

DANSK DATA ARKIV / DANISH DATA ARCHIVES

Adresse/Address	Islandsgade 10, DK-5000 Odense
	Tlf: +45 6611 3010 Telefax +45 6611 3060
e-mail	mailbox@dda.sa.dk
Internet	http://www.sa.dk/dda/

Udgiver **Dansk Data Arkiv** Redaktør **Bernhard Hansen** Produktion **Erik Steenstrup Dyhr**
Tryk **Riber Reklame AS Fåborg** Oplag **1200** ISSN **0105-3272**

Indhold

Redaktionelt	3
Med sproget som måleredskab – Om udforskning, testning og forbedring af vilkår for forståelse i spørgeskemabaseret forskning	4
Bevaring af viden – Om oparbejdning af data i Dansk Data Arkiv	24
Folketællingen 1801 en guldgrube	36
Enkestanden i Danmark 1801 – En analyse af en specifik social gruppes leveforhold, baseret på den komplette 1801 folketælling	38
Færdiganmeldelser	
Fritidsundersøgelsen 1993	50
Odense Kommune og borgerne 1990	51
Fysisk handicappedes vilkår 1994-1995	52
Odense Kommune og borgerne 1995	52
Unge andengenerations indvandrere 1999	53
Borgerne og lovene 1997	54
Trafik og levevilkår i Kgs. Enghave 2001	54
Valgdeltagelse ved kommunalvalget i 1997 i København og Århus	55
Oplevet diskrimination 1998	56
Trivsel, livsvilkår og livskvalitet under krisen 1982-1994	57
De danske partimedlemmer 2000	58
Fritidsundersøgelsen 1998, børn	59
Fritidsundersøgelsen 1998, voksne	59
Arbejde og hjemmeliv før og nu 1950-1988	60
Beretning for Dansk Data Arkiv 2002	61
English Summary	65

Om Metode & Data

Dansk Data Arkiv udgiver fremover tidsskriftet *Metode & Data* én gang årligt. Tidsskriftet bringer dels artikler om aktuelle metodiske og videnskabsteoretiske overvejelser indenfor samfundsvidenskab, sundhedsvidenskab og historie, dels artikler med resultater af empiriske studier inden for samme fagområder.

Tidsskriftets ambition er at udfylde en niche med en faglig og videnskabelig diskussion af metoder, metodeanvendelse og dataanalyse indenfor samfundsvidenskab, sundhedsvidenskab og historie.

Metode & Data udkom første gang i 1999 som en videreførelse af DDA-Nyt, og tidsskriftet rummer fortsat færdiganmeldelser af studier, der er oparbejdet i henhold til DDA's standard (studiestatusklasse FOD) samt nyt fra DDA.

Metode & Data modtager meget gerne manuskripter inden for de ovennævnte fagområder.

Redaktionelt

Af BERNHARD HANSEN, DDA

I forrige nummer af Metode & Data præsenterede Dorte Gyrd-Hansen og Richard Jones resultaterne af en spørgeskemaundersøgelse angående befolkningens ønsker om sundhedsmæssige forebyggelsespakker. Forfatterne viste, at svar på spørgsmål med samme indholdsmæssige betydning om interesse for sundhedsmæssige forebyggelsespakker i høj grad varierede afhængigt af den kontekst, de blev stillet i. Et beslægtet tema behandles af Hanne-Pernille Stax i dette nummer. I artiklen „Med sproget som måleredskab“ vises det mere generelt, hvorledes spørgsmål med samme indholdsmæssige betydning kan føre til ganske store utilsigtede (eller tilsigtede!) variationer i svarmønstre alene som følge af ganske små sproglige og designmæssige variationer i spørgeskemaet. Hanne-Pernille Stax nøjes dog ikke med at opregne problemer forbundet med spørgeskemadesign, men giver også forskellige anvisninger på metoder til testning og kvalitetsudvikling af spørgeskemaers sproglige kvalitet med henblik på at undgå de værste faldgruber. Det er metoder, som alle, der arbejder med design af spørgeskemaer og analyser af spørgeskemadata, bør forholde sig til. Dansk Data Arkiv designer ikke selv spørgeskemaer; til gengæld arbejdes der meget med spørgeskemadata i Dansk Data Arkiv. Den væsentligste del af dette arbejde består i den såkaldte oparbejdning af de data, som forskere og andre dataindsamlere lader deponere i Dansk Data Arkiv. Oparbejdningssprocessen, der dokumenterer data og sikrer langtidssikret opbevaring i systemuafhængige formater, beskrives nøjere i artiklen „Bevaring af viden – Om oparbejdning af data i Dansk Data Arkiv“ af Anne Sofie Fink Kjeldgaard et al. Artikelens hensigt er at præsentere læseren for det arbejde, som deponerede data gennemgår. Arbejde

med og oparbejdning af spørgeskemadata og andre typer af samfundsvidenskabelige data udgør den ene af Dansk Data Arkivs to hovedaktivitetsområder; det andet er organiseringen af indtastning af folketællingsdata. Hvad angår dette arbejde efterlyser Ole S. D. Hansen i et kortere indlæg flere seriøse studier med udgangspunkt i data fra den færdigindtastede 1801 folketælling; han opregner selv flere mulige afsæt for sådanne analyser. I den efterfølgende artikel præsenterer Nanna Floor Clausen resultater af nogle indledende analyser af enkestandens leveforhold netop på baggrund af data fra 1801 folketællingen. Nanna Floor Clausen undersøger blandt andet variationer med hensyn til enkestandes køns-, alders-, erhvervs- og geografiske fordelinger.

Udover disse artikler indeholder nummeret beskrivelser/færdiganmeldelser af fjorten datamaterialer af forskellig karakter, der siden udgivelsen af Metode & Data nr. 88 har gennemgået Dansk Data Arkivs oparbejdningsprocedure. Det sidste indslag i dette nummer er Dansk Data Arkivs årsberetning for 2002.

Afslutningsvis skal det nævnes, at Metode & Data fremover kun vil udkomme fast en gang årligt mod de nuværende to gange. Dette skyldes dels ressourcemæssige hensyn og dels et bevidst ønske om at opgradere arbejdet med andre former for formidling i form af nyhedsbreve, temanumre, samt intensiveret formidling via Dansk Data Arkivs hjemmeside. Selvom Metode & Data først udkommer næste gang i efteråret 2004, modtages manuskriptudkast (på maksimalt 10.000 ord) med metodiske overvejelser og/eller resultater af empiriske studier inden for samfunds- og sundhedsvidenskab og historie fortsat særdeles gerne. God læselyst.

Med sproget som måleredskab

Om udforskning, testning og forbedring af vilkår for forståelse i spørgeskemabaseret forskning

Af HANNE-PERNILLE STAX, INSTITUT FOR ALMEN OG ANVENDT SPROGVIDENSKAB, KØBENHAVNS UNIVERSITET

1. Introduktion

„Too much can be inferred from answers taken at face value to questions of dubious merit. [] all answers depend upon the way a question is formulated. Language is not a clean logic tool like mathematics that we can use with precision“ (Schuman 1982: 22).

Sikring af kvalitet i spørgeskemabaserede undersøgelser er en central fordring indenfor både forskning, evaluering og markedsanalyser. Men hvor for eksempel stikprøveudvælgelse og statistisk bearbejdning af spørgeskemadata har udviklet sig til selvstændige videnskabelige discipliner, så foregår formulering af spørgeskemaspørgsmål og svarmuligheder hyppigt på baggrund af erfaringsbaserede anbefalinger og fornemmelser for, hvad der er mest hensigtsmæssigt. I Danmark er empiriske undersøgelser af, hvorledes spørgeskemaspørgsmål bliver forstået, normalt ikke en naturlig del af det metodologiske arbejde med at sikre kvalitet i spørgeskemaundersøgelser. Der er få danske studier om forståelse af spørgeskemaspørgsmål (se: Andersen & Christoffersen 1982, Christoffersen 1984, Kjølner 1975, Körmendi m.fl. 1986, Olsen 1998, 2001), og der er et ret begrænset fokus på området indenfor surveybranchen¹.

I Danmark består en prætest af et spørgeskema hyppigt af en såkaldt pilot-test, hvor nogle interviewere afprøver et spørgeskema på nogle respondenter og giver tilbagemeldinger om udtalte problemer med bestemte spørgsmål. I for eksempel USA er det derimod en udbredt praksis at inddrage kognitionspsykologer og sprogforskere i udviklingen af spørgeskemadesign samt at teste konkrete skemaer, før de bringes i anvendelse² (se for eksempel Dippo & Norwood 1992, Schwarz & Sudman 1996). Ved testning anvendes primært såkaldte ekspert-reviews, kodning af interviewer-respøndent, interaktion eller introspektion, hvilket vil sige kvalitative interviews med respondenter om forståelse af udvalgte begreber, spørgsmål eller spørgeforløb. I forbindelse med mere generel metodeudvikling gennemføres der også større komparative studier af svarfordelinger ved forskellige udformninger af spørgsmål eller spørgeforløb.

