem tillader og lagrer delvise besvarelser.

15.1.2 Tertiære data i besvarelserne

Første trin i forbindelse med anvendelse af metoden, er at undersøge hvorvidt at nogle af de events og tertiære data der observeres og registreres i forbindelse med besvarelsesprocessen, forekommer i de to datasæt. Tid vil, selvsagt, altid være en faktor – men med henblik på de andre tertiære variable, kan vi således undersøge hvorvidt at der i spørgeskemaerne er forekommet svarskeft samt ændringer i progressionen. Det vil altså sige, om nogle respondenter i det ene eller andet spørgeskema har foretaget svarskeft og i så tilfælde hvor mange der er for hvert skema. Ligeledes er det værd at undersøge hvorvidt at progressionen for respondenterne udelukkende har været lineær, på den måde om de kun har klikket sig frem i spørgeskemaet eller om nogle af respondenterne har bevæget sig tilbage til et tidligere punkt i skemaet.

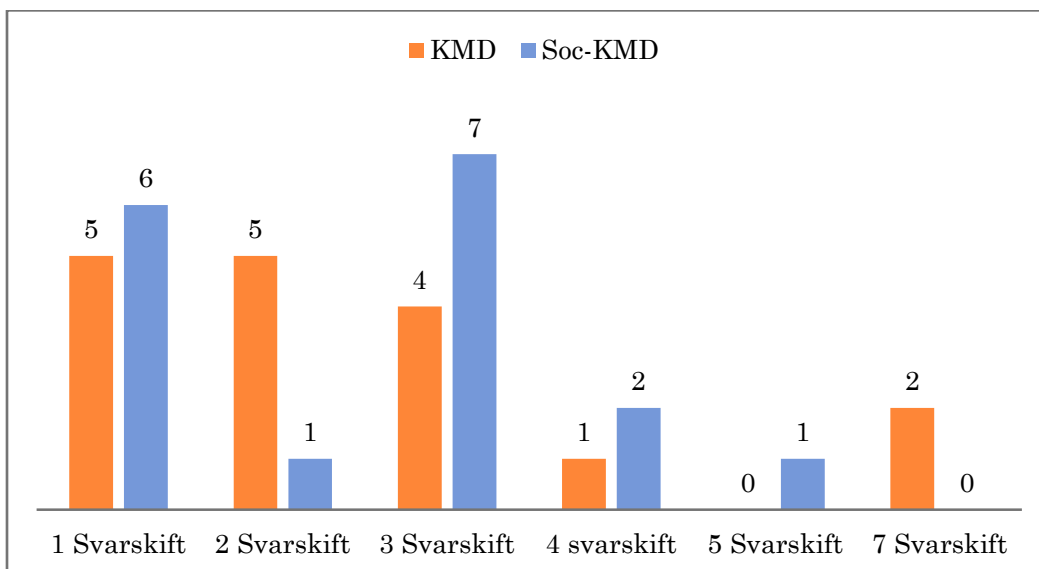
¹⁶ Skemaet er selvsagt ikke omdelt, på samme vis som ved post- eller mail-omdeling – men i stedet delt via forskellige medier.

Målet er således at undersøge hvorvidt at disse hændelserne forekommer, så der vil derfor ikke blive analyseret videre på eventuelle årsagsforklaringer eller lignende – men blot summeres op hvor ofte, hvis nogen, de forskellige forekommer i de respektive skemaer.

	KMD	Soc-KMD	Total
Svarskift	44 (17)	42 (17)	86 (34)
Progression	1 (1)	4 (3)	5 (4)

Tabel 15.1.2.1 – Tertiære variable i tal - (respondentantal i parentes)

Tabellen alene beretter kun om det samlede antal af svarskift og respondenter med svarskift i besvarelserne. Søger vi i stedet at undersøge nærmere, hvordan disse svarskift er fordelt, vil det i dette tilfælde kunne illustreres som følgende;



Figur 15.1.2.2 – Respondenter med svarskift (antal)

Vi kan på baggrund heraf aflæse at der er tale om en fordeling, hvor de fleste af de respondenter der har foretaget svarskift i højere grad vil have mellem 1 og 3 skift gennem deres forløb. Hvor kun enkelte har bevæget sig længere op – og hvor to respondenter i KMD skemaet har 7 svarskift hver. Det er værd at pointere, at spørgeskemaernes længde selvsagt vil kunne have indflydelse på antallet af svarskift – i det at der vil være flere spørgsmål at kunne skifte svar mellem. Dette udelukker dog ikke, at der kan foretages

indtil flere svarskift af den samme respondent i et enkelt spørgsmål – hvorved at dette vil have en væsentlig indflydelse på det endelige resultat.

Ud fra dette kan vi se, at der i begge spørgeskemaer forekommer en række svarskift i besvarelserne – hvilket giver årsag til at undersøge de underliggende problemstillinger. Imens der kun findes et fåtal af progressionsskifte, hvor respondenterne er gået tilbage til en tidligere side. Overordnet set kan vi således se at de formodede hændelserne forekommer i skemaerne – nogle oftere end andre. Det er på baggrund af den begrænsede samplestørrelse vanskeligt at konkludere hvorvidt der er tale om generelle tendenser eller om dette scenarie blot er en tilfældighed. Det bør i den forbindelse atter engang understreges, at ingen af respondenterne på nogen måde var klar over at selve tertiære data registreres og analyseres – hvorved at man bør kunne afvise at der er tale om et Hawthorne-lignende resultat, hvori at respondenterne søger at hjælpe med at understrege det undersøgte.

15.2 Respondentforløb

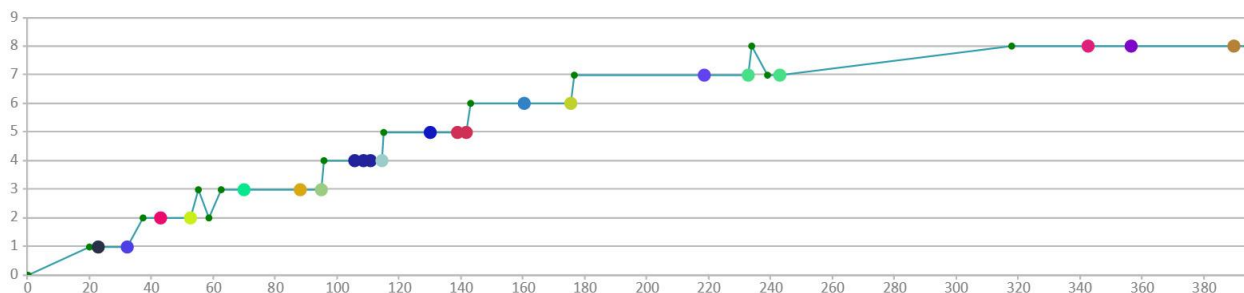
Som tidligere nævnt, er det muligt at undersøge respondenternes forløb mere nøje på baggrund af den indsamlede tertiære data. Afsnittet her har til formål at illustrere den mulige detaljegrad som kan opnås ved at dykke ned i enkelte respondentforløb. Formålet er altså at belyse i hvor høj grad, man er i stand til at reproducere og danne overblik over de enkelte forløb. Forhåbningen er at danne et afsæt, hvorfra der kan planlægges og etableres indsigt i mulige analysemodeller. Det enkeltstående forløb, er hvor en enkelt respondents besvarelse analyseres nøje. Denne metode er ikke den bedst egnede som udgangspunkt i forhold til en begyndelse på analysen af et større datasæt, men præsenteres i dette afsnit først for derved at introducere den grundlæggende idé bag.

15.2.1 Enkeltstående forløb

De indsamlede datasæts størrelse er relativt begrænset og i dette tilfælde er det derfor blot om at vælge en besvarelse der kan have interesse for på den måde at kunne illustrere de forskellige events i respondentens forløb – og hvorfor at netop denne søges undersøgt. Med udgangspunkt i den tidligere oversigt over de tertiære variable, vil det give mening at vælge en respondent som har de forskellige events i sit forløb, for på den måde at kunne illustrere de forskellige hændelser.

Ved at udvikle en model der analyserer forskellige tertiære events, er det derfor muligt at finde en respondent som matcher vores kriterier. Der er kun tale om et fåtal af tilfælde, hvor respondenterne går tilbage i undersøgelsen, dog er en enkelt af respondenterne ansvarlig for to af de fire progressionsskift der forekommer i Soc-KMD spørgeskemaet.

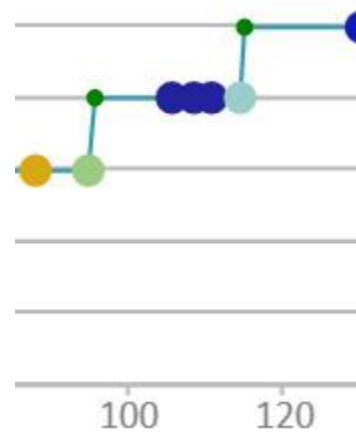
Hvis respondentens forløb visualiseres kunne det eksempelvis se således ud;



Figur 15.2.1.1 – Enkeltrespondentforløb uden legend.

Grafen forstås på den måde at y-aksen angiver den side med spørgsmål som respondenter befinder sig på, hvor side 0 er velkomstteksten, imens kan der af x-aksen aflæses tid i sekunder. Det vil altså sige at, ved at aflæse grafen kan man aflæse at respondenter bruger omtrent 20 sekunder på forsiden, formodentlig for at læse introteksten.

Hver boblerne i grafen er et event. Bobler der er knyttet til det samme spørgsmål, vil have den samme farve, som illustreret nærmere i figuren til højre. Der er i det givende tilfælde tale om at respondenter har skiftet svar indenfor et givent spørgsmål flere gange over en periode på cirka ti sekunder. Det er i den interaktive version muligt at holde musen over hvert enkelt af boblerne og se den gældende x-værdi/tid, hvilket spørgsmål det drejer sig om, såvel som hvilken svarmulighed respondenter har valgt, som ligeledes er illustreret til højre herfor. Der er således tale om at respondenter over en periode på ti sekunder har skiftet svar fra *positivt* til *hverken positivt eller negativt* og tilbage til *positivt*.



[105.518]: forhold-til udvikling:
4 - Positivt

Som det ligeledes ses i grafen herover, er der to forskellige typer af events. De store bobler relaterer til spørgsmål – imens at de små relaterer til progressions-events, såsom start, sideskift og afslutning. På samme måde som ved besvarelse af spørgsmål, er det således muligt at vide præcist hvornår respondenterne skifter side - samt hvilken side de skifter til. Der vil ved enkelte af disse events være knyttet bestemte typer af info – såsom *missing parameters*, hvilket betyder at respondenterne har forsøgt at skifte side – men ikke har udfyldt samtlige af de påkrævede felter, hvorved at de er blevet promptet af en fejl der informerer dem om hvilke felter de mangler. Ligeledes vil der ved *start*-eventet kunne ses en række af informationer, som er knyttet til respondenterens styresystem, skærmopløsning, browser og lignende. Disse informationer bevirker således at vi er i stand til at identificere at denne respondent har besvaret skemaet fra sin iPhone – hvilket muligvis gør os i stand til at sige mere om besvarelsesforløbet.

Grafen giver mulighed for at danne sig et hurtigt overblik over events i respondenterens forløb – og gør det muligt at skabe sig idé om hvornår der sker svarskeft eller progressionsskifte. Det kan dog være vanskeligt at skulle summere alle hændelser, således at de kan aflæses af grafen i papirformat – hvorfor at jeg har vedhæftet en tabel med oversigt over de forskellige hændelser i den enkelte respondentbesvarelse. Tabellen giver mulighed for at opnå en samlet indsigt i de specifikke hændelser i respondenterens forløb (som illustreret på næste side).

I tabellens første kolonne er der et ID som gør det lettere at identificere de forskellige events i forløbet. Næste kolonne er tid, målt fra start af skemaet til den gældende hændelse. Dernæst findes der i tabellen en række forskellige events i form af *start*, *movement*, *choice* og *end*. Start og end markerer henholdsvis begyndelsen og afslutningen af spørgeskemaet, hvilket betyder at alle der har påbegyndt besvarelse vil have en *start*-markør, imens kun de som har gennemført hele skemaet og klikket *Afslut* vil have en *end*-markør. Dette gør det hurtigt at søge igennem datasættet for at finde ud af hvorvidt at respondenterne har gennemført skemaet – eller afsluttet det præliminært. Choice er i stedet at respondenterne markerer et svar i skemaet, imens at movement relaterer til at respondenterne bevæger sig enten frem eller tilbage i skemaet.

ID	Tid	Type	Event / Spørgsmål	Værdi
0	0	Start	OS: iOS 7.0.6 Browser: FBCR Mobile: true Cookies: true Screen Size: 320 x 480	
1	19.744	Movement	Next	1
2	22.835	Choice	Køn	1 – Mand
3	32.127	Choice	Fødselsår	1988
4	37.079	Movement	Next	2
5	42.898	Choice	hvilken-region er du bosat i?	1 - Nordjylland
6	52.605	Choice	Arbejdssituation	4 - Arbejdssøgende
7	55.033	Movement	Next	3
8	58.439	Movement	Prev	2
9	62.555	Movement	Next	3
10	69.826	Choice	Computeregenskaber	1 - Altid let at anvende
11	88.089	Choice	Arbejdsudvikling	2 - Computere vil få større betydning
12	94.665	Choice	Smartphones	2 - Smartphones og tablets vil få større betydning
13	95.76	Movement	Next	4
14	105.518	Choice	forhold-til udvikling	4 – Positivt
15	108.396	Choice	forhold-til udvikling	3 - Hverken positivt eller negativt
16	110.553	Choice	forhold-til udvikling	4 – Positivt
17	114.455	Choice	smartphones-positiv-negativ	4 – Positivt
18	115.009	Movement	Next	5
19	130.025	Choice	Udviklingsforhold	1 - Udviklingen passer mig fint og jeg kan let følge med
20	138.849	Choice	teknologi-i fritiden	1 - Flere gange dagligt
21	141.571	Choice	teknologi-i fritiden	2 – Dagligt
22	142.934	Movement	Next	6
23	160.287	Choice	bliver-det offentlige mere effektiv	2 - Ja, den offentlige sektor bliver noget mere effektiv.
24	175.353	Choice	offentligt-bedre kvalitet	6 - Ved ikke
25	176.539	Movement	Next	7
26	218.566	Choice	offentligt-vs privat	3 - Den offentlige sektor og den private sektor er stort set på samme niveau
27	232.64	Choice	kontakt-til det offentlige	3 – Telefon
28	233.949	Movement	Next	8
29	238.84	Movement	Prev	7
30	243.029	Choice	kontakt-til det offentlige	1 - Digital Selvbetjening
31	317.908	Movement	Next	8
32	342.495	Choice	rigtige-beslutning at digitalisere	3 - Hverken/eller
33	356.351	Choice	behov-for-undervisning	2 - Ja, det er der behov for
34	389.724	Choice	offentlig-undervisning	3 - Ved ikke
35	394.138	End		

Tablet 15.2.1.1 – Enkelt Respondentforløb

Spørgsmål/info relaterer hvilket spørgsmål der er tale om – såvel som information om en given hændelse, eksempelvis i forbindelse med *movement* eventet, om det er en bevægelse frem eller tilbage i skemaet. Spørgsmålene som er listet, er i dette tilfælde noteret som

keyword-referencer, som er anvendt i forbindelse med udarbejdelsen af skemaet. Årsagen til at anvende de forkortede spørgsmål er udelukkende for at spare på mængden af data der videregives af respondentens browser. Derudover har start-eventet en række informationer vedhæftet om respondentens browser, som tidligere omtalt, hvilket ses øverst.