I det følgende vil jeg præsentere og diskutere forskellige tilgange, der anvendes til udforskning og testning af respondents spørgsmåls- og opgaveforståelse. Jeg vil give eksempler på indsigter, som er tilvejebragt gennem disse tilgange for at illustrere, hvad man kan opnå ved at udforske sproglige, kognitive og interaktionelle aspekter af spørgeskemabaserede interviewsamtaler.

1. Der var for eksempel ingen panels om sproglig kvalitet på „The International Conference on Improving Surveys“ (Danmarks Statistik & Socialforskningsinstituttet, 2002), mens for eksempel de seneste konferencer under American Association for Public Opinion Research samt konferencerne Q2001 (Statistics Sweden) og „5th Conference on Logic and Methodology“ (International Sociologist Association, 2000) alle havde panels om blandt andet spørgeskemadesign, interviewinteraktion og testning af spørgsmålsformuleringer.
2. For eksempel har Bureau of Labor Statistics, U.S. Bureau of the Census og National Center for Health Statistics permanente laboratorier, der gennemfører tests af respondents forståelse af konkrete spørgsmål og spørgeforløb samt bidrager til generel metodeudvikling indenfor formulering og testning af spørgeskemadesign. „The Joint Program in Survey Methodology“ (Westat) udbyder kurser i spørgsmålsdesign og testning på lige fod med kurser i statistik og sampling.

Jeg vil først præsentere nogle klassiske komparative studier, som er gennemført med sigte på generel metodeudvikling. Fund fra komparative studier kan sjældent omsættes direkte til generelle regler for 'god spørgsmålsformulering,' men studierne demonstrerer, hvor følsomt et måleinstrument sproget kan være. Dernæst vil jeg behandle kognitive forståelsesinterviews, som primært anvendes til at teste og kvalitetsudvikle konkrete spørgeskemainstrumenter. En tredje tilgang til udforskning og testning af spørgeskemaers sproglige kvalitet er analyser af interaktion mellem interviewere og respondenter i autentiske interviews. Jeg vil her behandle to forskellige metodologier: Adfærdskodning og sekventiel mikroanalyse. Sidst vil jeg behandle såkaldte ekspertreviews, hvor man evaluerer og kvalitetsudvikler spørgeskemaer og spørgsmålsformuleringer ved at integrere forskellige indsigter opnået gennem empiriske studier af vilkår for forståelse i spørgeskemabaseret forskning.

Nogle af de behandlede tilgange fokuserer primært på spørgeskemaundersøgelser, hvor dataindsamlingen foregår ved interviews, enten i form af besøgsinterviews eller telefoninterviews. Men de fleste af de diskuterende fund er også relevante i forhold til spørgeskemaundersøgelser baseret på selvudfyldelse.

2. Baggrund: Om spørgsmål, sprog og forståelse i spørgeskemaundersøgelser

„A survey trades in meaning, and meanings are embodied in language. A survey consists of a transfer of meaning between the participants through the medium of language“ (Bateson 1984).

Formålet med en spørgeskemaundersøgelse er, på baggrund af sammenlignelige informationer fra mange respondenter, at kunne belyse udbredelse og distribution af for eksempel holdninger, adfærd eller leveforhold i en bestemt population.

Det har længe været almindeligt accepteret, at selv små ændringer i et spørgsmåls formulering kan resultere i markant anderledes svarfordelinger (se for eksempel Rugg 1941). Denne viden er omsat i en generel fordring om, at alle respondenter i en spørgeskemaun-

dersøgelse skal præsenteres for nøjagtig de samme spørgsmålsformuleringer; ellers kan deres svar ikke sammenlignes (se for eksempel Hansen & Andersen 2000: 107). Nyere forskning i forskellige aspekter af respondents forståelse af spørgsmål og respondentopgaver har imidlertid rejst tvivl om, hvorvidt denne fordring virkelig er det mest centrale element i sikring af sammenlignelige forståelser af spørgsmål i spørgeskemaundersøgelser. Det, at alle respondenter får samme verbale input, sikrer ikke nødvendigvis, at de alle forstår det, der var forskerens intention med et spørgsmål. Man kan naturligvis anlægge det synspunkt, at hvis man bare spørger tilstrækkeligt mange, så vil effekterne af eventuelle individuelle forståelsesforskelle udligne hinanden. Men det ændrer ikke ved det forhold, at små sproglige detaljer i et spørgsmåls formulering kan have stor betydning for, hvordan svarene fordeles sig.

En del forskere sværger vedholdende til, at det altid er at foretrække at fastholde selv en dårligt fungerende spørgsmålsformulering i undersøgelse efter undersøgelse for at sikre, at de genererede svar er sammenlignelige over tid. Andre pointerer, at hvis et spørgsmål ikke bliver forstået og besvaret som intenderet, så er den sammenlignelighed, der opnås ved gentagne anvendelser, ret beset en sammenlignelighed mellem to lige ubekendte: De genererede svar er nok sammenlignelige i en vis formel forstand, men vi ved bare ikke, hvad de er svar på.

Som sprogpsykolog med en baggrund i forskning i forståelse og interaktion er det oplagt interessant at beskæftige sig med spørgeskemaundersøgelser. Det er svært at forestille sig en interaktionsform, hvor det er mere afgørende, at de enkelte spørgsmål og de enkelte svarmuligheder bliver forstået sådan, som forskeren har intenderet. Hvis det ikke er tilfældet, ved forskeren jo ikke, hvad svarene er svar på. Men hvordan får man indsigt i, hvordan respondenter forstår de spørgsmål, de svarer på? Og hvordan sikrer man sig, at de forstår det samme ved de stillede spørgsmål, som forskeren har ment med dem?

I forbindelse med spørgeskemaundersøgelser anvendes begrebet 'sort boks' som en metafor for faser eller processer, som forskeren ikke har nogen direkte indsigt i. En sort boks er som en tryllekunstners hat.

Man putter noget ned i den og tager noget andet op, og man ved ikke, hvad der skete undervejs – hvordan det ene blev til det andet. Det, der foregår fra et spørgeskema forlader forskerens skrivebord, og til det kommer tilbage i udfyldt tilstand – eller omsat til tabeller og grafer – kan beskrives som én stor sort boks for den forsker, der ikke er involveret i processen. Men processen fra skema til data kan også betragtes som en række sorte bokse. Nogle af disse sorte bokse er det umiddelbart muligt at 'kigge' ind i. For eksempel kan en forsker, der designer en interviewbaseret spørgeskemaundersøgelse, selv deltage i instruktionen af interviewere, og forskeren kan overvære eller selv gennemføre nogle interviews for at få indblik i, hvordan sådan nogle forløber i praksis (se for eksempel Hansen & Andersen 2000: 139).

Men der er også sorte bokse, som man ikke umiddelbart kan kigge ind i. For eksempel kan man ikke få indblik i, hvordan respondenter forstår og producerer svar på spørgsmål blot ved at overvære eller deltage i et interview. Det, der foregår i respondentens hoved under et interview eller en udfyldelse af et spørgeskema, kan man kun få adgang til ad omveje. Komparative studier af svarfordelinger ved forskellige spørgsmålsformuleringer og kognitive interviews om respondenters spørgsmålsforståelse er to bud på, hvordan man kan få indsigt i, hvordan respondenter forstår forskellige spørgsmålsformuleringer. Adfærdskodning og sekventiel mikroanalyse af interviewer-respondent interaktion er bud på, hvordan man kan få indsigt i, hvad der foregår mellem interviewer og respondent i spørgeskemabaserede interviews.

3. Komparative studier – spørgsmålsdesign og svarfordelinger

„Two priests, a Dominican and a Jesuit, are discussing whether it is a sin to smoke and pray at the same time. After failing to reach a conclusion, each goes off to consult his respective superior. The next week they meet again. The Dominican says ‘Well, what did your superior say?’ The Jesuit responds ‘He said it was all right.’ ‘That’s funny,’ the Dominican replies, ‘my superior said it was a sin.’ Jesuit: ‘What did you ask him?’ Reply:

‘I asked him if it was all right to smoke while praying.’ ‘Oh,’ says the Jesuit, ‘I asked if it was alright to pray while smoking.’“ (Sudman & Bradburn 1982: 1)

I komparative studier anvendes spørgeskema-metoden som instrument i eksperimentelle tests af konkrete spørgsmålsdesign. Fremgangsmåden er at teste forskellige – mulige eller faktiske – sproglige operationaliseringer af samme erkendelsesinteresse. To sammenlignelige respondentgrupper præsenteres for eksempel for forskellige formuleringer af 'det samme spørgsmålsindhold' for identiske spørgsmålsformuleringer placeret i forskellige rækkefølger eller for svarskalaer i forskellige udformninger. Hvis disse eksperimentelle variationer resulterer i forskellige svarfordelinger i sammenlignelige respondentgrupper, er der påvist en såkaldt svareffekt (se for eksempel: Sudman & Bradburn 1974).

Komparative studier har godtgjort, at selv ganske små variationer i spørgsmålsformulering, i spørgsmåls interne rækkefølge eller i udformning, antal og rækkefølge af svarmuligheder kan resultere i signifikant – og ofte markant – forskellige svarfordelinger.

I det følgende præsenteres nogle klassiske komparative studier, hvor forskellige typer af svareffekter er påvist. Disse undersøgelser illustrerer, at små sproglige detaljer i en spørgeskematekst kan være afgørende for, hvordan svarene fordeler sig. Det er naturligvis ikke overraskende, at forskellige spørgsmål giver forskellige svar. Men komparative studier kan illustrere, at to forskere med samme erkendelsesinteresse – tilfældigt eller overlagt – kan operationalisere deres erkendelsesinteresser i minimalt forskellige sproglige formuleringer og få markant forskellige svarfordelinger.

3.1 Enighedseffekt

Eksempel 1 illustrerer en type svareffekt, der er påvist i en række komparative studier. Effekten kan beskrives som en præference for at udtrykke enighed. De to spørgsmålsversioner A og B præsenterer to komplementære udsagn om årsagen til kriminalitet i USA. Men andelen af respondenter, der erklærer sig enige, er ret konstant – uanset hvilken påstand, der fremsættes i spørgsmålsformuleringen. 'Det omvendte spørgsmål' resulterer altså ikke i 'den omvendte svarfordeling'.

A	B
Individuals are more to blame than social conditions for crime and lawlessness in this country	Social conditions are more to blame than individuals for crime and lawlessness in this country
Agree 59,6% (individuals)	Agree 56,8%
Disagree 40,4%	Disagree 43,2% (individuals)

Eksempel 1.
Enighedseffekt

Schuman & Presser 1981: 208 (Gennemført i USA i 1974. N=473/472, „Ved ikke“ udeladt)

Implikationerne af enighedseffekter er, at den rapporterede tilslutning til et undersøgt synspunkt tilsyneladende påvirkes af, hvordan anmodningen om stillingtagen formuleres. Det, at præsentere et synspunkt i en deklarativ spørgsmålsformulering synes i sig selv at øge den rapporterede tilslutning til synspunktet.

Enighedseffekter betragtes generelt som fejlkilder, men effekten kan også udnyttes positivt, for eksempel i forsøg på at motivere respondenter til at tilkendegive deres eventuelle sympatier for synspunkter, der anses for at være 'politisk ukorrekte,' og som det ellers kan være svært at få frem. Det er dog indlysende, at man må være varsom med at drage konklusioner om den 'reelle' tilslutning til et synspunkt – og dets eventuelle implicite modsynspunkter – på baggrund af svar på deklarativ spørgsmål. Svarfordelinger på deklarativ spørgsmål kan nok indikere relative holdningsbevægelser over tid eller mellem forskellige grupper. Men isoleret set vil svarfordelinger på spørgsmål som de ovenstående næppe kunne siges at afspejle etablerede holdninger til for eksempel kriminalitetens årsager.

Når man betragter de to spørgsmålsversioner i denne komparative kontekst, så kan de måske begge forekomme overlagt 'skæve' og ledende. Men hvis man fokuserer på dem ét ad gangen, så vil man kunne gen-

kende et spørgsmålsformat, som er meget udbredt i spørgeskemabaserede holdningssundersøgelser: Nemlig serier af påstande eller deklarativer, som respondenter skal erklære sig enige eller uenige i. På grund af risikoen for enighedseffekter anbefaler en del surveymetodologer, at man helt undgår deklarativ spørgsmål med enig/uenig svar og i stedet søger at formulere mere balancerede spørgsmål som for eksempel: „Which in your opinion are more to blame for crime and lawlessness in this country – individuals or social conditions?“

Enighedseffekter i forbindelse med deklarativ spørgsmålsformuleringer er påvist i et stort antal komparative studier. Ikke desto mindre er spørgsmålsformatet fortsat meget udbredt i såvel danske som internationale spørgeskemaundersøgelser.

3.2 Forbudt/tilladt-effekt

Eksempel 2 illustrerer en anden type svar-effekt, som er fundet i en række komparative eksperimenter, nemlig en såkaldt 'forbudt/tilladt-effekt'. Effekten består i en større tilbøjelighed til at udtrykke modstand imod at tillade noget, der er 'potentielt følsomt,' end til udtrykke støtte til at forbyde det. Nogle af de første komparative eksperimenter, der belyser denne effekttype, blev gennemført i USA i 1940'erne og handlede om amerikaneres villighed til at tillade/forbyde

A	B
Do you think the United States should forbid public speeches in favor of communism?	Do you think the United States should allow public speeches in favor of communism?
Yes 39,3%	No 56,3%
No 60,1% (not forbid)	Yes 43,8% (allow)

Eksempel 2.
Tilladt/forbudt-effekt

Schuman & Presser 1981: 281 (Gennemført i USA i 1976. N=409/432, „Ved ikke“ udeladt)

offentlige taler imod demokrati (se for eksempel: Roper 1940; Rugg 1941). Eksemplet nedenfor er af lidt nyere dato, men udfaldet ligger i forlængelse af resultaterne af de tidligere studier.

Ud fra en rent logisk betragtning skulle man forvente, at et 'nej' til at forbyde kommunistiske taler i version A ville være indholdsmæssigt ækvivalent med et 'ja' til at tillade disse taler i version B. Men det komparative studie gendriver denne tese. Og de to versioner kunne danne grundlag for helt forskellige – ja ligefrem modsatrettede – konklusioner om holdninger til ytringsfrihed eller kommunister i den undersøgte population.

Når de to spørgsmålsformuleringer A og B står overfor hinanden som her, så forekommer det måske umiddelbart indlysende, at den første version er en mere 'barsk' formulering end den anden, og dermed kan den fundne svareffekt forekomme forudsigelig. Men hvilken spørgsmålsformulering giver mon det mest 'sandfærdige' billede? Og af hvad? En forsker, der ønsker at belyse en populations holdninger til for eksempel kommunister eller ytringsfrihed, må vælge konkrete formuleringer. Og to forskere med den samme erkendelsesinteresse kunne vælge hver sin version og – overlagt eller uoverlagt – få to meget forskellige resultater.

Den såkaldte 'forbudt/tilladt-effekt' er påvist i adskillige komparative studier (se for eksempel: Hippler & Schwarz 1986; Holleman 1999). Eksempel 3 er fra Olsen 1998.

Dette komparative eksperiment viser samme tendens. Respondenterne har en større tilbøjelighed til at sige 'nej' til at forbyde end til at sige 'ja' til at tillade den samme ting. Også i dette eksempel genererer forbudt-versionen altså en større rapporteret tolerance overfor noget potentielt følsomt,

end tilladt-versionen gør.

Men forbudt/tilladt-effekter er ikke stabile – hverken over tid eller på tværs af forskellige spørgsmålssemner. Et eksperimentelt studie i 1979 viste for eksempel ingen signifikant forskel i udtrykt holdning til cigaretreklamer i spørgsmålsformuleringerne: „Do you think the government should forbid cigarette advertisements on television?“ henholdsvis „Do you think the government should allow cigarette advertisements on television?“ (Schuman & Presser 1981: 283).

Gentagne studier har demonstreret, at forbudt/tilladt effekter er udbredte – men også, at deres forekomst og effekt er umulig at forudsige (se for eksempel: Hippler & Schwarz 1986; Holleman 1999). 'Effekternes effekt' varierer over tid og fra emne til emne. De afhænger for eksempel af emnets aktualitet samt af de enkelte respondents engagement i emnet. Endelig er det vist, at denne type svareffekt i nogle tilfælde er stærkere blandt respondenter med kortere uddannelser end blandt respondenter med længere uddannelser. Det sidstnævnte fund har indlysende problematiske implikationer i studier, der sammenholder rapporterede holdninger med uddannelsesniveau. Det er for eksempel muligt, at en spørgsmålsformulering med 'tillade' vil vise stor forskel i svarfordelinger mellem respondenter med kort, henholdsvis lang uddannelse, mens en formulering med 'forbyde' vil vise en mindre forskel (se for eksempel Schuman & Presser 1981: 279).

Effektens uforudsigelighed betyder, at det i hvert enkelt tilfælde er nødvendigt at teste, om et konkret tilladt/forbudt-spørgsmål er 'sensitivt' overfor effekttypen. Og for hvert eneste forbudt/tilladt-spørgsmål man møder på sin vej – som forsker, journalist eller lægman – må man overveje, hvorvidt det

Eksempel 3.
Tilladt/forbudt-effekt

A	B
Mener De, at foreninger og partier, der udbreder rasediskriminerende (racistiske) synspunkter, bør være forbudt i Danmark?	Mener De, at foreninger og partier, der udbreder rasediskriminerende (racistiske) synspunkter, bør være tilladt i Danmark?
Ja: 56%	Nej: 66%
Nej: 31% (ikke forbudt)	Ja: 20% (tilladt)

Olsen: 1998:366 (Gennemført i Danmark i 1996. N = 968/920, „Ved ikke“ = 13-14%)

kunne have genereret en markant anderledes svarfordeling, hvis 'forbudt' var erstattet med 'tilladt' eller omvendt.