Sidst i tabellen er '*Værdi*', som rummer svar der er afgivet af respondenterne eller information om progression. Det vil sige, at når respondenterne går enten frem eller tilbage i skemaet vil værdi fortælle om hvilken side respondenterne går til – hvilket kan ses efter cirka 19 sekunder i event ID 1, hvor respondenterne går videre i skemaet for første gang.

Udgangspunktet for at vælge denne respondent som eksempel, var fordi at han som den eneste har flere progressionsskifte – og heldigvis har respondenterne også nogle svarskift, så der er tale om et godt eksempel til at kunne illustrere en række forskellige scenarier der kan opstå i forbindelse med tertiære variable. Ud fra de indsamlede data ved vi at respondenterne har besvaret skemaet fra sin mobiltelefon – igennem den browser der er integreret i Facebook applikationen (FBCR).

Respondentens forløb i besvarelse af baggrundsspørgsmålene er lineært, i den forstand at der ikke er noget i forløbet der skiller sig ud. Det tager ikke respondenterne bemærkelsesværdigt lang tid at besvare spørgsmålene – alle spørgsmål besvares inden for nogle få sekunder af hinanden og der er ingen uventede pauser eller svarskift. Efterfølgende, i events 7, 8 og 9, går respondenterne videre til næste side med spørgsmål, tilbage til forrige side og frem til den næste side atter engang. Hver bevægelse varer mellem tre og fire sekunder – og det er nu vi kan forsøge at belyse en eventuel årsag herfor. Det visuelle design af skemaet bevirker, at skemaet ikke er unikt optimeret til mobiltelefon. Det betyder at skemaet ser ensartet ud på mobil og computer og ikke vises som en decideret *mobiludgave*. Det er derfor tænkeligt at respondenterne kan komme til at klikke forkert, i det at berøringsfølsomme skærme gør det lettere at lave en decideret klikfejl og at det visuelle design ikke nødvendigvis akkommoderer dette ved at anvende markant større navigationsknapper. Rent fysisk vil der i dette tilfælde være tale om at respondenterne først trykker på en knap i højre side af skærmen, for derefter at skulle trykke på en knap i venstre side af skærmen for at gå tilbage i skemaet. Det synes derfor ikke som en sandsynlig mulighed i dette tilfælde – men i stedet noget der formodes at være foretaget

intentionelt, for at genlæse eller bekræfte de svar respondenterne havde afgivet. Havde der være tale om et fejlklik, må det formodes at sandsynligheden for at respondenterne kommer til at klikke videre i skemaet to gange i træk synes mere plausibelt, end at de først skulle klikke i den ene og derefter den anden side af skærmen med nævneværdige mellemrum. Havde det været tilfældet at respondenterne havde klikket videre to gange i træk grundet den berøringsfølsomme skærm, ville der i stedet have været lagret information om at respondenterne havde haft *missing parameters*, i deres *movement*. I og med at de derfor ville mangle at udfylde spørgsmål i det indeværende spørgsmålsbatteri. Dog er vi ikke i stand til at identificere nøjagtigt hvad respondenterne foretager sig, i og med at respondenterne ikke foretager sig noget aktivt med selve skemaet – og det for os ikke er muligt at identificere hvor respondenterne har sit blik rettet imod.

Ved at opnå viden om det medie som respondenterne besvarer spørgeskemaet gennem, bliver det muliggjort for forskeren at kunne få yderligere indblik i respondenternes adfærd. Som illustreret må det formodes at der forekommer scenarier og derigennem problemstillinger som er unikke for de forskellige medier - dette er dog en undersøgelsesfase i sig selv og vil ikke blive uddybet yderligere gennem projektet. Det bør dog inkluderes at måden hvorpå spørgeskemaet omdeles ligeledes vil kunne have indflydelse på hvordan respondenterne besvarer det. Hvis spørgeskemaet søges omdelt via Facebook, bør det inddrages i overvejelserne, at op imod 65 % af al trafik på Facebook foregår fra mobil (Facebook 2014). Det må derfor ligeledes formodes at en del af de respondenter som vælger at besvare spørgeskemaet vil gøre det gennem deres smartphone – og med en årlig vækst af online trafik via mobil på omtrent 55 % årligt, må det ligeledes formodes at der vil være tale om en voksende tendens (NetmarketShare 2014)(Cisco 2014).

Hvis vi atter engang dykker ned i besvarelsen, ses det at respondenterens svar på side 3 forekommer lineært og uden afbrydelser eller svarsift – det er dog værd at bemærke at hvis man sammenligner med spørgsmålene vedrørende baggrundvariable, ses det at de holdningsorienterede spørgsmål tager respondenterne længere tid at besvare. Dette må formodes at skyldes at spørgsmålene er længere og knapt så standardiserede som de tidligere. Forstået på den måde, at baggrundvariable ofte er meget ensartede og at genkalde fakta såsom bopælsregion og alder, vil formodes at være hurtigere end at skulle genkalde en holdning.

På side 4 oplever vi respondentens første svarskifte i events 14, 15 og 16 i spørgsmålet vedrørende respondentens holdning til den teknologiske udvikling. Det ses at respondenterne vælger mellem svarmulighederne *Positivt* samt *Hverken positivt eller negativt*, for derefter at vælge svaret positivt. Der er cirka 3-4 sekunder mellem hver korrektion af besvarelsen, hvilket atter engang henleder opmærksomheden på at det ikke formodes at der er tale om et fejltryk, men en reel overvejelse fra respondentens side.

På næste side oplever vi atter engang svarskifte i events 20 og 21, i spørgsmålet om i hvor høj grad at respondenterne anvender teknologi i fritiden – hvor han først angiver at der er tale om noget han *gør flere gange dagligt*, for derefter at korrigerer det til *dagligt*. Der kan være tale om en refleksion fra respondenterne – om deres egentlige brug. Der kan ligeledes være tale om social desirability, i det at der fortsat kan være et hvis stigma tilknyttet folk der bruger meget tid med deres computer. Præcis hvilken af de to hændelser der er tale om, er det ikke muligt for os at identificere – dog giver det yderligere belæg for at der er tale om en aktiv kognitiv handling fra respondenterne. Forstået på den måde, at der er tænkt over hvad der bliver spurgt om i spørgeskemaet – såvel som deres position i forhold hertil.

Besvarelsen forløber lineært derfra og frem til events 27 til 31, hvor respondenterne bliver spurgt om hvilken kontaktform de primært anvender når de tager kontakt til de offentlige. Respondenterne vælger først at telefon er måden hvorpå de normalvist retter henvendelse, for derefter at gå videre til næste side. Der går herefter et par sekunder, hvorefter at respondenterne går tilbage til den forrige side og korrigerer sit svar på spørgsmålet til, at deres primære kommunikationsform med det offentlige i stedet foregår via digital selvbetjening, for derefter at fortsætte i spørgeskemaet, efter at have haft et kortvarigt ophold på knapt 25 sekunder på siden hvor han netop havde rettet svar.

Det er yderst interessant at se at den adfærd forekommer – i det at respondenterne mellem spørgsmålene reflekterer over deres egentlige position, for derefter at gå tilbage og rette svaret. Havde respondenterne ikke haft mulighed for at gå tilbage i skemaet, ville vi under normale omstændigheder ikke have adgang til det svar som respondenterne mener, repræsenterer deres erfaringer. Dette kan være frustrerende for respondenterne – men ligeledes også for forskeren, i det at man derfor ikke får det korrekte billede på det der søges undersøgt. Det skal selvfølgelig pointeres at respondenterne i denne undersøgelse

kun ganske få gange går tilbage i skemaet. Grundet den meget lille stikprøve, er det ikke muligt at belyse hvorvidt at der er tale om en generel tendens eller om det i større undersøgelser forekommer sjældnere eller oftere.

Resten af besvarelsen forekommer for denne respondent lineær og uden afbrydelse eller bemærkelsesværdige pauser eller lignende.

Denne respondenten har været et yderst interessant tilfælde og har illustreret en række af de forskellige situationer som kan opstå i forbindelse med indsamling af tertiære data. Overordnet set, ser det ud til at respondenten har besvaret skemaet på en sådan vis – at det må formodes at der er lagt refleksion og overvejelser bag svarene. Dette med baggrund i at respondenten har brugt tid på at gå tilbage i skemaet for at sikre sig at de svar der har været afgivet har været korrekte – og tænkt over spørgsmålene, selv efter at det første svar er afgivet. Udover et enkelt stop på knapt 25 sekunder, har der været en relativ forventelig rytme i besvarelsene, forstået på den måde at der er gået tilsvarende længere tid med længere spørgsmål, sammenlignet med den kortere tid i korte/hurtige spørgsmål.

Der er mange muligheder for at illustrere de indsamlede data, valget faldt på disse former da de gjorde det muligt hurtigt at opnå indsigt og oversigt over den enkelte respondents forløb – såvel som en dybdegående indsigt gennem tabellen.

Som nævnt er formålet med afsnittet at illustrere enkeltrespondentforløb, for at skabe indsigt i hvad en besvarelse rummer, samt hvordan disse kan visualiseres, læses og tolkes. Dette er derfor ikke forventeligt at der er tale om et startpunkt for en eventuel undersøgelse, men en af de sidste led såfremt man vælger at dykke ned i de enkelte respondenters besvarelser for derved at opnå indsigt heri.

15.3 Designforløb

Som tidligere omtalt, er en af underspørgsmålene der er knyttet til metoden at den evner at belyse problemer der vedrører spørgeskemaets design – i den forstand at selve den måde som skemaet er sammensat på bevirker at nogle respondenter besvarer spørgsmålet anderledes eller har anderledes adfærd i forhold til hvad der egentlig var forventet.

Der er en række af muligheder for at gå til værks, når man søger at undersøge problemer i designet, hvilket selvfølgelig også skyldes at design spænder over en lang række elementer. Der kan være tale om spørgsmål der er formuleret forvirrende eller svarmuligheder, hvor det ikke er muligt for respondenterne at finde en svarmulighed som er repræsentativ for deres holdning eller erfaring. Ligeledes kan der være tale om at svarmuligheder der er valgt eventuelt er præsenteret i en skala, som ikke giver mening for respondenterne eller at de eventuelt forventer at der er tale om en anden type af svarmuligheder end der egentlig er; forstået på en sådan vis at hvis respondenterne forventer at der er tale om et spørgsmål der tillader multiple svar, men at designet i stedet kun tillader at man vælger en enkelt kategori.

Den grundlæggende idé er at henvise forskerens opmærksomhed til områder, som bør undersøges nærmere. Det ville være fordelagtigt såfremt at alle analyseprocesserne kunne automatiseres, dog er det på nuværende tidspunkt ikke data eller viden nok til at gøre dette – hvilket overlader det analytiske arbejde til forskeren selv.

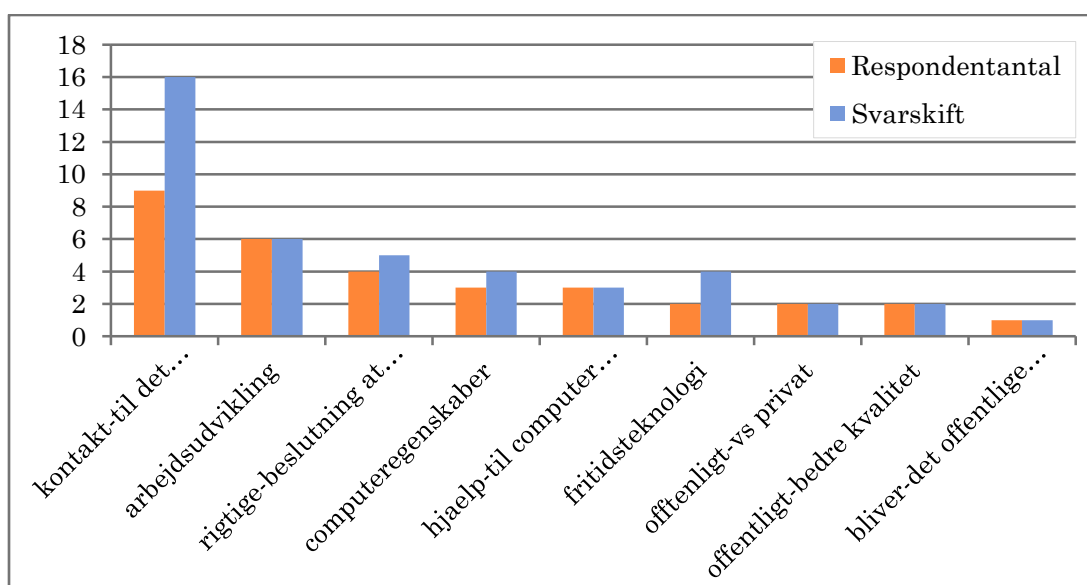
Fejl og mangler i spørgeskemaet er således en refleksion af respondenternes besvarelser og adfærd. Det er derfor heller ikke et forventeligt scenarie at samtlige fejl i skemaet fremhæves, men i stedet at der på baggrund heraf er større tendens til at belyse mulige problemområder. Der vil i denne del af analysen søges at afskærme sig fra komplekse modeller, hvor kombinationer af tertiære data anvendes til at skabe en samlet model – for i stedet at belyse hvad de forskellige tertiære data evner at henlede til, som enestående værdi.

15.3.1 Svarkift som henledende faktor

En af mulighederne for at blive opmærksom på mulige komplikationer i spørgeskemaets design, er ved at anvende eventet *svarkift* til at fremhæve eventuelle problemområder. Ud fra tabel 15.1.3.1, samt figur 15.1.3.2, kan vi se at der er et næsten identisk antal svarkift i besvarelserne – og at disse ligeledes er fordelt over samme antal respondenter, dog med forskellig fordeling hvad antal af svarkift angår. Men det i sig selv, kaster ikke lys over spørgeskemaets design.

Ved i stedet at kigge på mængden af svars kift i de forskellige spørgsmål i besvarelserne, bliver det væsentligt lettere at blive opmærksom på eventuelle problemområder. Det er dog vigtigt at der ikke blot bides mærke i mængden af svars kift, men også antallet af respondenter det er fordelt på. En af årsagerne hertil er, at der kan i et enkelt spørgsmål forekomme mange svars kift fra en enkelt respondent. Eksempelvis, hvis en respondent navigerer igennem listen med fødselsdatoer ved at benytte piletasterne på tastaturet, vil det i systemets nuværende form registreres som en hel række svars kift fra start-årstal til deres valgte årstal. Såfremt at det er tilfældet, vil der således være tale om en skævvridning i det at et enkelt spørgsmål vil have en overvægt af svars kift, i de enkelte spørgsmål, fordelt på meget få respondenter.