Det skal her bemærkes at enighedseffekter og tilladt/forbudt-effekter kan påvirke det samme spørgsmåls svarfordelinger i hver sin retning. En spørgsmålsformulering som: „Også racistiske grupper bør have ret til at holde møder og tale deres sag“ (SFI/Togeby 1999) vil for eksempel kunne påvirkes af begge effekttyper. Og hvor enighedseffekt forventeligt vil motivere den rapporterede tilslutning til synspunktet, vil formuleringen med 'ret til' (frem for for eksempel 'Man skal forbyde...') forventeligt mindske den rapporterede tilslutning til racistiske grupper ytringsfrihed.

3.3 Konteksteffekter

Komparative eksperimenter har også belyst, hvordan den interne rækkefølge, som spørgsmål stilles i, kan have betydning for den svarfordeling, de genererer. Spørgsmålet nedenfor blev i én skemaversion stillet i en 'neutral' kontekst, det vil sige efter nogle spørgsmål uden emnemæssig relation til hverken kommunister, journalister eller ytringsfrihed.

I en anden skemaversion blev det samme spørgsmål stillet efter et spørgsmål om amerikanske journalisters arbejde i kommunistiske lande:

I denne kontekst svarede omtrent dobbelt så mange respondenter 'ja' til at tillade journalister fra kommunistiske lande at arbejde i og rapportere fra USA. Det komparative

Kontekst A

Do you think the United States should let Communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?

Yes: 36,5%

Schuman & Presser 1981: 29 (Gennemført i USA i 1948. N= 581, „Ved ikke“ udeladt)

Eksempel 4A.
Konteksteffekt

Kontekst B

Do you think a Communist country like Russia should let American newspaper reporters come in and send back to America the news as they see it?

Do you think the United States should let Communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?

Yes: 73,1%

Schuman & Presser 1981: 29 (Gennemført i USA i 1948. N = 635, „Ved ikke“ udeladt)

Eksempel 4B.
Konteksteffekt

	1948	1980
Communist reporter question FIRST (or 'alone')	Yes: 36,5%	Yes: 54,7%
Communist reporter question AFTER American reporter question	Yes: 73,1%	Yes: 74,6%

Schuman & Presser 1981: 29 (Gennemført i USA i 1980. N = 342 / 335, „Ved ikke“ udeladt)

Eksempel 4C.
Konteksteffekt over tid

eksperiment blev gentaget i 1980 (eksempel 4C).

Som det ses, er konteksteffekten for dette specifikke spørgsmålspar reduceret over tid. Eksperimentet viser således, at også rækkefølgeeffekter er ustabile. Men eksperimentets resultater illustrerer en mere generel problematik angående indlån og genbrug af tidligere anvendte spørgsmålsformuleringer i nye sammenhænge: Hvis vi for eksempel forestiller os, at spørgsmålet om journalister fra kommunistiske lande blev stillet i kontekst A i 1948 og derefter gentaget i den samme kontekst i 1980, så ville vi finde en relativ stigning i tilslutningen til at tillade disse journalister at arbejde i USA. Men hvis spørgsmålet om journalister fra kommunistiske lande blev stillet i kontekst B 1948 og derefter gentaget i kontekst A i 1980, så ville man komme til det modsatte resultat, nemlig et fald i tilslutningen til at tillade disse journalister at arbejde i USA. Endelig er det værd at bemærke, at hvis spørgsmålet om journalister fra kommunistiske lande både i 1948 og i 1980 blev stillet i kontekst B, så ville man kunne 'vise,' at holdningen til at tillade kommunistiske journalister at arbejde i USA slet ikke havde udviklet sig i de forløbne 32 år.

Konteksteffekter er påvist i en lang række komparative eksperimenter med varierende resultater (se for eksempel Schuman & Presser 1981, Schwarz & Sudman 1992). Konteksteffekter er således ligesom enighedseffekter og tilladt/forbudt-effekter en type svareffekter, hvis implikationer spørgeskemaforskere altid må være opmærksomme på. Der er afdækket en del ret robuste konteksteffekttyper, og disse fund er udmøntet i mere generelle regler for

god spørgeskemaudformning. Som eksempel kan nævnes, at specifikke spørgsmål om et emne stillet FØR et mere generelt spørgsmål om samme emne vil påvirke besvarelsen af det generelle spørgsmål. Det anbefales derfor som hovedregel at stille generelle spørgsmål før specifikke spørgsmål om relaterede emner (se for eksempel Schuman & Presser 1981, Schwarz m.fl 1991; Schwarz m.fl 1991b; Gendall m.fl. 1997).

3.4 Svarskalaeffekt

Komparative eksperimenter er også anvendt til at belyse effekter af forskellige udformninger af svarskalaer. Det er for eksempel vist, at sproglig 'forankring' af svarskalaer, det vil sige valg af sproglige formuleringer ved numeriske svarskalaers (ende)punkter, kan være afgørende for, hvorledes svarene fordeler sig på disse skalaer. Nedenfor ses to forskellige fempunktsskalaer. I et komparativt eksperiment rapporterede respondenter, der fik skala A, for eksempel en væsentligt lavere gennemsnitlig tilfredshed end respondenter, der fik skala B (se: Guterbock & Hubbard 2000).

Også nummerering af punkterne på en svarskala kan være afgørende for, hvorledes svarene fordeler sig. Schwarz m.fl. (1991a) viste, at en 11-punkts-skala i version B nedenfor genererede flere svar 'over middel,' det vil sige til højre for skalaens midtpunkt end skalaen i version A. De verbale formuleringer ved skalaernes endepunkter var i begge versioner: „Not at all successful“ henholdsvis „Extremely successful“ (se Schwarz 1995; Schwarz m.fl. 1991a). Det er også vist, at frekvensskalaers udform-

Eksempel 5.
Svarskalaeffekt: Sproglig
forankring

A				
Not at all satisfied				Very satisfied
☐	☐	☐	☐	☐
1	2	3	4	5

B				
Very dissatisfied				Very satisfied
☐	☐	☐	☐	☐
1	2	3	4	5

A										
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
0	1	2	3	4	5	6	7	8	9	10

Eksempel 6.
Svarskalaeffekt:
Nummerering

B										
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
-5	-4	-3	-2	-1	0	1	2	3	4	5

A: Timers tv-forbrug pr dag	B: Timers tv-forbrug pr dag
Up to ½ hour ☐	Up to 2 ½ hour ☐
½ hour to 1 hour ☐	2 ½ hour to 3 hour ☐
1 hour to 1 ½ hour ☐	3 hour to 3 ½ hour ☐
1 ½ hour to 2 hour ☐	3 ½ hour to 4 hour ☐
2 hour to 2 ½ hour ☐	4 hour to 4 ½ hour ☐
More than 2 ½ hour ☐	More than 4 ½ hour ☐

Eksempel 7.
Frekvensskalaeffekt

Schwarz 1995: 159

ning kan påvirke den rapporterede frekvens af en undersøgt adfærd. I en undersøgelse fra 1985 af tyskeres daglige tv-forbrug fik to sammenlignelige grupper henholdsvis skala A og skala B nedenfor. I gruppen, der fik skala A, rapporterede kun 16,2 pct., at de så fjernsyn mere end 2 1/2 time pr dag, mens 37,5 pct. af respondenterne, der fik skala B, rapporterede et dagligt TV-forbrug på mere end 2 1/2 time.

For en kritisk gennemgang af studier af svarskalaeffekter se Krosnick & Fabrigar 1997.

3.5 Komparative studier af andre effekttyper

Komparative eksperimenter har belyst en lang række andre effekttyper, end de ovenfor nævnte. Her følger nogle referencer til studier af nogle undersøgte effekttyper: Effekter af *spørgsmålslængde*, se: Catania m.fl. 1996; Bradburn & Sudman 1979. Effekter af *svarrækkefølge*, se: Harrison & McLaughlin 1993; Krosnick & Alwin 1987. Effekter af *indsamlingsmåde*, for eksempel telefoninterview versus selvudfyldelse eller besøgsinterview, se: Bradburn & Sudman 1979; Körmendi m.fl 1986, Christoffersen 1984. Effekter af strengt *standardiseret ver-*

sus konversationel interviewing, se: Schober & Conrad 1997. Effekter af at præsentere *en eller to sider af en sag* i samme spørgsmål, se: Bishop m.fl. 1982. Effekter af, at interviewer formulerer *opfølgende spørgsmål* (probes) med alle relevante henholdsvis udvalgte relevante svarmuligheder, se: Smit m.fl. 1997. Effekter af at *instruere respondenter* om svaropgaven undervejs, se: Cannell m.fl. 1981.

3.6 Styrker og svagheder ved komparative studier

Komparative studier af svareffekter kan være en øjenåbner med hensyn til, hvor stor en rolle sproglige valg kan spille for svarfordelinger i spørgeskemaundersøgelser. Men hvilke antagelser som sprog og interaktion hviler disse studier selv på? Af praksis i komparative studier kan det udledes, at metodologien arbejder med en distinktion mellem 'sprogligt indhold' og 'sproglig form.' Dét, der angiveligt undersøges i komparative eksperimenter, er, om 'det samme spørgsmålsindhold' i forskellige 'sproglige udformninger' eller forskellige kontekster genererer forskellige svarfordelinger. Det er indlysende, at der ikke er noget nyt i at vise, at forskellige spørgsmål giver forskellige svar. Dét, der er den implicite grundantagelse bag

de komparative eksperimenter er, at det rent faktisk er 'det samme spørgsmål,' der testes på to sammenlignelige grupper. Hvorledes man afgør, hvorvidt to spørgsmålsvarianter er varianter over 'det samme spørgsmål' eller blot to forskellige spørgsmål, er ikke ekspliciteret. Det antages at kunne afgøres ved hjælp af almindelig sund fornuft. Frem for at eksplicite en bagvedliggende ækvivalensteori, søger komparative studier ved eksemplets magt at anskueliggøre, at to forskere, der ønsker at belyse samme fænomen – i god eller ond tro – kunne vælge to forskellige formuleringer og dermed nå til to helt forskellige resultater.

Opsummerende har komparative eksperimenter demonstreret, at små variationer i spørgsmålsdesign kan afstedkomme store variationer i de genererede svarfordelinger. Men studier af forskellige svareffekter har også demonstreret, at de er uforudsigelige og dermed svære at foregribe eller kompensere for. Styrken ved komparative studier er, at de demonstrerer, hvor følsomt et måleinstrument sproget er, og hvor varsom man skal være med at drage sikre konklusioner om etablerede holdninger og forhold og den undersøgte virkelighed på baggrund af målinger, der er genereret gennem sproglig interaktion. Set fra et anvendt perspektiv er den største svaghed ved komparative studier, at de nok kan demonstrere, at forskellige design giver forskellige svarfordelinger, men ikke hvilke design, der giver de 'bedste' eller mest valide svar.

Men hvordan opnår man så en bedre forståelse af respondents spørgsmålsforståelse og svarproduktion i spørgeskemabaserede undersøgelser? Kognitive forståelsesinterviews er en mulig tilgang til belysning af, hvordan respondenter forstår og besvarer konkrete spørgeskemaspørgsmål. Kognitive forståelsesinterviews anvendes dels i kvalitetstestning af konkrete spørgeskemainstrumenter, men metoden anvendes også i generel metodeudvikling med det sigte at kunne generere generelle regler for god spørgsmålsformulering og dermed at kunne forudsige og foregribe uønskede svareffekter.

4. Kognitive interviews – individuel spørgsmålsforståelse og besvarelse

Prætests af konkrete spørgsmålsformuleringer ved hjælp af kognitive interviews med respondenter om deres forståelse af spørgsmål og svaropgaver har ikke så lang en tradition som komparative studier. Det samarbejde mellem kognitionspsykologer og spørgeskememetodologer, som nu er etableret i især USA, blev indledt i begyndelsen af 1980'erne. Centrale manifestationer af det begyndende samarbejde var de to konferencer „Cognitive Aspects of Survey Research“ i Maryland, USA (se Jabine m.fl. 1984) og „International Conference on Social Information Processing and Survey Methodology“ i Mannheim, Tyskland (se: Hippler m.fl. 1987).

Den 'sorte boks,' som kognitive interviews søger indblik i, er respondentens individuelle kognitive processer. Helt overordnet er formålet med kognitive interviews at belyse, hvorvidt respondents opfattelse og forståelse af spørgsmål og respondentopgaver svarer til forskerens intention med dem.

De fleste kognitive interviewstudier tager udgangspunkt i en generel model af de kognitive opgaver, som en respondent skal løse i forbindelse med besvarelse af et spørgsmål. Disse opgaver er: Spørgsmålsforståelse, informationsgenskabelse (eller -konstruktion), informationsselektion og svarformulering (eller -valg) (se for eksempel Sudman m.fl. 1996, kap. 3; Tourangeau m.fl. 1990.)

Kognitive interviews gennemføres for at undersøge, hvordan disse kognitive opgaver løses i forbindelse med besvarelse af konkrete spørgeskemaspørgsmål, og hvorvidt opgaverne løses som intenderet.

Mange spørgeskemaundersøgelser søger at belyse, om der er forskelle i holdninger eller livsvilkår mellem for eksempel mænd og kvinder, yngre og ældre eller respondenter med lange vs. korte uddannelser. Men en meningsfuld sammenligning af de svar, som forskellige respondentgrupper giver på samme spørgsmål, fordrer naturligvis, at disse grupper har sammenlignelige forståelser af de spørgsmål, de svarer på. Kognitive interviews anvendes også til at belyse, om forskellige respondentgrupper evt. forstår og besvarer de samme spørgsmål på usammenlignelige måder.

Kognitive interviewmetoder kan groft inddeles i følgende: Verbale protokoller – også kaldet tænke-højt-interview, parafrasering og kognitiv probing. Disse tilgange kombineres hyppigt i kognitive interviewstudier. Nedenfor behandles de dog hver for sig.

4.1 Tænke-højt-protokoller – tilgange og centrale fund

En verbal protokol er en rapport om en respondents 'indre mellemregninger' i forbindelse med forståelse og besvarelse af en opgave (se for eksempel Ericsson & Simon 1980). Disse rapporter kan enten produceres sideløbende med respondentens løsning af svaropgaven, altså samtidig med at vedkommende forstår og besvarer konkrete spørgeskemaspørgsmål, eller man kan lade respondenterne besvare spørgeskemaets spørgsmål først og så efterfølgende – eller retrospektivt – bede dem om at sætte ord på deres mellemregninger og eventuelle problemer i forbindelse med svarprocessen. Det er en udbredt opfattelse, at retrospektive protokoller er nemmere at producere for respondenterne og ikke risikerer at influere de kognitive processer, som der ønskes information om. Til gengæld er der risiko for, at respondenterne glemmer nogle mellemregninger, hvis der først skal rapporteres om dem efter, at svaropgaven er løst (se for eksempel Schwarz 1997: 36, Redline m. fl. 1998).

Ideen med verbale protokoller er, at respondenterne skal fortælle interviewer, hvad vedkommende tænker i forbindelse med besvarelsen af konkrete spørgeskemaspørgsmål. Og interviewer skal være så lidt styrende som muligt. Verbale protokoller er for eksempel anvendt til at belyse respondenteres tankeprocesser ved besvarelse af spørgsmålet i eksempel 4 ovenfor om kommunistiske journalisters adgang til at arbejde i USA (se Bishop 1992). Efter besvarelse af spørgsmålet i kontekst A henholdsvis kontekst B (se eksempel 4) blev respondenter stillet følgende opgave (se eksempel 8):

De verbale protokoller viste en tendens til at respondenter, der besvarede spørgsmålet om kommunistiske journalisters adgang til

at arbejde i USA først, fokuserede på deres holdninger til kommunister, kommunisme og spionage i deres besvarelse. Respondenter, der først besvarede dette spørgsmål efter spørgsmålet om amerikanske journalisters adgang til at arbejde i kommunistiske lande, fokuserede derimod på lighed for alle journalister og på journalistisk frihed (Bishop 1992: 153ff).

Verbale protokoller overlader det til respondenterne at vælge, hvilke dele af svarprocessen, der skal rapporteres om og hvilke aspekter af de stillede spørgsmål, der skal kommenteres. Dette åbne design gør, at verbale protokoller kan give forskere information om problemer i svarprocessen, som de ikke selv havde nogen forventninger om. Til gengæld risikerer forskeren, at de enkelte respondenter fokuserer på helt forskellige ting, og at de ikke adresserer problemstillinger, som forskeren på forhånd er interesseret i at få belyst. En alternativ tilgang, der sikrer at respondenterne rapporterer om bestemte aspekter af svaropgaverne, er kognitiv probing.

4.2 Kognitiv probing – tilgange og centrale fund

Når man anvender kognitive probes i kognitive interviews om respondenteres spørgsmålsbesvarelse, beder man respondenterne om at forklare, hvorledes de har håndteret bestemte aspekter af svaropgaven. Man kan for eksempel spørge til, hvordan de har forstået udvalgte ord eller vendinger, hvorledes de har beregnet en frekvens eller et antal eller hvilke ting, de har lagt vægt på i en besvarelse. Kognitiv probing forberedes ved en tekstbaseret analyse af et spørgeskema. Her udpeges potentielle problemer i forbindelse med spørgsmålsforståelse, informationsgen-skabelse, informationssortering eller svarproduktion. Derefter udformes spørgsmål, der kan belyse, om respondenterne faktisk har problemer i forbindelse med løsningen af disse dele af svaropgaverne, eller om forskellige respondenter løser opgaverne på meget forskellige måder således, at svarene bliver usammenlignelige.

Now tell me all that you can remember about your thinking as you answered this question.

Bishop 1992: 148

Eksempel 8.
Verbal protokol

Johnson m.fl. 1997 anvender for eksempel kognitive probes til at belyse forskellige dele af respondents svarprocesser og relaterer fundene til forskellige respondent-karakteristika. Studiet viser blandt andet, at der er substantielle forskelle i, hvorledes respondenter med forskellig kulturel baggrund løser de samme svaropgaver, både hvad angår spørgsmålsforståelse, informationsgenskabelse, informations-sortering og svarproduktion. I eksempel 9 ses en probe, der er designet til at belyse spørgsmålsforståelse hos respondenter, der identificerede sig selv som henholdsvis 'Afro-American,' 'Hispanic' eller 'White non-Hispanic.'