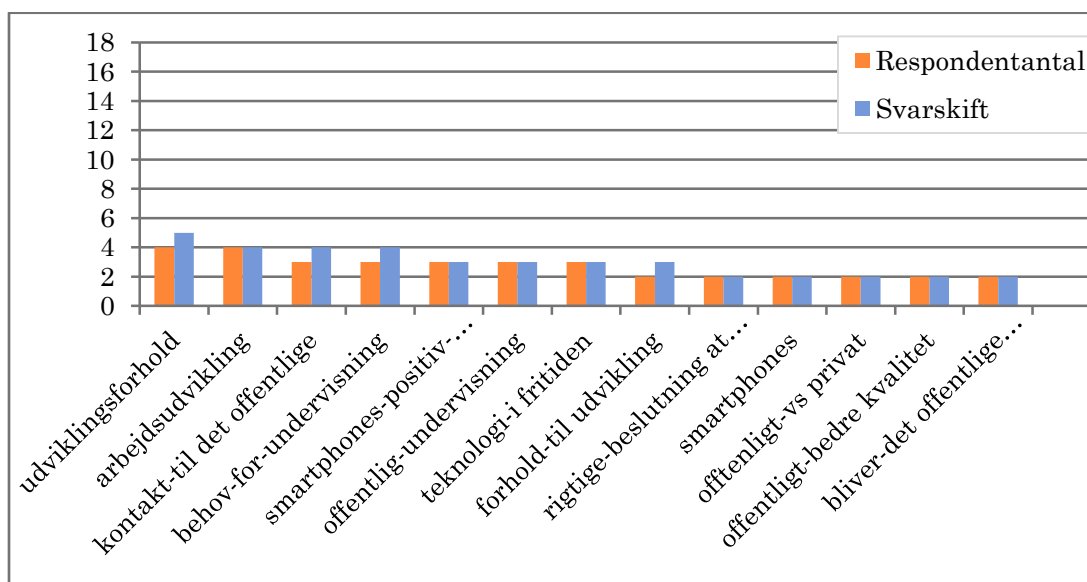
I forbindelse med bearbejdelsen af besvarelserne i de to datasæt, er der udviklet en algoritme som løbet samtlige besvarelser igennem og registrerer respondenternes svars kift i de enkelte spørgsmål. Efterfølgende er de få svars kift som er forekommet i spørgsmål vedrørende baggrundsdata blevet filtreret fra, i det at der kun var tale om et fåtal som fremstod som deciderede fejlklik – hvilket ligeledes gavner præsentationen af grafen, i det at mængden af data reduceres.



Figur 15.3.1.1 – Svars kift og respondenter med svars kift i KMD-skemaet

Som tidligere nævnt, er en række af spørgsmålene i Soc-KMD skemaet videreudviklinger og omstruktureringer af spørgsmål fra KMD skemaet. Det betyder at nogle af spørgsmålene kun er ændret i så lille en grad at de ville kunne paralleliseres, med henblik på at belyse eventuelle designmæssige fejl.

Der er dog fortsat en større forskel mellem især det spørgsmål som er blevet omstrukturerede fra et til to spørgsmål, som har været tilfældet for flere af spørgsmålene. Disse forskelle er ligeledes årsag til at det ikke har været muligt at samle begge skemaer i samme graf. I stedet er graferne præsenteret i samme skala, for på den måde visuelt at illustrere omfanget af forskellene i svarskeft, som illustreret på næste side.



Figur 15.3.1.2 – Svarskift og respondenter med svarskift i Soc-KMD

Ved at aflæse grafen for KMD spørgeskemaet, ses det at spørgsmålet med absolut flest svarskift er spørgsmålet "Hvordan kommer du i kontakte med det offentlige?"¹⁷. Samme spørgsmål er i Soc-KMD skemaet ligeledes blandt de spørgsmål med højest svarskift, her er formuleringen i stedet "Hvordan kommer du primært i kontakt med det offentlige?"¹⁸.

Begge spørgsmål præsenterer en række identiske svarmuligheder, strækkende fra digital selvbetjening, telefon, personligt fremmøde til at respondenter aldrig har været i kontakt

¹⁷ Forkortet som kontakt-til-offentlig

¹⁸ Forkortet som kontakt-til det offentlige

med det offentlige. Det er kun muligt for respondenter at markere en enkelt svarmulighed på listen over kontaktmuligheder.

I KMD spørgeskemaet forekommer der sammenlagt 16 svarskeft i denne kategori, fordelt på 9 respondenter, mens der i Soc-KMD kun forekommer 4, fordelt på 3 respondenter. Det er værd at bemærke at understrege at spørgsmålet i Soc-KMD lægger vægt på at der er tale om den *primære* kontaktform, mens det oprindelige spørgsmål ikke på samme måde illustrerer dette. Ved at gennemgå en detaljeret oversigt over besvarelserne, kan det aflæses at der ved flere af svarskeftende i KMD skemaet er tale om at respondenterne markerer én mulighed, skifter til en anden for derefter at vælge det oprindelige svar igen.

Som illustreret i tabel 15.3.1.3 herunder;

Respondent	Svarskeft
15	1 - Digital Selvbetjening 2 - E-mail 1 - Digital Selvbetjening
16	1 - Digital Selvbetjening 3 - Telefon 1 - Digital Selvbetjening
21	1 - Digital Selvbetjening 2 - E-mail
35	1 - Digital Selvbetjening 4 - Personligt Fremmøde
37	1 - Digital Selvbetjening 3 - Telefon
44	2 - E-mail 1 - Digital Selvbetjening 2 - E-mail
46	3 - Telefon 4 - Personligt Fremmøde 3 - Telefon 4 - Personligt Fremmøde
50	1 - Digital Selvbetjening 7 - Har aldrig været i kontakt med det offentlige 2 - E-mail
62	3 - Telefon 1 - Digital Selvbetjening 3 - Telefon

Tabel 15.3.1.3 – Detaljeret svarskeft i kontakt-til-offentlig - KMD Spørgeskema

Anerledes er der kun en enkelt respondent i Soc-KMD skemaet, som efterfølgende har valgt sit oprindelige svar igen¹⁹. Det ses i de besvarelser, hvor respondenterne ikke har skiftet til sit oprindelige svar at deres andet svar befinder sig længere nede på listen over svarmuligheder. Hvilket rejser spørgsmålet ved, om respondenterne har været opmærksomme på at de kun kunne vælge en enkelt svarmulighed og at der er foretaget et egentlig svarskeft. Der kan ligeledes være tale om at respondenterne ikke har været opmærksomme på at markeringen fra deres første svar er blevet fjernet, i det at de valgte et svar længere nede på listen.

Der er, som nævnt, en forskel i formuleringen af de to spørgsmål, hvoraf det som er mere specifikt vedrørende mulighederne for svar, har væsentligt færre svarskeft end det som ikke har udspecificeret antal. Det er yderst vanskeligt at belyse hvorvidt at der er tale om at respondenterne genkender formen i spørgsmålet og antager at der er tale om et spørgsmål der tillader multiple svar – eller om forskellen i antallet af svarskeft, skyldes forskellen i formuleringen af de to spørgsmål. Hvorvidt at dette har nogen reel betydning eller at der er tale om et tilfælde er vanskeligt at afgøre – det essentielle er blot at påpege at det er den eneste forskel i spørgsmålenes formulering.

Dette skaber stor usikkerhed omkring data der er indsamlet fra spørgsmålet, i og med, at det må antages at når frekvensen af svarskeft tilbage til den oprindelige svarmulighed er så høj, tyder det på at respondenterne har troet at de kunne vælge flere svar. Normalvis vil der i det tilfælde være en annotation eller formulering der indikerer muligheden for flere svarmuligheder, såsom "*Sæt gerne flere krydser*".

De mange svarskeft bevirker ligeledes at der kan sættes spørgsmålstegn ved de besvarelser, hvori respondenterne *ikke* har valgt sit oprindelige svar igen, men i stedet et svar længere nede på listen. Det må formodes at der kan være tale om at respondenterne ikke har været opmærksomme på at markeringen blev fjernet fra deres tidligere svar. Der kan ligeledes være tale om at respondenterne besvarer som de læser skemaet og faktisk har valgt den mulighed der repræsenterer deres erfaringer bedst. Dette synes umiddelbart ikke muligt at identificere hvorvidt at der er tale om den ene eller den anden form for besvarelse – hvilket

¹⁹ Ikke illustreret.

giver forskeren mulighed for at overveje pålideligheden af de indsamlede data, hvorved at man kan vælge at udelade disse fra en eventuel analyse eller konklusion.

Spørgsmålet med næst-flest svarskeft på tværs af de to skemaer er spørgsmålet vedrørende arbejdsudvikling i forhold til computere og smartphones. Som nævnt i et tidligere afsnit, er dette et af de spørgsmål som i Soc-KMD skemaet er blevet brudt op i to, hvor det ene rummer spørgsmål vedrørende computerens påvirkning på respondentens arbejdssituation og det andet vedrørende påvirkningen af smartphones og tablets.

Opdeling af disse spørgsmål bevirker at disse svarskeft i projektets eget skema er delt over to kategorier, hvilket ligeledes er årsagen til at diagrammerne er opdelt for hvert spørgeskema. Dette skyldes vil være vanskeligt at parallelisere spørgsmål der ikke er formuleret ens. Spørgsmålene er i diagrammerne noteret som henholdsvis *arbejdsudvikling* og *smartphones*, hvoraf den sidste kun fremgår i undersøgelsens eget skema.

I KMD skemaet er der forekommet 6 svarskeft i spørgsmålet – og ved at undersøgere disse nærmere, er vi i stand til at identificere at samtlige af disse svarskeft er skiftet til et svar senere i listen. Det vil altså sige at respondenterne i dette tilfælde først har valgt en svarmulighed der lå højere på listen over svarmuligheder end den de endte med. Spørgsmålene og svarmulighederne forventes ikke at kunne mistolkes ud fra en typeforståelse, som var tilfældet i tidligere eksempel. Svarmulighederne er opdelt i en skala der strækker sig fra at Computere og Smartphones vil få meget større betydning til at de slet ingen betydning vil få.

Fælles for alle svarskeft i dette spørgsmål er at det svar respondenterne er skiftet til, er svaret der befinder sig lige under det først valgte. Det vil altså sige, at de respondenter der først valgte svar nummer et; ”Computere og smartphones vil få meget større betydning”, efterfølgende har valgt næste svar; ”Computere og smartphones vil få større betydning”. Samme adfærd gør sig gældende for respondenter der har valgt svar der befinder sig tættere på midten, hvorefter at de har valgt et svar lavere.

Iagttagelse af de to affødte spørgsmål i undersøgelsens eget spørgeskema nærmere, kan vi se at de udviser nøjagtigt samme tendens, hvor respondenterne først vælger en svarmulighed der ligger øverst på listen, for derefter at skifte til svaret lige under. Dette

kan selvfølgelig have rod i at respondenterne foretager en moderat satisficing, hvori at de stiller sig tilfredse med det første svar de støder på – for derefter at korrigere det til et svar der passer bedre når de læser det.

Det kan være yderst vanskeligt, uden at foretage en specifik undersøgelse, at konkludere hvorvidt at der er tale om satisficing eller om der i stedet skulle være tale om en anden form for respondentadfærd. Alternativt kunne man i en anden undersøgelse lave et identiske skalaer, men hvor respondenterne blev præsenteret for at skalaen kunne være enten højst først eller lavest først. Det vil dog formodes at der selvfølgelig kan være en uønsket effekt, såfremt at denne rækkefølge ikke er ensartet gennem hele respondentens besvarelsesforløb – hvorved at skiftende rækkefølge kan virke forvirrende eller frustrerende for respondenterne. Der vil derfor i stedet være tale om at nogle respondenter præsenteres med en gennemgående højst først rækkefølge, imens andre bliver præsenteret for en lavest først. Det vil på den måde være lettere at belyse hvorvidt at der er tale om at respondenter vælger svarmuligheder som de læser dem – eller om de generelt har en tendens til at korrigere ned i svar. Det bør dog nævnes, at der i Soc-KMD besvarelserne, som rummer de to spørgsmål, kun er en enkelt af de der har foretaget svarskift som er genganger i begge. Det betyder altså, at det er forskellige respondenter der har svarskift i de to spørgsmål.

Som udgangspunkt synes dette dog at være et spørgsmål om respondentadfærd og ikke noget der umiddelbart synes at kunne spores tilbage til spørgeskemaernes design. Spørgsmålene rummer ikke ekspertviden og der bliver kun spurgt til en holdning eller formodning, som er fordelt ud over flere spørgsmål – og der forekommer omtrent lige så mange svarskift i alle af de tre spørgsmål. Alle tre spørgsmål er formuleret ud fra samme *formel* og er struktureret og opsat ens – hvilket selvfølgelig kan give anledning til at overveje, hvorvidt en anden skala havde givet et andet resultat – eller om resultatet der er opnået ikke er ønskeligt. Forstået på den måde, om der er noget i de svarskift som indikerer at der skulle være tale om et uønskeligt resultat eller om det antal af svarskift og måden de er forekommet på, vil formodes at tilvejebringe viden om respondenternes holdninger.

På samme måde som i ovenstående spørgsmål, blev spørgsmålet *”Synes du, at den offentlige sektor skal gøre mere for at hjælpe borgerne med at anvende computere og smartphones?”* fra KMD skemaet omformuleret og omstruktureret til to spørgsmål. Årsagen hertil var at det oprindelige spørgsmål ikke mulige svarkategorier der matchede det der blev spurgt, hvorved at enkelte holdninger blev udelukket. Netop dette scenarie blev jeg kontaktet vedrørende, af en enkelt respondent. Respondenten nævnte at de ikke følte at der en svarmulighed i spørgsmålet, som repræsenterede af deres holdning. Det var derfor forventet inden udsendelse af spørgeskemaer, at dette spørgsmål ville forvolde problemer eller give anledning til svarskeft fra respondenterne. Dette viste sig dog ikke at være tilfældet, men i stedet blev lige mange svarskeft fordelt ud over alle spørgsmålene, hvor både det ene spørgsmål i KMD skemaet, såvel som de to spørgsmål i Soc-KMD skemaet hver især havde henholdsvis tre og fire svarskeft. Spørgsmålet vedrørende behovet for undervisning i Soc-KMD skemaet er det spørgsmål med flest skift. Denne mulige designfejl er således ikke reflekteret tydeligt udelukkende i respondenternes svarskeft, hvorved at man kan søge til en kombination af tertiære data i stedet.

Det bør ligeledes pointeres at der er tale om en enkelt holdning, som ikke nødvendigvis vil være repræsentativ for ret mange. Det forventes derfor at kun et fåtal af respondenter vil studse over den manglende mulighed. Havde det være muligt at akkumulere en større sample i begge spørgeskemaer, formodes det at det ville have været lettere at kunne belyse om der er tale om gennemgående tendens vedrørende spørgsmålet.

15.3.2 Events som henledende faktor

På samme måde som når forskerens opmærksomhed rettes mod enkelte spørgsmål grundet antal af svarskeft, er det muligt at undersøge hvorvidt at respondentens sidste handling kan være indikativ på en mulig problemstilling. Tanken er at hvis en respondent på et tidspunkt støder på et spørgsmål som ikke giver respondenterne mulighed for at afgive svar der repræsenterer deres holdning eller erfaring, eller et spørgsmål på anden vis er uønsket, grænseoverskridende eller uinteressant, kan bevirke at respondenterne vælger at stoppe besvarelsen.

Ved at undersøge om der respondenterne stopper deres besvarelse på en given side, er det muligt at blive opmærksom på sider med spørgsmål som bør undersøges nærmere. Grundet

den måde hvorpå events registreres og overføres til databasen, vil det sidste event for en respondent altid være enten starten på skemaet, et sideskift, manglende udfyldelse af spørgsmål²⁰ eller en afslutning på undersøgelsen. Dette skyldes at data overføres fra respondenten besvarelse i det øjeblik de foretager en af disse progressions events. Det betyder derved at det ikke er os muligt at se hvad respondenten foretager sig – om de eventuelt har afgivet svar inden skemaet lukkes. Denne problemstilling adresseres i et senere afsnit.