Johnson m.fl. (1997: 96 ff.) fandt, at respondenter, der identificerede sig selv som 'Hispanic', hyppigere nævnte økonomiske problemer, når de skulle angive betydningen af 'stress' end respondenter med anden kulturel baggrund. Eksperimenter viste i øvrigt ingen sammenhæng mellem respondents indkomst og fokusering på økonomisk stress. Samme eksperiment viste også, at respondenter, som identificerede sig selv som 'White non-Hispanic', mindre hyppigt end respondenter med anden baggrund nævnte problemer i sociale relationer, når de skulle angive betydningen af 'stress'.

I et dansk studie (Olsen 2001) afdækkede kognitive probes blandt andet, at respondenter kunne forstå ordet 'værelser' ret forskelligt, sådan at nogle respondenter, der typisk boede i lejlighed, talte alle opholds- og soverum med som værelser i deres bolig, mens andre respondenter, der typisk boede i huse, ikke talte stuer, men kun soveværelser og børneværelser med (Olsen 2001: 1: 2004).

Studiet viser altså, at selv om ordet værelse ikke kan betegnes som et 'vidt begreb', så kan de to forskellige mulige forståelser give ret forskellige svar (se Olsen 2001: II: 95 ff.).

Implikationen af forskelle i spørgsmålsforståelse kan for eksempel være at forskelle i svarfordelinger mellem forskellige respondentgrupper faktisk ikke afspejler reelle forskelle i respondenternes virkelighed (for eksempel deres stressniveau eller boligforhold) men alene, at de forstår et anvendt ord forskelligt og derfor ikke inkluderer de samme informationer i deres svar.

Johnson m.fl (1997: 98 ff) viste også at den måde, hvorpå respondenter genskaber information, sorterer information og producerer svar, kan variere med forskellige respondentkarakteristika. I eksempel 10 ses en kognitiv probe, der er designet til at belyse respondenternes informationsgenskabelse.

Studiet viste, at ca. halvdelen af respondenterne forsøgte at genskabe og tælle hver enkelt relevant episode i det forløbne år. Denne procedure betegnes episodisk genskabelse. Godt 40 pct. af respondenterne baserede deres svar på en generel frekvensberegning. Denne genskabelsesstrategi betegnes semantisk genskabelse, og da kan respondenter for eksempel tage udgangspunkt i, at han 'normalt' går til lægen ca. hver anden måned og derfor svare: 6 gange. Godt 5 pct. af respondenterne opgav, at deres svar var baseret på et gæt.

Studiet viste en systematisk sammenhæng mellem genskabelsesstrategi og rapporterede lægebesøg: De respondenter, som svarede ud fra episodisk genskabelse, opgav gennemsnitligt knap 3 lægebesøg. De respondenter, der svarede ud fra semantisk genskabelse,

Eksempel 9. Kognitiv probing

Spørgsmål:	During the past 2 weeks, would you say you experienced a lot of stress, a moderate amount of stress, relatively little stress, or almost no stress at all?
Probe:	What does the word 'stress' mean to you?

Johnson m.fl. 1997: 97

Eksempel 10. Kognitiv probing

Spørgsmål:	During the last year, how many times did you see or talk to a medical doctor?
Probe:	How did you arrive at this answer?

Johnson m.fl. 1997: 99

opgav gennemsnitligt ca. 10 lægebesøg, mens de respondenter, der gættede gennemsnitligt, opgav ca. 7 lægebesøg.

Den anvendte kognitive probe kan ikke afdække, om respondentens valg af genskabelsesstrategi faktisk påvirker det rapporterede antal lægebesøg, eller om hyppigheden af lægebesøg omvendt determinerer respondentens valg af genskabelsesstrategi. Men studiet sætter spørgsmålstegn ved, om de afgivne svar er reelt sammenlignelige. Og det indikerer, at det kunne være formålstjenligt, at forskere i deres spørgsmål specificerer den ønskede genskabelsesstrategi. Altså, om der ønskes information om en 'normal' eller gennemsnitlig frekvens eller om konkret adfærd i en helt specifik periode.

4.3 Parafrasering

I kognitive interviews anvender man også parafraseringsopgaver, hvor man beder respondenterne om at gentage konkrete spørgsmål med deres egne ord. Parafraseringsopgaver betragtes hyppigt som en særlig form for kognitiv probing (se for eksempel Forsyth & Lessler 1991). Sigtet er at få informationer om, hvordan respondenter forstår spørgsmålet og svaropgaven, og metoden har ligesom verbale protokoller den fordel, at respondenter selv sætter dagsordenen. Metoden kan altså give informationer om problemtyper, som forskeren ikke selv var opmærksom på. Metoden har imidlertid også de samme ulemper som verbale protokoller. Den sikrer ikke, at man får informationer om aspekter af svarprocessen, som på forhånd anses for potentielt problematiske. Dertil er det ofte uklart, om et evt. misforhold mellem spørgsmålsintention og parafraseret formulering skyldes et misforhold mellem forskerens spørgsmålsintention og respondenterens spørgsmålsforståelse eller mellem respondenterens spørgsmålsforståelse og respondenterens *efterfølgende parafrasering*. En yderligere svaghed ved tilgangen er, at nogle respondenter simpelthen gentager spørgsmål mere eller mindre ordret, og det giver jo ikke megen ny information.

4.4 Styrker og svagheder ved kognitive interviews

Kognitive prætest-interviews kan give spørgeskemaforskere indblik i aspekter af, hvordan forskellige respondenter opfat-

ter de anmodninger om information, som er spørgeskemaundersøgelsens bærende medium. Ved at gennemføre blot et lille antal kognitive interviews er det muligt at afdække uklarheder og utilsigtede fortolkningsmuligheder, som kunne resultere i svar, der ikke er sammenlignelige eller ikke belyser det, som forskeren ønsker at belyse. Dermed er kognitive interviews et uvurderligt redskab i sikring og udvikling af sproglig kvalitet i spørgeskemaundersøgelser.

En hyppigt fremsat indvending imod værdien af kognitive interviewstudier er dog, at det ikke er indlysende, at de kognitive opgaver, som respondenter udfører og rapporterer om i kognitive interviews, svarer til de opgaver, som de ville udføre i et autentisk spørgeskemabaseret interview. De informationer, man får fra kognitive interviews, afhænger meget af respondenterne evne og parathed til at sætte ord på deres mellemregninger. Dertil stiller et kognitivt interview respondenter 'til regnskab' for sine svar, og at dette forhold antages at kunne påvirke hele svarprocessen således, at den vil afvige fra en svarproces i et almindeligt interview. For behandlinger af validitet i forskellige typer kognitive interviews se for eksempel Russo m.fl. 1989, Willis m. fl. 1991.

Man kan også anføre, at kognitive prætest-interviews kendetegnes ved et ret ensidigt fokus på respondenter og dennes – angiveligt – individuelle kognitive arbejde i forbindelse med spørgsmålsforståelse og svarproduktion. Man kunne formulere det sådan, at belysning af svarprocesser ved kognitive interviews forudsætter, at det faktisk er respondenter alene, der foretager alle mellemregninger, fra et spørgsmål er præsenteret, og til svaret er formuleret. Men forudsætningen om, at respondenter interagerer 'rent' med forskerens spørgsmålsformuleringer, holder hyppigt ikke i spørgeskemaundersøgelser, der gennemføres ved interview. En lang række studier af interaktionen i spørgeskemabaserede interviews har vist, at der foregår megen anden interaktion mellem interviewer og respondenter end 'korrekt' oplæsning af de præformulerede spørgsmål og produktion af svar, der umiddelbart kan registreres som en præformuleret svarmulighed. Respondenter og interviewere interagerer for eksempel hyppigt i forbindelse med afklaring af spørgsmåls betydning eller af hvilke formuleringer, der

kan gælde som svar. For eksempel finder Fowler & Cannell (1996: 25) i en analyse af 250 autentiske interviews, at kun 33 pct. af alle spørgsmål oplæses ordret (se også for eksempel Cannell m.fl. 1981; Oksenberg m.fl. 1991; Houtkoop-Steenstra 2000; Maynard m.fl. 2002).

Dermed må man slutte, at kognitive interviews belyser svarprocesser i idealiserede interviewforløb, hvor opgaveforståelse, spørgsmålsforståelse og svarproduktion antages at være respondentens individuelle bedrifter. Når man betragter data fra kognitive interviews som vidnesbyrd om, hvordan en respondent rent faktisk forstår spørgsmålsformuleringer, erindrer og udvælger relevant information og formulerer registrerbare svar, så har man forudsat, at det rent faktisk er individuelle kognitive opgaver, som respondenter løser alene. Og det er det altså hyppigt ikke. Hvorvidt lokale afvigelser fra det præformulerede manuskript for det spørgeskemabaserede interview i sig selv udgør et metodologisk problem skal ikke berøres her (men se for eksempel: Beatty 1995, Schober & Conrad 1997, Suchman & Jordan 1990). På dette sted skal det blot påpeges, at hvis opgaveforståelse og svarproduktion er opgaver, som respondenter hyppigt løser i interaktion med interviewer, så må man behandle fund og anbefalinger, der alene baserer sig på studier, der ikke inddrager denne interaktion med rettidigt forbehold.