Fra tabel 15.1.1 kan vi aflæse at der i KMD undersøgelsen forekommer et langt større frafald af respondenter end i projektets eget spørgeskema. Respondentfrafaldende vil derfor se således ud;

Event	Side	Antal
end		28
next	missing_param	2
next	1	1
next	3	5
next	4	2

Tabel 15.3.3.1 – Respondentfrafald - KMD

Event	Side	Antal
end		25
next	missing_param	1
next	1	1
next	7	1

Tabel 15.3.3.2 – Respondentfrafald – Soc-KMD

Opsætningen i tabellerne ovenfor er opdelt således at der i venstre kolonne er angivet hvilken type event der forekommer, midterste kolonne angiver side imens den sidste kolonne angiver antallet af respondenter som er stoppet ved den gældende event. De forskellige typer events er, som nævnt, *end*, *next* og *prev*. Dog er der ingen respondenter der i undersøgelsen er gået tilbage til en tidligere side og derefter afsluttet undersøgelsen. Årsagen hertil formodes at der kun er et fåtal af respondenter der har gået tilbage i spørgeskemaet og at de fleste skift tilbage, vil formodes at være knyttet til en intention, såsom at kontrollere sine svar eller ændre svarmulighed – hvilket er illustreret i tidligere afsnit. Forskellen i proportionen af frafald er ligeledes testet, hvilket dog viste sig ikke at være signifikant med en *p*-værdi på 0,115 – selvom det er tæt på.

End-markøren angiver respondenter hvis sidste event var afslutning af undersøgelsen, imens at next i stedet angiver at deres sidst registrerede handling, samt den side de har været skiftet til. Ud fra dette kan vi aflæse at der i KMD undersøgelsen har været en overvægt af respondenter der har afsluttet undersøgelsen efter at have nået side tre.

²⁰ I tabellerne angivet som *missing_param*.

Spørgeskemaet er struktureret således at side 0 er en introduktion til undersøgelsen, hvorefter side 1 og 2 består af spørgsmål vedrørende baggrundsvARIABLE, såsom fødselsår, køn, bopælsregion og jobsituation. Side tre er den første side hvor respondenter oplever holdningsspørgsmål. Spørgsmålene på siden er;

”Tror du, at den offentlige sektor kan blive mere effektiv og spare penge ved at anvende digitale løsninger og selvbetjening?”

”Tror du, at den offentlige sektor kan øge kvaliteten af de offentlige serviceydelser ved at anvende digitale løsninger og selvbetjening?”.

Det første af de to spørgsmål er i undersøgelsens eget spørgeskema omformuleret, i det at der er tale om et *double-barrel* spørgsmål. Det vil sige at der i spørgsmålet bliver spurgt om to ting samtidigt, men svarmulighederne kun giver mulighed for at respondenter kan besvare det ene af disse spørgsmål. I dette tilfælde er der tale om at spørgsmålet vedrører om det offentlige kan blive mere effektivt samt et spørgsmål vedrørende om der ligeledes kan spares penge ved at anvende digitale løsninger og selvbetjening.

Det er muligt at argumentere for at effektivisering og økonomisk fordel kan kobles sammen – imens at det for andre må formodes at være adskilte begreber, i det at man godt kan være mere effektiv uden at det samtidig bliver mere økonomisk. Det er dog værd at påpege at der i undersøgelsens eget spørgeskema ikke ses tendens til respondentfrafald ved dette spørgsmålsbatteri.

Med de indsamlede data er det ikke muligt at identificere hvorvidt at der er tale om at spørgsmålene har haft effekt på respondentfrafaldet, i det at der kun er registreret op til sidste progressions event. Det er ligeledes tænkeligt at respondenterne bliver stillet spørgsmål, som de ikke tidligere har haft holdning til eller simpelthen ikke ønsker at deltage længere, af ren og skær mangel på interesse. Det er dog yderst tankevækkende adfærd, som undersøges nærmere i et senere afsnit, såvel som problemstillingen med ikke at være bekendt med handlinger på sidste side vil diskuteres og præsenteres en mulig løsning herpå i et af de afsluttende afsnit.

15.3.3 Tid som henledende faktor

Tid kan være en vanskelig værdi at arbejde med, i det at der ikke er fastlagte grænser for hvor længe det bør tage at gennemføre et spørgeskema, i det at læsehastighed, beslutningsevne og fokus varierer fra person til person. Det er derfor ikke muligt at fastlægge rammer for hvordan dette nødvendigvis håndteres bedst muligt. Dette afsnit har til formål at forsøge at belyse nogle forskellige muligheder som kan skabe indsigt i mulige komplikationer i undersøgelsesdesign.

For at skabe overblik over respondenternes forløb som helhed, ses der i ovenstående tabel en samling af alle respondenters besvarelser i Soc-KMD skemaet. Progression er i tabellen noteret som eksempelvis *next-1*, hvilket refererer til retningen såvel som den side der skiftes til. Det vil altså sige, at hvis der står *prev-0*, vil respondenteren have gået tilbage fra side 1 til side 0 – som er introduktionsteksten. Imens er events som forekommer yderst sjældent og som kun opleves for et fåtal respondenter placeret nederst i tabellen, for derved fremme læsevenligheden og den lineære opsætning af spørgsmål. Der er her tale om events som skift til forrige side – såvel som ”*missing_param*”, hvor respondenteren mangler at udfylde alle spørgsmål ved sideskift.

På grund af de tertiære data der indsamles i forbindelse med respondenternes besvarelser, har det været nødvendigt nøje at overveje, hvordan disse events håndteres i forbindelse med illustration og tabulering af data. Det betyder at der ikke blot skal tages højde for respondentens svar, men også de events der støder op til. Sammenlignet med de *traditionelle* metoder, hvor der vil være tale om lineære forløb, hvor forskeren kun er bekendt med slut-svaret, hvorved at der ved hver spørgsmål kun vil være maksimalt et event per respondent.

Besvarelser hvori respondenteren har et eller flere svarskeft i samme spørgsmål er således blevet slået sammen til at tælle en enkelt besvarelse, hvorved at den tid de har anvendt på spørgsmål er indgået i regnskabet som én samlet tidsmåling. Argumentation her bag er at der er tale om at opnå indsigt i den tid der er anvendt på de enkelte spørgsmål – og at der igennem analysen af de forskellige data har været tendens til at respondenterne bruger væsentligt længere tid på første besvarelse af spørgsmålet, hvorefter at svarskeft forekommer relativt kort tid herefter. Dette skyldes selvfølgelig at der forud for afgivelse af

svar, først skal læse selve spørgsmålet og de efterfølgende svarskift vil som oftest være en korrektion af holdning og kræve mindre læsning.

Event	Median	Mean-tid	Min. tid	Maks. tid	1. Kvartil	3. Kvartil	Respondenter	Standard afvigelse
next-1	8.52	233.61	2.08	5058.36	4.84	16.76	28	943.77
køn	3.04	4.78	1.81	41.89	2.37	4.2	26	7.52
fødselsår	6.07	6.64	2.65	14.8	4.78	7.18	26	2.77
next-2	1.7	2.55	0.67	5.81	1.29	3.44	26	1.69
hvilken-region er du bosat i?	3.04	4.87	1.85	35.26	2.45	3.88	26	6.47
arbejdssituation	4.54	5.32	1.58	15.64	3.09	7.01	26	3.1
next-3	2.21	4.33	0.79	35.02	1.53	3.84	26	6.81
computeregenskaber	5.73	8.02	3.55	23.78	4.85	8.12	26	5.09
arbejdsudvikling	8.97	12.34	1.92	32.64	5.49	18.6	26	8.76
smartphones	6.59	7.82	2.36	26.12	5.55	8.55	26	5.2
next-4	1.79	2.51	0.83	9.96	1.43	2.75	26	2.03
forhold-til udvikling	7.33	9.49	1.61	39.58	5.76	10.36	26	7.11
smartphones-positiv-negativ	4.15	5.83	0.61	23.87	2.49	6.89	26	5.13
next-5	1.01	1.19	0.53	4.14	0.88	1.15	26	0.7
udviklingsforhold	8.03	24.55	1.9	376.01	6.61	14.48	26	70.59
teknologi-i fritiden	5.71	6.59	2.03	16.77	3.2	9	26	4.14
next-6	1.86	2.13	0.78	4.44	1.37	2.43	26	1.05
bliver-det offentlige mere effektiv	9.87	13.95	3.3	76.51	6.36	16.69	26	14.75
offentligt-bedre kvalitet	8.43	8.82	1.3	28.8	5.23	10.24	26	5.91
next-7	1.61	2.71	0.72	9.04	1.31	2.7	26	2.4
offentligt-vs privat	13.96	15.86	4.96	42.03	11.75	20.03	25	7.84
kontakt-til det offentlige	9.23	11.91	1.22	46.66	6.59	15.01	25	9.01
next-8	2.64	5.95	1.16	76.19	1.95	4.17	25	14.42
rigtige-beslutning at digitalisere	14.35	13.75	3.54	34.19	6.03	18.59	25	7.51
behov-for-undervisning	5.86	6.72	1.43	19.06	3.35	7.86	25	4.53
offentlig-undervisning	7.98	11.87	1.58	36.83	5.9	13.87	25	9.57
prev-2	3.41	3.41	3.41	3.41	3.41	3.41	1	0
prev-7	4.89	4.89	4.89	4.89	4.89	4.89	1	0
prev-0	1.95	1.95	0.66	3.23	1.3	2.59	2	1.28
next-missing_param	7.73	7.73	7.73	7.73	7.73	7.73	1	0

Tabel 15.3.3.1 – Tidsforbrug fordelt over spørgsmål i Soc-KMD

Af tabellen kan der aflæses mean- og median-tid, såvel som korteste, længste samt første og tredje kvartil af besvarelsestiden i de respektive spørgsmål. Dette giver således indsigt i

hvor længe respondenterne har været om at besvare spørgsmålene – såvel som et overordnet indblik i fordelingen af besvarelsener. Derudover er det ligeledes muligt at aflæse antallet af respondenter der har den givende event – hvilket betyder, at når respondentantallet falder, skyldes det at respondenterne har ophørt deres besvarelse af skemaet.

Det er værd at have i mente, at de tider som er knyttet til progressionsevents, såsom ”next”, vil være vanskelige at tolke. Dette skyldes at det i de data som indsamles via metoden, på nuværende tidspunkt, ikke er muligt at identificere hvad respondenterne foretager sig inden de skifter side, såfremt at de ikke foretager svarsift eller lignende. Denne problemstilling diskuteres nærmere i et senere afsnit.

Grundet samplens størrelse, vil enkelte outliers have en langt større effekt på resultaterne end hvis der eksempelvis var tale om en større sample, såfremt at antallet af outliers forblev det samme. Denne outlier-effekt kan aflæses af begge tabeller, især i mean-tiden såvel som standardafvigelsen – hvilket ligeledes er årsagen til at mediantiden anvendes, for på den måde at give et mere nøjagtigt billede på fordelingen af respondenter i datasættene (Moore & Notz 2006)(de Vaus 2002). Ved at anvende meanen, samt de tilstødende værdier som min, maks og de to kvartiler er det muligt at få et bedre billede af det egentlige tidsforbrug i de enkelte spørgsmål.

Tabellen åbner op for muligheden for at belyse de mest ekstreme tilfælde, her især med henblik på maksimum tiden. Med henblik på disse såkaldte outliers, kan vi se at det er en respondent som har opholdt sig på den første side i knapt 85 minutter. Det forventes at respondenterne har haft åbnet skemaet via linket, hvorefter at dette gemt i baggrunden, imens de har foretaget sig noget andet. Det er vigtigt at være opmærksom på at denne adfærd udelukkende registreres, når respondenterne bevæger sig videre til næste side. Der er derfor ikke tale om en respondent som har haft skemaet åbnet i længere tid, for derefter at lukke det ned helt uden at bevæge sig videre i forløbet – da vi, som tidligere nævnt, ikke vil opnå indsigt i et sådan tilfælde.

Outliers med usædvanligt lange pauser eller uventede stop i besvarelsesforløbet har en effekt, som ses både i form af hvordan median og mean udarter sig – her dog tydeligst indvirkning på meanen. Det er i forbindelse med analysen udarbejdet en algoritme, som

tager hånd om netop disse uventede pauser – hvilket har stor effekt på både mean såvel som standardafvigelse. Dog er forskellen på medianen ikke bemærkelsesværdig, da den kun varierer med få millisekunder fra det nuværende – disse tabeller er derfor ikke inkluderet i projektet.

Event	Median	Mean-tid	Min. tid	Maks. tid	1. Kvartil	3. Kvartil	Respondenter	Standard afvigelse
next-1	8.12	23.06	2.26	508.56	3.45	14.47	38	80.14
køn	3.32	5.94	1.56	92.98	2.26	4.15	35	15
fødselsår	5.44	33.38	3.43	952.43	4.35	7.96	35	157.64
next-2	2.02	3.24	0.42	23.92	1.09	2.88	35	4.28
hvilken-region er du bosat i?	3.33	3.74	1.72	17.84	2.59	3.8	35	2.67
jobsituation	3.83	4.35	1.64	15.82	2.74	5.33	35	2.49
next-3	1.87	2.66	0.79	11.81	1.18	2.65	35	2.31
bliver-det offentlige mere effektiv	13.38	16.39	5.16	56.41	9.47	17.79	30	10.82
offentligt-bedre kvalitet	9.28	13.48	1.34	62.73	5.82	14.78	30	13.1
next-4	2.09	3.62	0.74	20.44	1.25	3.36	30	4.41
offtenligt-vs privat	15.3	17.69	3.06	45.95	10.51	20.71	28	10.59
kontakt-til det offentlige	8.18	10.8	2.32	32.55	5.66	12.24	28	7.99
next-5	2.71	3.3	0.8	8.62	1.79	4.6	28	2.05
rigtige-beslutning at digitalisere	13.83	15.78	4.03	56.13	8.85	19.19	28	10.13
computeregenskaber	5.23	5.65	1.92	12.98	3.87	6.4	28	2.82
next-6	2.24	2.81	0.68	9.71	1.4	3.02	28	2.05
arbejdsudvikling	11.22	14.7	1.85	69.42	8.41	16.85	28	12.71
fritidsteknologi	10.55	14.75	0.88	57.06	6.22	18.12	28	12.56
next-7	2.32	2.86	0.65	13.75	1.45	3.17	28	2.53
hjaelp-til computer smartphone	8.26	10.58	1.28	46.14	5.48	10.49	28	10.1
prev-6	3.93	3.93	3.93	3.93	3.93	3.93	1	0
next-missing_param	5.24	9.84	4.06	20.2	4.65	12.72	3	7.35

Tabel 15.3.3.2 - Tidsforbrug fordelt over spørgsmål i KMD skema

På samme vis som det opleves i Soc-KMD skemaet, kan det ses at der i KMD skemaet ligeledes ses en respondent som *hviler* på forsiden, før de går videre i besvarelsen. Derudover er det af tabellen lettere at aflæse hvornår i forløbet at der opleves respondentfracfald. Af Tabel 15.3.3.2 kan det aflæses at der i KMD spørgeskemaet oplever størst respondentfracfald på første side, samt ved overgangen til de holdningsorienterede spørgsmål. Altså, at der er et større antal af respondenter der springer fra så snart at der

ikke længere blot at tale om spørgsmål vedrørende baggrundsdata, men i stedet undersøgelsens *egentlige* spørgsmål. Ligeledes opleves der i projektets eget spørgeskema også frafald på første side, hvorefter det sidste frafald opleves på skemaets næstsidste side.

Ud fra tabellerne kan vi aflæse at den tid der anvendes på at besvare spørgsmål vedrørende baggrundsdata, er væsentligt kortere end holdningsprægede spørgsmål. I begge skemaer, er tiden der er anvendt på forskellige baggrundsspørgsmål meget ensartet og befinder sig i de fleste tilfælde indenfor et sekund af hinanden. Der kan ligeledes i samme tematiske spørgsmålsbatterier ses at det at baggrundsspørgsmålene som respondenterne har anvendt mest tid på, er det vedrørende fødselsdato – hvilket kan forklares, ved at der i det tilfælde er tale om en liste af årstal, i en såkaldt *dropdown-menu*. Respondenten skal derfor søge igennem for at finde det korrekte årstal, hvorved at alle de givende svarmuligheder for respondenterne ikke er synlige på skærmen samtidigt – som er tilfældet i de resterende baggrundsspørgsmål. Denne tendens bliver yderligere tydeliggjort i KMD-skemaet, i det at spørgsmålet *"Hvordan vurderer du selv dine evner til at anvende en computer?"* forekommer i et spørgsmålsbatteri med holdningsspørgsmål. Spørgsmålet er i Soc-KMD blevet betragtet at være vedrørende baggrundsdata, hvorfor det er placeret i sammenhæng med de andre baggrundsspørgsmål. Det kan i tabellen over KMD-skemaet aflæses at respondenterne bruger væsentligt kortere tid på at besvare dette spørgsmål, sammenlignet med de omkringliggende holdningsspørgsmål. Det kunne i forlængelse heraf være interessant at undersøge, hvorvidt at der er tale om en generel tendens – hvor at erfaringsmæssige og baggrundsspørgsmål altid besvares på væsentligt kortere tid end holdningsspørgsmål.

Derudover kan det aflæses at der i begge datasæt er en median for progressionsskiftet *next* på lige omtrent to sekunder. Dog har der i KMD skemaet oplevedes væsentligt flere respondenter som har taget længere tid om dette skifte, end i Soc-KMD skemaet.

Af de to tabeller kan det aflæses, at der i nogle spørgsmål forekommer visse forskelle. KMD-skemaet har dobbelt så mange spørgsmål med en median på over 10 – hvilket alle er holdningsspørgsmål. Dog er forskellen i de to datasæt ikke større end at udfaldet kunne være gjort af et fåtal af respondenter. Spørgsmål såsom *"Offentlig-vs-privat"* såvel som *"Rigtig-beslutning-at-digitalisere"* har i begge skemaer taget omtrent lige lang tid at besvare. Det er her interessant at spørgsmålene i begge spørgeskemaer er ens – og at de

ikke er blevet omformuleret eller struktureret anderledes. Forskellen mellem de to forekommer udelukkende i tematisk rækkefølge.

15.3.4 Sammenligning af datasæt

Årsagen til at anvende to spørgeskemaer, var for på den måde at kunne belyse forskellige aspekter og facetter af processen. Dette åbner ligeledes op for mulighed for at kunne sammenligne de to datasæt, for på den måde at belyse hvorvidt at der opleves forskel i besvarelserne de to skemaer imellem.

Den primære forskel mellem de to spørgeskemaer, beror sig på den tematiske sammensætning såvel som antallet af spørgsmål. Indtil flere af spørgsmålene i KMD skemaet er i Soc-KMD blevet omformuleret eller delt i to, for på den måde at undgå *doublebarrel* spørgsmål og samtidigt sørge for at de svarmuligheder respondenterne blev præsenteret for, matchede det spørgsmål der blev stillet.

En af de oplagte ting at undersøge, er således hvorvidt at der er blevet anvendt væsentligt længere tid på at besvare det længere spørgeskema – eller om den tematiske struktur, såvel som opdelingen af spørgsmål har bevirket at respondenterne lettere forstår spørgsmålet og finder den svarmulighed der repræsenterer deres holdning.

	Median (Komplet)	1. Kvartil	3. Kvartil	Median (- baggrund)	1. Kvartil	3. Kvartil
KMD	157,5	115,25	241,25	116,5	84,75	161,5
Soc-KMD	194	141	325	140	101	225

Tabel 15.3.4.1 – Oversigt over tidsforbrug med og uden baggrundsdata.

I tabellen er to forskellige målinger, hvor den ene er den samlede besvarelsestid for respondenterne i spørgeskemaet, er der i det andet tilfældet blevet fjernet data fra baggrundsspørgsmål, i det at de alle fremstod som ensartede uden interessante forandringer. Det har gennemsnitligt taget respondenterne længere tid at besvare det længere spørgeskema, til trods for at det var formodet at mening og rækkefølge ville

forbedre dette. Selv når baggrundsspørgsmålene udelades, har besvarelsen af Soc-KMD skemaet taget længere tid.

I og med at det tidligere blev belyst, at spørgsmål med længere tekst tager længere tid at besvare – bliver det således et spørgsmål om hvorvidt at forskellen mellem de to besvarelser kan findes ene og alene på baggrund af de flere spørgsmål. Ser vi udelukkende på holdningsspørgsmålene, er der i KMD skemaet 9 spørgsmål, imens at der i Soc-KMD er 14 holdningsspørgsmål. Uden at det har nogen sikkerhed på sig, kan vi dog med det udgangspunkt se at respondenterne gennemsnitligt har anvendt kortere tid per spørgsmål – hvilket også kan aflæses af tabellerne i tidligere afsnit. Forskellen mellem de to skemaer er således at der i Soc-KMD skemaet er anvendt 10 sekunder i snit per spørgsmål – mens der i KMD skemaet er benyttet næsten 13 sekunder per spørgsmål. Hvorvidt at der er tale om en signifikant forskel er vanskeligt at belyse – dog er det interessant at der opleves en sådan forskel – i det at det giver anledning til at tro at ændringerne i spørgsmålene har haft en reel effekt.

Som tidligere nævnt er forskellen i frafald kun tæt på signifikans, til trods for at forskellen mellem de to er et frafald på 27 % i KMD-skemaet, mod små 11 % i Soc-KMD. Især i den udstrækning af majoriteten af frafaldet i KMD-skemaet forekommer på første side med holdningsorienterede spørgsmål.

15.4 Respondenttyper

Et andet spørgsmål i projektet vedrører hvorvidt at metoden evner at gøre det lettere at belyse og fange forskellige respondenttyper, her især med fokus på de som vil kunne have en mulig negativ indflydelse på det undersøgte. Afsnittet har til formål at analysere, hvorvidt at det i de indsamlede datasæt er muligt at fremhæve nogle af disse.

15.4.1 Speeders, cheaters og repeaters

For at kunne identificere de forskellige respondenttyper, vil jeg først kortlægge hvilke sammensætninger af tertiære data, der formodes at være i stand til at fremhæve de forskellige respondenttyper.

Speeders er betegnede som respondenter der søger at bevæge sig hurtigst muligt gennem undersøgelsen, hvilket bevirker at de under traditionelle webbaserede metoder vil identificeres på baggrund af deres samlede svartid. I forbindelse med denne undersøgelse, vil der dog i stedet søges at undersøge hele forløbet som helhed. Dette gøres ved først og fremmest at undersøge om der i besvarelsen forekommer svarskeft – eller om forløbet er fuldt lineært. I det at det ikke er forventeligt at en traditionel *speeder* vil tage sig tiden til at korrigere de svar der afgives.

Såfremt at respondentens besvarelse er lineær, er det muligt ved at kortlægge hvor længe man formoder at det som minimum vil tage at gennemføre besvarelse af de forskellige spørgsmål. Er det eksempelvis logisk at forvente at en respondent vil være i stand til at læse og forstå både spørgsmål og svarkategorier på to til tre sekunder?

Alternativt er det muligt programmatisk at sammenholde respondentens besvarelse med den resterende sample – eller en gruppe af respondenter – for på den måde at undersøge om respondenter blot besvarer spørgsmålene hurtigere eller om der er tale om en decideret *speeder*. Hvis man, i sammenligning med andre besvarelser, ser at den formodede *speeder* bruger procentvis længere tid på spørgsmål, tilsvarende til den stigning andre respondenter oplever.

Ligeledes vil brud på den lineære adfærd, såsom at respondenter bevæger sig frem og tilbage mellem siderne/slides i undersøgelsen også kunne fungere som tegn på, at der ikke nødvendigvis er tale om en *speeder*.

Skellet mellem *speeders* og *cheaters* synes ikke særlig stort – i det at den primære forskel der opleves mellem de to, er varigheden af besvarelsen. Der vil i tilfælde af *cheaters* oftest være tale om respondenter, der har et fuldstændigt lineært forløb – uden brud i forhold til svarskeft eller progression. Det vil antages at der i forløbet enten bruges relativt kort tid – eller at der vil opleves længere pauser ved hver besvarelse.

Repeaters kan på nuværende tidspunkt allerede fanges i mange tilfælde, i det at deres besvarelser typisk vil være plaget af ensartede svar og minimal spredning. Igennem metoden er det dog muligt at belyse hvorvidt at der er tale om en bevidst adfærd, i kraft af

at det bliver muligt at se om respondenter vælger svarmuligheder der befinder sig uden for den normale spredning, for derefter at skifte sit svar til et der befinder sig indenfor.

Igen er det et spørgsmål om, hvorvidt at undersøgelsen kan betragtes som værende lineær – eller om der i stedet vil opleves brud og forandringer undervejs.

Som tidligere omtalt, er det forventeligt at disse respondenttyper fremgår hyppigere i undersøgelser, hvori der udloves præmier eller hvor respondenter på anden vis belønnes for deres deltagelse. Da dette ikke tilfældet med de skemaer der er omdelt i forbindelse med projektet, vil det derfor forventes at der sandsynligheden for at disse respondenttyper findes være begrænset. Det er ligeledes værd at bemærke at den begrænsede samplestørrelse ligeledes bevirker at det formodes at disse respondenttyper vil være vanskeligere at finde.

15.4.1.1 Speeders

På baggrund af de karakteristika som er kortlagt i den teoretiske beskrivelse, ses det at en speeder er en respondent der søger at gennemgå besvarelsen af spørgeskemaet så hurtigt og lineært som muligt. Forstået på den måde, at speeders forløb forventes at være ensrettede, uden brud på kronologien eller svarskeft. Derudover forventes de at gennemføre deres besvarelse "*hurtigst muligt*".

Definitionen af "*hurtigst muligt*", synes dog vanskelig at præcisere, i det at der ikke er tale om et fast interval at det i stedet varierer fra respondent til respondent. Dog må det formodes at de respondenter som har besvaret væsentligt langsommere end resten af samplen, ikke har skyndt sig gennem forløbet i den forstand som definerer *speeders*.

For at fremhæve de mulige speeders i de to datasæt, er der således udviklet en fleksibel algoritme, som tillader forskellige justering af forskellige søgeparametre, hvilket illustreres og inddrages senere i dette afsnit. Algoritmen og dennes muligheder, samt de anvendte indstillinger kan summeres op (som følger på næste side).

	Svarskift	Progression	Pause-multiplikator	Tids-multiplikator
Grundindstilling	Nej	Ja	Nej	Nej
U. Progression	Nej	Nej	Nej	Nej
Tillad pauser	Nej	Ja	10x	Nej
Korrigeret svartid	Nej	Ja	Nej	1.5x
Korrigeret svartid samt pauser	Nej	Ja	10x	1.5x

Tabel 15.4.1.1.1 – Oversigt over algoritmeindstillinger

Årsagen til at svarskift udelades i samtlige indstillinger, er at svarskift indikerer at respondenterne ikke søger at gennemføre skemaet hurtigst muligt, hvorved at det vil formodes at de ikke vil have svarskift.

Den grundlæggende opsætning af algoritmen, er sammensat således at den vil fremhæve respondenter hvis besvarelsestid af samtlige spørgsmål befinder sig under eller er lig median-tiden og som samtidigt har fuldt ud lineære besvarelse, uden svar- eller progressionsskift. Denne indstilling er således den strengest mulige, i det at det kræver at samtlige af respondenterens events befinder sig under mediantiden.

Med denne opsætningen findes der ingen respondenter i de to datasæt, som matcher samtlige kriterier – hvilket giver anledning til i stedet at søge at anvende en af de korrigerede modeller. Der findes indenfor justeringen af algoritmen en række muligheder. I og med at mediantiden i progressionsevents, såsom *next*, er meget lav – synes det oplagt at undersøge, hvorvidt at der findes speeders, såfremt at den udelades. Det betyder dog samtidigt at der er mulighed for at respondenterne har haft et komplet lineært forløb, men har reflekteret over deres svar inden sideskift. Denne justering fremhævede heller ingen respondenter, hvilket betyder at der heller ikke findes respondenter der matcher disse kriterier.

At opsætte kriterier som påkræver at samtlige events skal befinde sig under mediantiden, kan synes urealistisk, i og med at det kun kræver at en enkelt af besvarelserne er udenfor rammerne førend at de frasorteres. Dette betyder således, at vi forbigår de respondenter som har haft længere pauser eller ophold på forskellige sider – hvilket har vist sig at forekomme hyppigere end tidligere forventet, som illustreret i et tidligere afsnit.

En justering af algoritmen betyder således, vil gøre det muligt at finde de respondenter som har haft de uventede og unaturlige pauser, imens at besvarelsestiden af de resterende spørgsmålene fortsat er under mediantiden. Det betyder at algoritmen i stedet korrigeres til at betragte spørgsmål eller skift til næste side, som har varet mere end ti gange så længe som mediantiden, vil blive betragtet som unaturlige pauser. Det vil sige, såfremt at et sideskift har taget mere end tyve sekunder, i gennemsnit – eller at besvarelsen af enkelte spørgsmål har varet mere end eksempelvis 80 sekunder, i spørgsmål der har varet otte. Med disse indstillinger, fremhæves en enkelt respondent fra KMD datasættet, som matcher vores kriterier.

ID	Tid	Median	Type	Event	Værdi
0	0	0	start	OS: Windows 8.1 (...)	
1	2.96	8.12	movement	next	1
2	1.76	3.32	choice	køn	2 - Kvinde
3	952.43	5.44	choice	fødselsår	1978
4	0.81	2.02	movement	next	2
5	2.48	3.33	choice	hvilken-region er du bosat i?	3 - Syddanmark
6	1.75	3.83	choice	jobsituation	3 - Selvstændig
7	1.03	1.87	movement	next	3
8	10.21	13.38	choice	bliver-det offentlige mere effektiv	4 - Nej, den offentlige sektor bliver noget mindre effektiv.
9	8.17	9.28	choice	offentligt-bedre kvalitet	4 - Nej, den offentlige sektor giver en dårligere service.
10	0.77	2.09	movement	next	4
11	3.06	15.3	choice	offentligt-vs privat	1 - Den offentlige sektor er langt bagefter den private sektor.
12	6.25	8.18	choice	kontakt-til det offentlige	1 - Digital Selvbetjening
13	0.97	2.71	movement	next	5
14	5.45	13.83	choice	rigtige-beslutning at digitalisere	4 - Forkert beslutning
15	2.44	5.23	choice	computeregenskaber	1 - Altid let at anvende
16	1.08	2.24	movement	next	6
17	3.9	11.22	choice	arbejdsudvikling	1 - Computere og smartphones vil få meget større betydning
18	2.22	10.55	choice	fritidsteknologi	1 - Udviklingen passer mig fint, og jeg anvender de tilbud, der passer til mig
19	1.18	2.32	movement	next	7
20	6.58	8.26	choice	hjælp-til computer smartphone	4 - Nej, det er der ikke behov for
21	1.15	3.93	end		

Tabel 15.4.1.1.2 – Respondent under median med ophold i KMD-skema.