Analysen af den faktiske interaktion mellem interviewere og respondenter er bud på, hvordan man kan belyse interaktive aspekter af forståelses- og svarprocesser i spørgeskemabaserede undersøgelser.

5. Interaktionsanalyser – interaktion i samtale med manuskript

Analysen af den faktiske interaktion mellem respondenter og interviewere er siden 1970'erne blevet anvendt i forbindelse med kvalitetsudvikling af spørgeskemabaserede interviews. Såkaldt *adfærds-kodning* anvendes til at identificere problemer i spørgsmålsformuleringer og interviewerpraksis. *Sekventielle mikroanalyser* af interaktionen mellem interviewere og respondenter anvendes endvidere til at belyse, hvilke konsekvenser standardisering af interviewets interaktion har for deltagerens mulighed for at forstå inten-

tionen med de stillede spørgsmål og de producerede svar. Adfærds-kodning er en normativ tilgang, hvor spørgsmål-svar-sekvenser, der består af et korrekt oplæst spørgsmål efterfulgt af et umiddelbart registrerbart svar, er idealet og betragtes som udtryk for problemfrie spørgsmålsformuleringer og forbilledlig intervieweradfærd. I sekventielle mikroanalyser anlægges der derimod det perspektiv, at sekvenser, der afviger fra denne norm, kan give indsigt i spørgeskemainterviewets generelle interaktionelle dynamik.

5.1. Adfærds-kodning – tilgange og centrale fund

Adfærds-kodning af interaktionen mellem interviewere og respondenter er i udgangspunktet blevet anvendt til at undersøge, om *interviewere* fulgte retningslinierne for god interviewpraksis. Det vil sige, om de oplæste spørgsmålene ordret, præsenterede alle relevante svarmuligheder og ikke påvirkede respondenteres spørgsmålsforståelse eller svarvalg. Idealet er her, at interviewererne siger præcis det og kun det, som er præformuleret i spørgeskemateksten.

Det materiale, som underkastes analyse ved adfærds-kodning, er båndoptagelser af autentiske spørgeskemainterview. Ved adfærds-kodning tildeles alle tilfælde af ideel eller paradigmatisk adfærd en bestemt kode, mens forskellige former for afvigelse fra spørgeskematekstens manuskript tildeles andre koder. Kodelisterne vil variere fra studie til studie, men der vil hyppigt være koder for intervieweradfærd som for eksempel: 'Korrekt stillet spørgsmål,' 'Spørgsmål oplæst med ubetydelig variation,' 'Spørgsmål oplæst med meningsforstyrrende variation,' 'Spørgsmålsfortolkning,' 'Neutral anmodning om specifikation af svar,' 'Ledende anmodning om specifikation af svar' (se for eksempel Kjølner 1975, Fowler & Cannell 1996). Ved hjælp af en sådan kodeliste kategoriserer man intervieweres adfærd og vurderer, om de følger de gældende retningslinier for standardiseret interviewing. I disse studier vil hyppige afvigelser fra spørgeskematekstens manuskript blive betragtet som indikation på mangelfuld interviewerinstruktion eller dårlig interviewerpraksis.

Adfærds-kodning anvendes også som tilgang til at belyse, om konkrete spørgsmålsformuleringer og spørgeforløb er 'velfun-

Spørgsmål: How much did you pay, or will you have to pay, out of pocket for your most recent visit? Do not include what insurance has paid for or will pay for. If you don't know the exact amount, please give me your best estimate.

Eksempel 11.
Spørgsmål, der
forårsagede mange afvigelser
fra manuskript

Oksenberg m.fl. 1991: 355

gerende.' Her inddrages også respondentens adfærd i adfærdskodningen. Typiske koder for respondentadfærd vil være for eksempel: 'Registrerbart svar,' 'Ikke registrerbart svar,' 'Anmodning om gentagelse,' 'Anmodning om hjælp til fortolkning af spørgsmål,' 'Afbrydelse af spørgsmål under oplæsning' (se for eksempel: Oksenberg m.fl. 1991).

Hvis en interviewer oplæser et spørgsmål 'korrekt', og respondenter på sin side besvarer spørgsmålet umiddelbart efter interviewerens første oplæsning og producerer et svar, der matcher en af de præformulerede svarmuligheder, så kan hele spørgsmåls-svarsekvensen betegnes som 'paradigmatisk'. En sådan sekvens indikerer, at spørgsmålet har fungeret uproblematisk for både interviewer og respondent. Hvis derimod interviewerens eller respondentens adfærd afviger fra spørgeskematekstens manuskript, kan det ses som en indikation på, at der er problemer i dette manuskript. I eksempel 11 ses et eksempel på et spørgsmål, som resulterede i mange tilfælde af afvigelser fra manuskriptet. Interviewere oplæste hyppigt ikke spørgsmålet ordret, og respondenter afbrød hyppigt interviewerne under oplæsningen, bad om hjælp til fortolkning eller producerede svar, der ikke matchede de præformulerede svarmuligheder.

Kodningen alene kan kun indikere, at der er problemer, men ikke afklare hvorfor. At respondenter hyppigt afbryder interviewere under oplæsning af et spørgsmål kunne dog tyde på, at spørgsmålet er formuleret sådan, at det *lyder* afsluttet, før det er det. Og Oksenberg m.fl. (1991) anbefaler da også, at ovenstående spørgsmål omformuleres, så det ikke kan forekomme at være afsluttet, før det rent faktisk er det.

Et eksempel på et 'generelt' fund fra adfærdskodninger er, at spørgsmål, der rummer mulighed for dobbelte negationer, hyppigt giver anledning til længere interaktioner end blot spørgsmål og svar. For eksempel vil en spørgsmålsformulering som „Vi skulle ikke have indført fri abort“ (SFI/Togeby

1999) hyppigt udløse svaret „Jo“. Og derefter kræver det en del interaktion at bringe svaret på en format, der passer en svarskala som 'Helt enig,' 'Delvis enig,' 'Delvis uenig' eller 'Helt uenig.'

Adfærdskodning er en oplagt tilgang til belysning af, om interaktionen mellem interviewere og respondenter forløber som planlagt. Og adfærdskodninger kan give et hurtigt indblik i, hvorvidt standardiserede spørgsmålsformuleringer rent faktisk omsættes i standardiserede interviewinteraktioner. Talrige studier, der har anvendt adfærdskodning på autentiske interview har vist, at det gør de meget hyppigt ikke.

En oplagt svaghed ved adfærdskodning er, at de alene kan lokalisere *erkendte* problemer. En adfærdskodning kan afdække, at en respondent har problemer med at forstå et spørgsmål. Men den kan for eksempel ikke afdække, om en respondent – uden erkendte problemer – misforstår et spørgsmål eller forstår noget andet ved det, end forskeren havde intenderet. Den dikotomiske skelnen mellem paradigmatisk og ikke-paradigmatisk sekvenser som adfærdskodning tager udgangspunkt i og er baseret på en antagelse om, at det paradigmatisk er det ideelle. Men denne antagelse er – netop – en antagelse, som tilgangen ikke selv kan underbygge.

En anden svaghed er, at adfærdskodning nok tager udgangspunkt i autentiske interviews, men meget hurtigt transformerer denne interaktion til en række separate præanalytiske kategorier. Det er oplagt, at den ene type adfærd fører den anden med, og at man mister megen information om interaktionens dynamik ved at behandle hvert enkelt træk i interviewet som en separat adfærd (se for eksempel Maynard & Schaeffer 2002a: 20). For eksempel kan man pege på, at det er en simplificering blot at kategorisere en respondentadfærd som 'afbrydelse' uden at inddrage, hvorledes denne afbrydelse forholder sig til det afbrudte interaktion (om sekventiel adfærdskodning se dog van der Zouwen & Dijkstra 2002). En anden tilgang, der søger

at belyse spørgeskemabaserede interview som en sekventielt organiseret interaktion, er *Conversation Analysis*, som jeg her vil omtale som sekventiel mikroanalyse.

5.2. Sekventiel mikroanalyse – tilgange og centrale fund

„Interviews in which there are ‘deviations’ from the standardized script and the paradigmatic question-answer sequences make visible the ways in which answering a question – and the cognitive processing it requires – can be fundamentally interactive and collaborative rather than a respondent’s own problem“ (Schaeffer & Maynard 1996: 67).

Sekventielle mikroanalyser arbejder ikke ud fra en præanalytisk norm for god eller paradigmatiske interaktion. Derimod behandles alle træk – eller al handlen – i en interaktion som deltagerens meningsfulde bidrag til kollektiv tilvejebringelse af fælles forståelse i interaktionen. Suchman & Jordan (1990) pegede i en tidlig artikel om sekventiel mikroanalyse af spørgeskemainterview på, at når interviewere afviger fra spørgeskemaets formuleringer, så kunne det være tegn på at et præformuleret spørgeskema ikke er et velegnet manuskript for en mundtlig samtale. Detaljerede analyser af *hvorledes* interviewere afviger fra manuskripter kunne således give indsigt i, hvad der – fra intervieweres perspektiv – skal til for at få spørgsmål til at fungere i praksis.