På samme vis som tidligere, kan der i tabellen ses oversigt over et enkelt respondentforløb, samt mediantiden for samtlige besvarelser i skemaet. Dette tilfælde er det eneste hvor respondenter matcher samtlige kriterier i algoritmen, hvor unaturlige pauser inkluderes. I dette tilfælde er der tale om at respondenter har gjort ophold før besvarelse af fødselsår – som har været knapt 16 minutter.

Der er i de to første holdningsspørgsmål tale om at respondenter anvender tid der er omtrentligt tilsvarende til mediantiden, med få sekunders variation – hvorefter at det kan synes at respondenter hæver tempoet i sin besvarelse, for derefter at besvare de resterende spørgsmål væsentligt under mediantid. Undtagelserne er dog i spørgsmålene vedrørende kontakt til det offentlige, såvel som det afsluttende *double-barrel* spørgsmål – som begge har taget bemærkelsesværdigt længere tid. Det er tænkeligt at det er denne respondenter der kontaktede mig vedrørende de manglende svarmuligheder i det afsluttende spørgsmål i KMD skemaet – dog uden at det er muligt eller ønskeligt at belyse dette med nogen som helst nøjagtighed.

Det er værd at anmærke at der i denne besvarelse ikke er tale om at respondenter kun har besvaret med svar, der befinder sig i bestemte dele af spørgsmålsbatterierne – som eksempelvis ville kunne opleves i andre tilfælde. De afgivne svar, virker ligeledes sammenhængende og kongruente – i det at respondenter udviser at være gennemgående uenig med digitaliseringen af det offentlige. Dog er det interessant at de spørgsmål, hvor respondenter har besvaret med den første svarmulighed, ligeledes er de som har taget kortest tid at besvare.

Ved at ændre algoritmen ved at anvende en tidsmultiplikator, vil man i stedet være i stand til at fremhæve respondenter hvor samtlige af deres besvarelsestider befinder sig under 1.5 gange mediantiden. På den måde åbnes der op for at enkelte tider varierer, således at det ikke længere er krævet at alle svartider skal befinde sig indenfor median. Det ville ligeledes være muligt at begrænse algoritmen til kun at tillade et vist antal af tilfælde, som går op til 1.5 gange median – dette er dog ikke testet i dette henseende.

I denne model, findes der i begge skemaer en enkelt respondenter hvor samtlige besvarelsestider befinder sig under den fastsatte margin. Den samlede besvarelslængde er for respondenter i KMD skemaet 97,9 sekunder – imens det i Soc-KMD skemaet er blot

78,8. Som tidligere illustreret, var den generelle tendens at besvarelsen af Soc-KMD skemaet tog længere tid. Det er derfor interessant at illustrere forløbet fra dette skema, da den ikke blot matcher vores kriterier men ligeledes også tager væsentligt kortere tid, selvom den generelle tendens påpeger at den burde vare længere.

På samme vis som i forrige tabel, er det muligt at se den oprindelige median for de givende spørgsmål – for på den måde at belyse hvilke af svarene, der bevirker at respondentens besvarelse ikke fremhæves under de strengeste kriterier.

ID	Tid	Median	Type	Event	Værdi
0	0	0	start	OS: Mac OS X 10_10_0 (...)	
1	2.23	8.52	movement	next	1
2	1.98	3.04	choice	køn	1 - Mand
3	4.22	6.07	choice	fødselsår	1990
4	0.67	1.7	movement	next	2
5	2.21	3.04	choice	hvilken-region er du bosat i?	1 - Nordjylland
6	2.95	4.54	choice	arbejdssituation	6 - Under uddannelse
7	0.91	2.21	movement	next	3
8	4.66	5.73	choice	computeregenskaber	2 - Ofte let at anvende
9	3.11	8.97	choice	arbejdsudvikling	2 - Computere vil få større betydning
10	2.47	6.59	choice	smartphones	2 - Smartphones og tablets vil få større betydning
11	1.6	1.79	movement	next	4
12	4.13	7.33	choice	forhold-til udvikling	3 - Hverken positivt eller negativt
13	1.8	4.15	choice	smartphones-positiv-negativ	3 - Hverken positivt eller negativt
14	0.93	1.01	movement	next	5
15	9.28	8.03	choice	udviklingsforhold	1 - Udviklingen passer mig fint og jeg kan let følge med
16	2.74	5.71	choice	teknologi-i fritiden	1 - Flere gange dagligt
17	1.12	1.86	movement	next	6
18	6.59	9.87	choice	bliver-det offentlige mere effektiv	2 - Ja, den offentlige sektor bliver noget mere effektiv.
19	2.43	8.43	choice	offentligt-bedre kvalitet	2 - Ja, den offentlige sektor (...)
20	0.85	1.61	movement	next	7
21	7.01	13.96	choice	offentligt-vs privat	2 - Den offentlige sektor er noget bagefter den private sektor
22	3.15	9.23	choice	kontakt-til det offentlige	1 - Digital Selvbetjening
23	1.16	2.64	movement	next	8
24	5.44	14.35	choice	rigtige-beslutning at digitalisere	2 - Rigtig beslutning
25	2.99	5.86	choice	behov-for-undervisning	1 - Ja, det er der helt sikkert behov
26	2.18	7.98	choice	offentlig-undervisning	1 - Ja, det er det offentliges (...)
27	1.07	3.41	end		

Tabel 15.4.1.1.3 – Respondent med svartider under 1.5x median i Soc-KMD

Den omtalte forskel i besvarelsestid er fremhævet i tabellen²¹. Det kan på den baggrunds aflæses at der er tale om at respondenterne kun har haft handlinger der har varet længere end mediantiden i et enkelt tilfælde. De resterende svartider befinder sig således under medianen – og i de fleste tilfælde væsentligt under.

Fordelingen af svar i holdningsspørgsmålene, befinder sig for denne respondent i midt, samt den øverste halvdel af svarmuligheder. Det vil altså sige, at respondenterne i samtlige holdningsspørgsmål har besvaret med 1, 2 eller 3 – hvoraf det sidste er et nul-svar, som ikke præsenterer nogen holdning på hverken den ene eller anden side. Svarene kan ligeledes synes ret systemiske, i den forstand at nogle af sider med spørgsmål rummer svar med samme placering – hvilket kan ses i 12, 13 – 15, 16 samt 18 og 19. Dette kan selvsagt forklares ved at spørgsmålene i Soc-KMD skemaet i flere tilfælde er brudt op i enslydende tematiske spørgsmål, hvilket bevirker at respondenterne lettere danner forståelse for sammenhæng. Det er ligeledes forventeligt at holdningerne til de respektive spørgsmål, ligeledes vil være ensartede, i det at de beror sig på tema eller emne.

Samme tendens med at besvare spørgsmål under samme temaramme, med samme svar ses ligeledes gentaget i andre besvarelser – som fremhæves, ved at anvende den sidste indstilling af algoritmen. Algoritmen tillader således pauser som har varet minimum ti gange så længe som mediantiden, på samme måde som tidligere. Derudover er det ligeledes tilladt for respondenterne at have besvarelsestider op til 1.5x af medianen i de enkelte hændelser. De respondenter, som er blevet fremhævet i de to modeller som er slået sammen i denne algoritme, vises selvsagt atter engang. Derudover fremhæves der to respondenter yderligere – en i hvert datasæt – hvor respondenterne i Soc-KMD skemaet ligeledes illustrerer den førnævnte tendens tydeligt. Det kan på baggrund af tabellen ligeledes ses at der er tale om de samme hændelser som tidligere, samt tilføjelsen af ID 8, 9 og 10.

²¹ ID 15

ID	Tid	Median	Type	Event	Værdi
0	0	0	start	OS: Windows 7 (...)	
1	390.9	8.52	movement	next	1
2	41.89	3.04	choice	køn	1 - Mand
3	4.54	6.07	choice	fødselsår	1984
4	0.94	1.7	movement	next	2
5	1.87	3.04	choice	hvilken-region er du bosat i?	2 - Midtjylland
6	3.36	4.54	choice	arbejdssituation	4 - Arbejdssøgende
7	0.99	2.21	movement	next	3
8	3.55	5.73	choice	computeregenskaber	1 - Altid let at anvende
9	1.92	8.97	choice	arbejdsudvikling	1 - Computere vil få meget større betydning
10	2.36	6.59	choice	smartphones	1 - Smartphones og tablets vil få meget større betydning
11	0.83	1.79	movement	next	4
12	5.97	7.33	choice	forhold-til udvikling	3 - Hverken positivt eller negativt
13	0.65	4.15	choice	smartphones-positiv-negativ	3 - Hverken positivt eller negativt
14	1.02	1.01	movement	next	5
15	4.27	8.03	choice	udviklingsforhold	1 - Udviklingen passer mig fint og jeg kan let følge med
16	3.04	5.71	choice	teknologi-i fritiden	1 - Flere gange dagligt
17	0.94	1.86	movement	next	6
18	3.79	9.87	choice	bliver-det offentlige mere effektiv	1 - Ja, den offentlige sektor bliver meget mere effektiv.
19	1.3	8.43	choice	offentligt-bedre kvalitet	1 - Ja, den offentlige sektor giver en bedre service.
20	2.04	1.61	movement	next	7
21	6.4	13.96	choice	offentligt-vs privat	2 - Den offentlige sektor er noget bagefter den private sektor
22	3.76	9.23	choice	kontakt-til det offentlige	4 - Personligt Fremmøde
23	1.78	2.64	movement	next	8
24	4.77	14.35	choice	rigtige-beslutning at digitalisere	2 - Rigtig beslutning
25	2.35	5.86	choice	behov-for-undervisning	1 - Ja, det er der helt sikkert behov for
26	6.45	7.98	choice	offentlig-undervisning	1 - Ja, det er det offentliges ansvar at tilbyde borgerne den fornødne undervisning og information.
27	0.74	3.41	end		

Tabel 15.4.1.1.4 – Respondent med svartider under 1.5x median samt pauser i Soc-KMD

På samme vis som tidligere, er de tilfælde hvor respondentens besvarelse har været længere end medianen fremhævet. De to første tilfælde kategoriseres dog som unaturlige pauser, i

det at varigheden heraf er mere end 10 gange mediantiden. På den baggrund kan det dermed påpeges, at muligheden som tidligere nævnt med progression, medvirker til at respondenterne ikke tidligere er blevet fremhævet. Respondenterne ville derfor ligeledes blive fremhævet, såfremt at modellen justeres til at tillade pauser samt at udelade tid fra progressionshændelser, såsom "next".

Men er datasættets median overhovedet anvendelig i jagten på speeders? Det er vanskeligt at sige – men den er i afsnittet anvendt som et pejlemærke, for at illustrere nogle af de måder hvorpå man kan gribe til værks. Der er dog intet i litteraturen der dikterer noget omkring respondenterne forventede svartider – blot at de forsøger at arbejde sig så hurtigt muligt igennem skemaet. Ville det i stedet være mere fordelagtigt at søge at indsnævre rammerne – ved i stedet at anvende dele af de indsamlede data, til at fastlægge margins for de respektive besvarelser. Dette diskuteres yderligere i et senere afsnit.

15.4.1.2 Cheaters

Mulige cheaters og speeders formodes at have et lignende lineært forløb, på den måde at de heller ikke forventes at have brud på kronologien. Cheaters adskiller sig derfor ved at bruge længere tid på deres besvarelser end speederen.

Udgangspunktet for denne del af analysen er derfor at fremhæve respondenter som har en komplet lineær besvarelse – uden at tage hensyn til tiden respondenterne har brugt på at besvare skemaet. Med den algoritme ser fordelingen således ud;

KMD	Eget skema
11	6

Tabel 15.4.1.2.1 – Oversigt over potentielle cheaters

Respondenterne har haft et komplet lineært forløb, uden svar- eller progressionsskift. Det i sig selv er dog ingenlunde problematisk i og med at der ikke er noget der forhindrer respondenterne i at foretage alle overvejelser inden de markerer det ønskede svar. I kraft af at modellerne for speeders og cheaters er så ensartede betyder det ligeledes, at de formodede speeders fra forrige afsnit ligeledes indgår i denne sammenligning, da de lever op til samme kriterier.

Formuleringen af cheaters er desværre begrænset – og disse respondenter matcher de forudbestemte kriterier. Men er der nogle tiltag vi kunne gøre os, for bedre at kunne identificere hvorvidt at der er tale om cheaters? En mulighed kunne være at ændre algoritmen til også at tage hensyn til hvordan respondenterne besvarer sine spørgsmål. Nærmere betegnet, hvorvidt at der kunne være tale om en høj satisficing, ved at respondenterne eksempelvis har et fast mønster eller bestemte tendenser i deres besvarelse.

I dette tilfælde var der dog intet umiddelbart system at spore – i det at alle respondenterne har variationer i deres holdningsspørgsmål, dog ofte med tendens til at afgive ensartede svar i spørgsmålsbatterier. Det vil altså sige at en række af respondenterne udviser tendens til at spørge om et emne i et batteri med ensartet emne og skala, vil have identiske svar. Eksempelvis i batteriet, i projektets eget skema, der vedrører den teknologiske udvikling på arbejdspladsen, hvoraf det ene spørgsmål vedrører computere og det andet vedrører smartphones og tablets. Der kan dog være tale om at spørgsmålene er så ensartede i deres udformning og indhold, at respondenterne i den forstand ikke adskiller udviklingen mellem de forskellige former.

En smule anderledes er det dog, hvis vi iagttager disse respondenteres svar i det batteri som består af to spørgsmål, der udsprang fra et enkelt spørgsmål i KMD skemaet, vedrørende det offentlige effektivitet og kvalitet. Her ses der tendens til at respondenterne enten vælger at svare ensartet i begge spørgsmål, således at digitaliseringen formodes at effektivisere det offentlige, såvel som at gøre kvaliteten bedre. Imens at andre har svaret at det vil gøre det offentlige mere effektivt, men samtidigt vil kvaliteten blive forringet. Som en sidenote kan man påpege at dette svarmønster indikerer at opdelingen af det oprindelige spørgsmål var nødvendig, da den sidstnævnte holdning ikke er repræsenteret i KMD skemaet.

Såfremt at vi skulle udpege enkelte af disse respondenter som mulige cheaters, ville det være de som bevæger sig i disse systemiske mønstre og besvarer batterier ensartet. Spørgsmålet er, hvorvidt at der er tale om høj satisficing eller om respondenterne blot afgiver svar for at afgive svar. Det er vanskeligt at afgøre – men det må påpeges at muligheden eksisterer.

15.4.1.3 Repeaters

I forbindelse med at systemet til dataindsamling blev udviklet, inkorporeres der et system der gjorde det stort set umuligt for en respondent at gentage spørgeskemaet fra den samme computer. Hvor mange ofte vil anvende cookies til formålet, blev der i stedet anvendt hvad der er kaldet *server side sessions* – som tilskriver respondenten et permanent id baseret på en række karaktertræk. Denne metode blev valgt frem for cookies, i det at cookies kan slettes af brugeren og kræver en synlig tilladelse – i og med at loven kræver at modtager accepterer anvendelsen af cookies. I stedet lagres der data om respondentens computer, browser og forbindelse i serverens database midlertidigt – hvorved at respondenter der bevæger sig ind på siden igen vil blive tilknyttet det samme ID som tidligere.

Det er selvsagt muligt for respondenterne at kunne have snydt systemet ved at ændre nogle af de parametre som blev anvendt for at forhindre repeaters – da dette kun vil kræve at respondenten eksempelvis søger at besvare fra en anden enhed, såsom deres mobiltelefon eller tablet. Dette betyder ligeledes at respondentens eneste motivation for at deltage i undersøgelsen, skulle være undersøgelsen selv – i det at der i forbindelse med besvarelse *ikke* var udlovet præmier eller anden belønning.

Disse hindringer formodes at gøre det tilstrækkeligt vanskeligt for mulige repeaters at gentage besvarelsen af skemaet, såvel som den manglende belønning, til at antage at repeaters ikke forventes at findes i skemaet. Det betyder selvsagt at det bliver yderst vanskeligt at belyse eventuelle repeaters i det at det formodes at de ikke forekommer i disse datasæt.

16 Diskussion

Igennem analysen er der blevet bragt en længere række informationer til overfladen – og metoden er nu blevet testet i et praktisk scenarie for første gang. Men hvor står vi nu – og hvad evner metoden egentlig? Det er nogle af de emner der vil blive diskuteret i dette afsnit.

16.1 Andre respondenttyper

I projektet er der taget afsæt i teorien om speeders, cheaters og repeaters som værende nogle af de primære respondenttyper der opleves i webbaserede spørgeskemaer. Men rummer Thompson Metoden mulighed for at fremhæve andre respondenttyper end netop disse?

Som tidligere omtalt, har der i undersøgelsen af metoden været et stort fokus på anvendelsen af de tertiære variable, både som enkeltstående men ligeledes også kombinationen af de indsamlede data. Det er dog muligt at dykke endnu dybere ned i anvendelsen af de forskellige tertiære variable – for på den måde at etablere en dybere forståelse og indsigt i respondenternes besvarelsesforløb. Et af de emner i metoden som fortsat viderebygges, er idéen om svarmønstre og tendenser – hvor man på baggrund af respondenternes besvarelestendenser, svarskeft, spredning samt baggrundsdata vil kunne etablere deciderede respondenttyper. Tanken er at sammensætningen af forskellige variable, kan fremhæve forskellige tendenser blandt respondenter. Det vil på baggrund heraf være muligt at fremhæve forskellige respondenttyper, som vil formodes at kunne opstå hyppigere end andre. Der kunne eksempelvis være tale om, at mandlige respondenter, som er under uddannelse, der befinder sig i alderen 18 til 24 – i langt højere grad vil have tendens til komplet lineære forløb på relativt kort tid, sammenlignet med de resterende respondenter. Det vil på den måde højne sandsynligheden for at kunne udelukke formodede respondenter som eventuelle speeders, i det at det i stedet kan ses som en generel tendens for denne respondenttype. Derved ikke at forstå at en respondentgruppe udelukkende vil være bestående af én bestemt respondenttype, men at de forskellige typer opstår med forskellig sandsynlighed indenfor bestemte kriterier.

Sådanne respondenttyper forventes dog ikke at kunne indsamles med nogen pålidelighedsgrad i en undersøgelse med en så lille stikprøve, som har været tilfældet i projektet. Det vil i stedet formodes at det er nødvendigt at have en væsentligt større og mere differentieret sample – eventuelt som kombination af flere undersøgelser.

Alt om alt leder dette op til, hvorvidt at den algoritme der har været anvendt til at fremhæve speeders, vil være den bedste. Den fremhæver ganske vist de respondenter som anvender kortest tid, til trods for pauser og uventede stop – men samtidigt skabes der en

del usikkerhed om hvorvidt at den måde at analysere dataene på korrekt fremhæver speeders. Dette især, er en af årsagerne til at fremhæve de mere grupperede respondenttyper, i det at der i stedet vil kunne dannes et mere detaljeret overblik end hvis der blot analyseres på tværs af samplen.

Dette vil ligeledes åbne op for muligheden for at fremhæve kontekstuelle problemstillinger i eksempelvis formulering eller indhold af spørgsmålet. Forstået på den måde, at det er tænkeligt at der i et spørgeskema kan være spørgsmål som er lettere forståeligt for respondenter i en bestemt aldersgruppe – eller med et bestemt uddannelsesniveau. En problemstilling som ellers vil blive overset, såfremt at der kun analyseres på tværs af den samlede sample.

Der er som led i denne diskussion foretaget en række simple lineære regressioner, med henblik på at belyse eventuelle muligheder i de to datasæt for konstruktion af et sæt præliminære respondenttyper. Der blev forsøgt undersøgt hvorvidt at respondenternes baggrundsdata ville have nogen indflydelse på deres besvarelse, her betragtet i samlet tid for besvarelse såvel som antal af svarskeft. Progressionsskift blev udeladt, i det at der var tale om så lille et antal. Der er dog ingen signifikante sammenhæng at spore i de forskellige regressioner, hvorfor at de er udeladt i analysen.

Udviklingen af respondenttyper vil dog formodes at kræve mere af forskerens tid – og må derfor antages primært at være anvendeligt på institutter, hvor der foretages en længere række undersøgelser. Således at der findes tid og mulighed for at skabe og belyse de forskellige respondenttyper. Det vil selvsagt kræve en megen finjustering og analyse, førend man har fuldt pålidelige respondenttyper – dog vil præliminære typer kunne beregnes på relativt enkelt, hvilket vil formodes at være bedre end ingen typer.

16.2 Metodens anvendelighed

Det store spørgsmål efter analysen er således. Er metoden anvendelig i et *virkeligt* forskningsscenario og hvilke komplikationer måtte der eventuelt være i forbindelse med anvendelsen?

Som med så mange andre metoder, er der endnu intet som er hugget i sten. Der er ingen øvre eller nedre værdier til at definere præcist hvornår emner bør fremhæves, enten i forhold til respondentens adfærd og besvarelse – eller designmæssige problemer i spørgeskemaet. Det er derfor ikke muligt, på nuværende tidspunkt, at automatisere processen. Det er fortsat nødvendigt at forskeren læser og tolker på de resultater som metoden tilvejebringer. Når det er sagt, har målet med metoden aldrig været at fratage forskeren sin beslutningsevne og overlade alt til computeren. Målet har i stedet været at skabe en metode som evner at fremhæve mulige problemområder, til nærmere undersøgelse og refleksion. I det henseende synes metoden så bestemt at have sin anvendelse – i det at den har formået at belyse en længere række, både komplikationer både i forhold til spørgeskemaerne der har været anvendt i projektet, såvel som respondenternes besvarelser.

Der har i begge af projektets spørgeskemaer, været en yderst lille sample – hvilket har afgrænset muligheden for at undersøge metodens skalerbarhed. Dette har ligeledes været hensigtsmæssigt, i den forstand at det har været enklere at gennemskue hvorvidt de modeller og beregninger der er foretaget, har været korrekte og fungeret efter hensigten.

Hensigten med modellen har hele tiden været, at den ikke skulle kræve for meget tid af forskeren, men i stedet skulle fungere så enkelt som muligt. Det skulle igennem resultaterne blive lettere for forskeren at lokalisere hvor det vil være fordelagtigt at dykke dybere ned i de indsamlede data. Det er således essentielt at der er forskel på den studerende, som foretager en spørgeskemaundersøgelse for første gang – og forskeren som ønsker at undersøge hvordan respondenter agerer i spørgeskemabesvarelser. De har hver deres ønskede *dybde* og forskellige interesserer i metodens udbytte – og netop i det henseende kan det på nuværende tidspunkt være vanskeligt at forenkles metodens resultater så meget, at de vil kunne anvendes af dem der kun ønsker den mest overfladiske indsigt.

Blandt overvejelserne for at simplificere metoden, har været hvorvidt at det ville være fordelagtigt med score-beregninger, således at respondenternes besvarelser og designproblemer i spørgeskemaet fremhæves på baggrund af en simpel og let-fortolkelig score. Problemet i udviklingen heraf, er at det kræver indsigt i en længere række

undersøgelse, for derved at kunne udvikle en scoreberegning der fungerer pålideligt. Dog forventer jeg ikke at det på noget tidspunkt, vil være muligt at automatisere processer fuldt ud – i det at forskerens indsigt og analyseevne kan tilvejebringer resultater og overvejelser som er stort set umulige at programmere. Iagttager vi eksempelvis den problemstilling, som tidligere blev omtalt, hvor et spørgsmål blev præsenteret i den forkerte skala i forhold til hvad respondenterne egentlig forventer, ville denne konklusion ikke kunne drages automatisk – i det at der lige såvel kunne være tale om et spørgsmål, hvor svarmulighederne overlappede eller skabte problemer.

Alt om alt, synes metoden bestemt at have sin anvendelse. Med tid og flere resultater, forventes det ligeledes at det vil blive væsentligt lettere at korte selve bearbejdsprocessen ned - og fremme de resultater som er mest relevante for forskeren.

17 Videreudvikling af metoden

I designet og udviklingen af den software der har været anvendt til dataindsamling og behandling, har der været nogle antagelser og idéer om hvordan og hvorledes der kunne genereres det bedst mulige output. Dette afsnit har til formål at diskutere mulige ændringer i systemet som det er nu – for på den måde at kunne forbedre det analytiske potentiale eller ændre nogle af de problemstillinger der har opstået i forbindelse med bearbejdelse af den indsamlede data.

17.1 Eventsampling

Eventsampling har overordnet set vist at frembringe en stor del af de formodede resultater. Der opstår dog en problemstilling i forbindelse med respondentfracfald, i det at vi efter sideskift er blinde indtil der forekommer et andet progressionsskift, i og med at information kun videregives fra respondenteren når de bevæger sig frem eller tilbage i spørgeskemaet. Det betyder at forskeren kun er i stand til at vide hvad der sker op til det punkt, hvor respondenteren går videre i skemaet – og ikke såfremt at de har foretaget sig noget på den side med spørgsmål de er skiftet til.

Dataregistreringsmetoden blev oprindeligt valgt for ikke at overbelaste serveren med for mange forespørgsler på for kort tid, hvorved at data kunne risikere at gå tabt. Efter at have gennemgået log-filer over ressourceanvendelse kan det aflæses at serveren på intet

tidspunkt har ladet sig påvirke af respondenternes besvarelser – og at *load* sagtens tillader en langt større respondentbelastning og dataindsprøjtning. Fremadrettet synes det derfor fordelagtigt at flytte registreringen til at dække over samtlige events, såsom når respondenter foretager svarskift. Det vil derfor give forskeren mulighed for at få indsigt i respondentens handlinger op til deres afslutning af skemaet. Det vil derfor være muligt at kunne identificere hvorvidt at respondenterne rent faktisk har afgivet svar – eller blot stoppet uden at have besvaret et eneste af spørgsmålene. Derudover bør man ligeledes søge at registrere når respondenteren påbegynder undersøgelsen²², i det at det ligeledes vil give anledning til at undersøge hvorvidt at der er tale om at respondenteren er stoppet allerede ved den indledende formulering af undersøgelsen. Dette kan således give anledning til revidering og nærmere undersøgelse af tekst, vilkår og andet der måtte fremgå på den første side.

Denne ændring vil ligeledes forbedre dataindsamlingsmulighederne i scenarierne hvor spørgeskemaerne rummer flere spørgsmål per side end der er tilfældet med skemaerne der er anvendt i dette projekt.

17.2 Scrolltracking

Efter at have gennemgået de indsamlede data, synes det oplagt at der gennem metoden skabes indsigt i en lang række ting. Spørgeskemaerne der har været anvendt i projektet, har været struktureret således at hver side med spørgsmål sjældent ville fylde mere end en enkelt sidelængde. Det betyder at respondenterne altid har været i stand til at se samtlige spørgsmål på siden samtidigt – med nogle enkelte undtagelser af respondenter med en tilstrækkelig lav opløsning på mobil. Netop disse respondenter ligger til grund for overvejelsen af denne videreudvikling af metoden, i det at der i andre undersøgelser ofte vil være langt flere spørgsmål per side, hvorved at hver enkelt side er længere end en enkelt skærmlængde. Det formodes at tilføjelsen af scrolltracking vil give forskeren yderligere mulighed for at opnå indsigt i respondenternes forløb.

Scrolltracking betyder kort fortalt, at man registrerer hvorvidt at respondenteren bevæger skærmbilledet længere ned på siden, for eksempelvis at læse hele siden samtidigt – eller om de eventuelt har scrollet ned til de næste par spørgsmål, før de påbegynder besvarelsen.

²² Ved *start*-eventet, som ikke tidligere har gemt besvarelser.

Det vil ligeledes give anledning til at undersøge hvorvidt at respondenterne scroller op på siden igen, for at læse eller genlæse spørgsmål, korrigerer svar og lignende. På samme måde som det opleves med progression, hvor respondenterne bevæger sig tilbage i skemaet – dog med forventning om at denne adfærd vil forekomme hyppigere end bagudgående progression, sammenlignet med de resultater der har været i forbindelse med projektet.

Selve handlingen af scrolle, kan synes som meget data, men det er muligt at bearbejde det således at den kun opdaterer når nye elementer kommer ind på skærmen. Det vil altså sige, at man undgår at skulle forholde sig til koordinater i forhold til respondentens placering, men i stedet vil være i stand til at se hvilke elementer respondenterne har scrollet for at kunne se.

17.3 Out of focus-tracking

Som først illustreret i min bacheloropgave fra 2012, var det en oplagt mulighed at inkludere viden om respondentens vinduesfokus, når skemaet var åbent. Det er ganske vist ikke muligt at registrere hvad respondenterne foretager sig uden om spørgeskemavinduet, men det er muligt at registrere at vinduet ikke længere er i fokus. Det betyder altså, at når respondenter har større pauser – som ses i tabellerne 18.3.3.1 og 18.3.3.2 – vil det være muligt at registrere hvorvidt at der er tale om en pause, hvor respondenterne har foretaget sig noget andet på computeren eller om vinduet fortsat har været i fokus på deres skærm.

17.4 Tid, sted og genoptagelse

Det har i forbindelse med besvarelse af spørgeskemaerne i projektet ikke været muligt at foretage genoptagede besvarelser, hvor en respondent eksempelvis bliver sendt et unikt link – der gør at de vil kunne lukke og åbne besvarelsen efter ønske. I det at undersøgelsens målsætning var at undersøge de grundlæggende elementer i Thompson Metoden, blev denne funktion undladt – en målsætning der også afspejles i længden af spørgeskemaerne. Begge de anvendte spørgeskemaer er begge så korte at det ikke formodes at der for en respondent vil være behov for at skulle genoptage besvarelsen. Dette vil dog formodes at være en nødvendighed når der arbejdes med væsentligt længere spørgeskemaer.

Det vil i forbindelse med besvarelse og især genoptagelse være fordelagtigt at registrere og arbejde med klokkeslæt. Hvornår på dagen er skemaet besvaret og hvornår er besvarelsen eventuelt genoptaget, såfremt at dette skulle være tilfældet. Det vil i forbindelse med genoptagelse ligeledes være fordelagtigt at anvende *delvis geolocation*, som gør det muligt for forskeren at se hvorvidt at respondenterne har genoptaget undersøgelsen samme sted som tidligere – eller om de eventuelt har flyttet sig, geografisk set. Det vil altså sige at det vil være muligt at identificere hvorvidt at respondenterne har foretaget besvarelsen et sted, for senere at besvare den et andet.

Har respondenterne stoppet og genoptaget besvarelsen, er det med ovenstående data muligt at identificere forskelle, såfremt at der skulle være ændringer i måden hvorpå respondenterne besvarer skemaet eller deres besvarelser.

18 Konklusion

Målet med projektet var at undersøge hvilken effekt det vil have at anvende Thompson Metoden, i forbindelse med onlinebaserede spørgeskemaer. Sideløbende med problemformuleringen var der således en række underspørgsmål, som dækker over nogle af de formodninger der har været for metodens kunnen forud for undersøgelsen.

18.1 Spørgsmål 1

Første underspørgsmål er *”Metoden evner at fremhæve mulige fejl og mangler i spørgsmålene i spørgeskemaet”*. Med et tilbageblik på analysen, kan det ses at der er en række forskellige indikatorer som kan henlede forskerens opmærksomhed på problemstillinger der bør adresseres. Eksempelvis opleves der i KMD spørgeskemaet et højt frafald så snart at respondenterne når den første side med holdningsspørgsmål. Der er på siden et såkaldt dobbelt-barrel spørgsmål, hvor nogle af respondenternes holdninger ikke ville kunne tilkendegives. Der kan dog ligeledes være tale om at respondenterne ikke længere ønsker at deltage i undersøgelsen, så snart at de bliver nødt til at anvende yderligere kognitiv egenskab på at besvare de forskellige spørgsmål. Det synes dog usandsynligt at der i de to skemaer vil være så stor forskel i respondenternes ønske til at deltage – hvorved at man må formode at der må være tale om en mulig fejlformulering eller

uoverensstemmelse mellem spørgsmål og svarmuligheder. På den baggrund vil det argumenteres at metoden evner at fremhæve mulige fejl eller mangler i spørgeskemaet.

18.2 Spørgsmål 2

Andet spørgsmål, ”*Metoden evner at fremhæve designfejl i spørgeskemaet.*”, er beslægtet med første – dog med den forandring at der er tale om en forskel som beror sig på udelukkende designmæssige komplikationer i henhold til spørgsmålstypevalg såvel som rækkefølge og opsætning af spørgsmålsbatterier. Med udgangspunkt i analysen vedrørende tidsforbrug i de forskellige spørgsmål, ses det især tydeligt at der er forskel på tidsforbruget i spørgsmål vedrørende holdning og spørgsmål vedrørende erfaring eller baggrund. Her skiller det eksempelvis sig ud ved, at der midt i undersøgelsen er et erfaringsbaseret spørgsmål som respondenterne bruger væsentligt kortere tid på end de omkringliggende spørgsmål. Dette giver således anledning til at forskeren nærmere kan evaluere rækkefølgen og opsætningen af spørgsmålene – om den placering af spørgsmål er den absolut mest hensigtsmæssige, eller om det eventuelt ville kunne flyttes til en anden placering i skemaet.

På samme måde, oplevedes det i begge datasæt at der var en stor tendens til at respondenterne havde svarskeft i et spørgsmål vedrørende kontaktformer til det offentlige. Den umiddelbare konklusion herpå, er at der var uoverensstemmelse mellem den skala og de muligheder der blev præsenteret og respondenternes forventning til skalaen. Dog med den ene forandring at der i det ene skema var udspecificeret at der var tale om en *primær* kontaktform – hvilket formodes at have haft indflydelse på at færre respondenter havde svarskeft heri. Dog er resultaterne på dette spørgsmål herefter så tvetydige at det bør være en overvejelse eventuelt at udelukke eller modificere anvendelsen af den indsamlede data. Men herigennem bliver det således belyst at der er en række muligheder for at fremhæve mulige designfejl i spørgeskemaet.

18.3 Spørgsmål 3

I tredje spørgsmål, forsøges der at kastes lys over metodens dybde med formuleringen ”*Metoden skaber en dybere indsigt i respondenternes besvarelsesforløb.*”. Sammenligner vi med en ordinær dataindsamling, er indsigten i selve forløbet mangedoblet. Det er på

baggrund af de indsamlede data muligt at skabe indblik i selve forløbet til sådan en grad at det kan syntetiseres både som graf – eller som animation af forløbet, såfremt det ønskes. Det er dog ikke uden sine undtagelser – i det at der, som nævnt, fortsat opleves blinde pletter hvor forskeren i det nuværende datasæt ikke har adgang – såvel som viden forskeren aldrig vil kunne være i stand til at indsamle.

Denne dybere indsigt åbner op for muligheden for at studere selve besvarelsen nærmere, således at respondentadfærd bliver et studium i sig selv, hvor man tilnærmer sig et observationsstudie i respondentens vante rammer. Ligeledes med den tilføjede bonus, at respondenterne ikke er bevidst om observationen – hvorved at anonymitet opretholdes og reaktivitet mindskes. Vi er derfor i stand til at konkludere at metoden evner at skabe dybere indsigt i respondenternes besvarelsesforløb – til trods for at der fortsat er rum og mulighed for udvidelser og forbedringer heraf.

18.4 Spørgsmål 4

I det sidste spørgsmål, ”*Metoden evner at fremhæve mulige speeders, cheaters og repeaters.*”, undersøges det hvorvidt at metoden evner at fremhæve de respondenttyper som der forud for analysen var formodet at være blandt de mest prominente og problematiske.

Som også omtalt i analysen, er et af de hovedsagelige problemer i dette henseende at der ikke er blevet anvendt nogen form for belønning for besvarelse af spørgeskemaerne – hvilket formodes at fremme frekvensen af speeders og repeaters. Ligeledes bevirker de små samplestørrelser at sandsynligheden for at have respondenter der matcher disse respondenttyper. Dog evnede vi gennem metoden at belyse de respondenter hvis kriterier matchede dem der var definerede for speeders og cheaters – dog uden at nogle af disse blev frasorterede i forhold til en eventuel indholdsanalyse. Antagelsen vil derfor siges at være delvis bekræftet, i det at den evnede at belyse respondenter der matchede de forudbestemte kriterier – dog uden at der i analysen blev testet for repeaters, i og med at muligheden for deres eksistens i datasættet var absolut minimal.

Dette spørgsmål bør dog videreføres, med henblik på diskussionen andre respondenttyper – i det at disse vil medvirke til at kunne definere intervaller og grænseværdier, for respondenter med lignende træk. Hvorved at vi på baggrund af det der er etableret i

analysen, kan se at vi vil være i stand til at kunne optegne disse respondenttyper til videre brug i en analyse.

18.5 Opsummering

Hvilken effekt, hvis nogen, vil det have at anvende Thompson Metoden?

Den overordnede problemformulering for projektet beror sig på at undersøge hvilken effekt anvendelsen af Thompson Metoden vil have. Denne kan dog opdeles og gradbøjes i flere henseende, i det at det at anvendelsen vil have forskellige effekt ved forskellige applikationer.

På baggrund af analysen kan de forskellige scenarier opdeles i to overordnede forgreninger, i form af de som søger at anvende metoden som en sikkerhed eller validering af deres indsamlede data og de som søger at undersøge respondentadfærd i spørgeskemaer.

Tager vi hånd om første scenarie, vil vi på baggrund af det analyserede konkludere at effekten vil være at forskerne har bedre mulighed for at vurdere deres spørgeskemadesign og formuleringer, såvel som arbejde sig hen imod respondenter som muligvis deltager med anden interesse end det undersøgte – hvilket især formodes at være tilfældet såfremt at der udloves præmier og andet i forbindelse med besvarelse. Derved ikke sagt at deres besvarelser er mindre værd i et sådan tilfælde, men det vil åbne op for at kunne opnå bedre indsigt i hvorvidt at der er tale om satisficing eller muligvis social desirability. Det bliver på den måde lettere at sikre sig, at det der søges undersøgt rent faktisk er det der bliver undersøgt.

Det forventes ikke at forskere i første scenarie vil søge at dykke ned i de enkelte besvarelser i andre tilfælde end pilotundersøgelser – men ville med fordel kunne konstruere respondenttyper, med det formål at belyse de gennemgående tendenser i besvarelserne.

Anderledes formodes det at de der søger at undersøge selve respondentadfærden i et rent spørgeskemateknisk regi, får åbnet op for en ny indsigt og nye muligheder – i det at det bliver lettere at måle og aflæse de konsekvenser forskellige designbeslutninger og formuleringer har. Det vil altså sige at man i undersøgelsen af adfærd, ville kunne stille spørgsmål som ikke tidligere har været betragtet som god praksis – ved eksempelvis at stille spørgsmål som er enten tabuiserede eller stigmatiserende²³ emner. Det åbner ligeledes op for at undersøge de tekniske indvirkninger, eksempelvis hvilken indvirkning en omvendt skala vil have på respondenternes tendenser til svarskeft. Det vil på den vis også være muligt at iagttage, om kontrolspørgsmål prompter en usædvanlig lang pause eller om respondenterne føler at der er tale om noget gentagende, for derved at gå tilbage i skemaet og undersøge det nærmere – såfremt at man i undersøgelsesdesignet har tilladt fri mobilitet.

Det formodes at for dem der undersøger respondentadfærd at det vil være interessant og relevant at dykke ned i de enkelte besvarelser, for på den måde at opnå indsigt i de forskellige facetter af respondenternes besvarelser. Det forventes ligeledes at disse vil anvende udvidede respondenttyper til at undersøge og etablere indsigt i de forskellige tendenser blandt respondentgrupperne.

Kort fortalt, så er effekten af at anvende Thompson Metoden en dybere indsigt i respondenternes adfærd, spørgeskemaets design og bedre mulighed for at validere at det undersøgte er forstået og modtaget af respondenterne.

²³ (Goffmann 2001)

19 Litteraturliste

- Adams, J. Stacey** (1965), "Inequity in social exchange". New York: Academic Press.
- Bethlehem, Jelke** (2010): "International Statistical Review". Oxford: Blackwell Publishing Ltd.
- Blasius, Jörg & Victor Thiessen** (2012): "Assessing the Quality of Survey Data". Bonn: SAGE Publications Ltd.
- Bryman, Alan** (2008): "Social Research Methods", New York: Oxford University Press.
- Cisco** (2014): "... Era of Mobile Data Consumption" - <https://gigaom.com/2014/02/05/cisco-the-u-s-is-officially-in-the-gigabyte-era-of-mobile-data-consumption>
- Couper, Mick P.** (2000): "Web Surveys". Oxford: Oxford University Press.
- Couper, Mick P.** (2011): "The future of modes of data collection" - Public Opinion Quarterly, Vol. 75, No. 5, 2011, pp. 889–908. Oxford: Oxford University Press.
- Couper, Mick & Peter Miller** (2008): "Web Survey Methods" - Public Opinion Quarterly, Vol. 72, No. 5 2008, pp. 831–835. Oxford: Oxford University Press.
- de Vaus, David A.** 2001: "Research Design in Social Research". London: SAGE Publications Ltd.
- de Vaus, David A.** 2002: "Surveys in Social Research". Singapore: Phoenix Print Media Pte Ltd. 5. udgave.
- Eckersley, Peter** (2014): "How Unique Is Your Web Browser". Electronic Frontier Foundation.
- Facebook** (2014): "Mobile usage statistics". <http://newsroom.fb.com> - <http://www.statista.com/statistics/277958/number-of-mobile-active-facebook-users-worldwide/>
- Freedman, David et al** (2007): "Statistics". New York: W.W. Norton & Company inc.
- Fricker, Scott et al** (2005): "An Experimental Comparison of Web and Telephone Surveys" - Public Opinion Quarterly, Vol. 69, No. 3, Fall 2005, pp. 370–392. Oxford: Oxford University Press.
- GDS UK GOV** (2013): "How many people are missing out on Javascript enhancement" - <https://gds.blog.gov.uk/2013/10/21/how-many-people-are-missing-out-on-javascript-enhancement/>
- Giddens, Anthony** 1997: "Social Theory and Modern Sociology". Cambridge: Polity Press.

Goffman, Erving 2009: "Stigma", Frederiksberg: Forlaget Samfundslitteratur. Grimen,

Groves, Robert M. (2011): "Three Eras of Survey Research" - *Public Opinion Quarterly*, Vol. 75, No. 5, 2011, pp. 861–871.

Groves, Robert M. et al. (2009): "Survey Methodology", 2nd Edition. New Jersey: John Wiley & Sons, Inc.

Heerwegh, D. & Looseveldt, G (2008): "Face-to-face versus web surveying in a high Internet coverage population: differences in response quality". Oxford: Oxford University Press.

Hox, Joop et al (2008): "International Handbook of Survey Methodology". Jena: European Association of Methodology.

Hviid Jacobsen, Michael & Kristiansen, Søren (2001): "Farligt feltarbejde". Aalborg: Aalborg Universitetsforlag.

Kenett, Ron S. & Salini, Silvia (2012): "Modern analysis of customer surveys". West Sussex, UK: John Wiley & Sons, Ltd.

Kreuter, Frauke et al (2008): "Social Desirability in cati, ivr and Web Surveys". Oxford: Oxford University Press.

Krosnick, J. (1991): "Response Strategies for coping with the cognitive demands of attitude measures in surveys". *Applied Cognitive Psychology* 5 (3): 213–236.

Krosnick, J. A. et al (1996): "Satisficing in surveys". Initial Evidence. San Francisco: Jossey-Bass.

Kvale, Steinar og Svend Brinkmann 2009: "Interview – introduktion til et håndværk". København: Hans Reitzels Forlag. 2. udgave, 3. oplag.

Moore, David S. og William I. Notz 2006: "Statistics – Concepts and Controversies". New York: W. H. Freeman and Company.

NetmarketShare (2014): "Mobile/Tablet Browser Trends" - <https://www.netmarketshare.com/browser-market-share.aspx?qprid=1&qpcustomb=1>

OECD (2013): "OECD Skills Outlook 2013: First Results from the Survey of Adult Skills". OECD Publishing. <http://dx.doi.org/10.1787/9789264204256-en>

Politiken (2014): Teleselskaber stopper deling af hemmelige mobilnumre - <http://politiken.dk/forbrugogliv/digitalt/internet/ECE2444978/teleselskaber-stopper-deling-af-hemmelige-mobilnumre/>

- Reis, H. T., & Gable, S. L.** (2000). "Event sampling and other methods for studying daily experience". New York: Cambridge University Press.
- Retsinformation** (2014): "Bekendtgørelse om krav til information og samtykke ved lagring af eller adgang til oplysninger i slutbrugeres terminaludstyr" - <https://www.retsinformation.dk/forms/R0710.aspx?id=139279>.
- Roster, Catherine et al** (2004): "International Journal of Market Research" Vol. 46 Quarter 3. New Mexico: University of New Mexico.
- Wiljas-Jensen, Justyna (2013): "It-anvendelse i befolkningen" - 2013. København: Danmarks Statistik.
- Treiman, Donald J.** 2009: "Quantitative Data Analysis". San Francisco: Jossey-Bass
- Ældresagen** (2014): Ældre i Tal. Danmarks Statistik.