I en sekventiel mikroanalyse arbejder man lige som ved adfærdskodning med båndoptagelser af autentiske interview. Men formålet er ikke at kode, kategorisere eller vurdere interviewerens og respondentens tale. I stedet analyseres detaljerne i de enkelte talerbidrag som meningsfuld handlen, uanset om de matcher eksternt definerede normer eller retningslinier. Og hvert enkelt talerbidrag analyseres som en demonstration af, hvordan taleren har forstået umiddelbart foregående talerbidrag. For eksempel kan en respondents afbrydelse af en interviewer under oplæsningen af et spørgsmål ses som en demonstration af, at respondenter mener at have hørt tilstrækkeligt til at kunne producere en respons. Omvendt kan en pause, efter at en respondent har produceret et svar, ses som en demonstration af, at intervieweren ikke har accepteret svaret som en fyldestgørende besvarelse.

Sekventielle mikroanalyser har lige som adfærdskodning i udgangspunktet været anvendt til at belyse aspekter af *intervieweres* adfærd i spørgeskemabaserede interviews. For eksempel har Houtkoop-Steenstra (2000) vist, hvordan intervieweres afvigelser fra spørgeskemateksten hyppigt afspejler orientering imod generel praksis for høflig og opmærksom samtaledeltagelse, og Maynard & Schaeffer (2002b) og Houtkoop-Steenstra & van den Bergh (2002) har studeret forskellige mønstre i og implikationer af intervieweres forsøg på at opnå respondenters tilsagn om deltagelse i interview.

Det kendetegnende for sekventielle mikroanalyser er dog, at de analyserer deltagerens bidrag som elementer i sekventielt organiserede forløb. Houtkoop-Steenstra (2002) har for eksempel analyseret sekvenser, hvor intervieweren ikke oplæser spørgsmål til deres præformulerede afslutning og vist, at sådanne utidige afbrydelser hyppigt sker som respons på, at en respondent begynder at svare på et spørgsmål før, det er oplæst i sin fulde længde. Og sådanne utidige besvarelser påbegyndes hyppigt lige omkring et sted i en spørgsmålsformulering, som kunne have været spørgsmålets afslutning. Dette studie viser, hvornår respondenter *oplever*, at et spørgsmål er afsluttet nok til at svare på, og analysen udpeger ’sårbare punkter’ i spørgsmålsformuleringer, som bør omformuleres, hvis man ønsker, at respondenterne skal høre hele spørgsmålet, før de svarer.

Stax (2003) har analyseret sekvenser, hvor interviewere kun tilbyder respondenterne én svarmulighed, selv om der er flere relevante muligheder i spørgeskemateksten. Dette studie viste, at i disse sekvenser begynder respondenterne hyppigt at svare lige omkring det sted, hvor intervieweren har oplæst én svarmulighed, og altså før intervieweren kan nå at præsentere flere. Studiet viser, at det ikke er intervieweren alene, der afgør, hvor mange svarmuligheder respondenterne skal præsenteres for før, vedkommende svarer. Det viser også, hvordan nogle præsentationer af svarmuligheder er mere udsatte for utidige afbrydelser end andre (om implikationer af selektiv præsentation af svarmuligheder se for eksempel Smit m.fl. 1997).

Sekventielle mikroanalyser kan give indblik i interaktionelle mekanismer i interview-samtaler, der er baseret på et præformuleret

manuskript. Og de kan gøre det anskueligt for forskere og brugere af spørgeskemaundersøgelser, hvordan interaktionen mellem interviewer og respondent er en betydende faktor i tilvejebringelsen af spørgeskemadata. En svaghed ved tilgangen er, at den kun kan give indblik i aspekter af svarprocesser, der er egentligt interaktive. Komparative studier og kognitive interviews (se ovenfor) kan siges at forudsætte, at respondenter i praksis løser alle dele af svaropgaverne alene. Disse studiers fund giver således indblik i aspekter af respondenters individuelle håndtering af svaropgaver i paradigmatisk eller idealiserede spørgsmål-svar-skevenser. Sekventielle mikroanalyser kan til gengæld ikke vriste megen information ud af paradigmatisk sekvenser. Hvis en respondent producerer et svar umiddelbart efter, at en interviewer har oplæst et spørgsmål, så demonstrerer respondenter blot at have forstået spørgsmålet som et spørgsmål, som vedkommende kan svare på. *Hvad* respondenter har forstået, kan en interaktionsanalyse ikke belyse.

Sekventielle mikroanalyser vil således hyppigt fokusere på sekvenser, der ikke forløber 'som planlagt'. Men sigtet er ikke at identificere, tælle eller kritisere afvigelser fra den præformulerede plan. Sigtet er at belyse, hvordan deltagerne i fællesskab løser de problemer, der skal løses. Sekventielle mikroanalyser kan således give input til at komme nogle af disse problemer i forkøbet enten ved at forbedre den præformulerede tekst eller ved at tillade en konstruktiv interaktion (se for eksempel Schober & Conrad 1997).

6. Ekspertreview – tekstbaseret analyse

De forskellige tilgange til udforskning og testning af sproglig og interaktionel kvalitet i spørgeskemabaserede interviewundersøgelser, som er diskuteret overfor, sigter alle på at tilvejebringe viden om respondenters og intervieweres forståelse af spørgsmål og svaropgaver. Og formålet med at tilvejebringe en sådan viden er at kunne sikre – eller i det mindste forbedre mulighederne for at etablere – fælles forståelse af interviewets spørgsmål og svar.

De diskuterede tilgange anvendes hyppigt i forbindelse med test og kvalitetssikring af konkrete spørgeskemaer, og fundene fra

sådanne studier kan være svære at generalisere til andre spørgsmålsformuleringer eller andre spørgeskemaundersøgelser. Empiriske studier har dog tilvejebragt en del robuste fund, som er udmøntet i en række generelle anbefalinger angående hensigtsmæssige spørgsmålsdesign, svarformuleringer og interviewerinstruktioner til brug i spørgeskemaundersøgelser (se Sudman & Bradburn 1982; Converse & Presser 1986; Foddy 1993).

Generelle anbefalinger baseret på empiriske studier kan fungere som støtte i arbejdet med at formulere spørgeskemaspørgsmål. De kan også anvendes som checklister ved evaluering og kvalitetsudvikling af eksisterende spørgeskemadesign. Det er blandt andet dét, man gør i et såkaldt ekspertreview.

Ekspertreviews er tekstbaserede problemanalyser af færdige eller næsten færdige spørgeskemadesign og spørgsmålsformuleringer. Disse analyser tager udgangspunkt i teoretiske forståelser af kognitive og interaktive processer i spørgeskemabaserede interviews og/eller erfaringer fra empiriske studier af sprog, kognition og interaktion i spørgeskemabaserede interviews.

I et komparativt studie af forskellige metoder til prætestning af spørgeskemainstrumenter finder Presser & Blair (1994), at ekspertreviews er mest reliable og mest 'produktive.' Det vil sige, at ekspertreviews resulterede i identifikation af flere potentielle problemkilder i et spørgeskemainstrument end pilottestning, adfærdskodning og kognitive interview. På den baggrund anbefaler Presser & Blair (1994: 95), at spørgeskemaer rutinemæssigt underkastes ekspertreviews, ligesom andre tekster underkastes sproglig kvalitetssikring før de publiceres.

En indlysende svaghed ved ekspertreview er, at de jo ikke kan identificere 'nye' problemer. De må nødvendigvis tage udgangspunkt i eksisterende teorier og analyser af andre instrumenter. En anden problematik er også nævnt ovenfor i forbindelse med behandlingen af konteksteffekter. Identiske spørgsmålsformuleringer kan fungere forskelligt i forskellige kontekster. Dette forhold betyder naturligvis, at man også må være varsom med ureflekteret at overføre fund fra empiriske studier af konkrete spørgsmålsformuleringer eller spørgsmålsformater fra én kontekst til en anden.

Endelig er det indlysende, at det at finde mange potentielle problemkilder naturligvis ikke er det samme som at fastslå, at de vil resultere i uønskede svareffekter. Ikke desto mindre afdækker ekspertreviews meget hyppigt talrige eksempler på formuleringer, der er åbenlyst uhensigtsmæssige og afviger fra de mest gængse anbefalinger om formulering af spørgeskemaspørgsmål. Ekspertreviews kan naturligvis ikke altid erstatte problematiske formuleringer med problemfrie. Men de kan altid bidrage til en øget indsigt i, hvad de indsamlede svar kan være svar på.

7. Perspektiver – evaluering og kvalitetsudvikling

Kvalitet i spørgeskemabaserede undersøgelser er ikke kun et spørgsmål om sampling, bortfaldsprocenter og forsvarlig statistisk behandling af data. Det er heller ikke kun et spørgsmål om at tilstræbe en idealsituation, hvor alle respondenter „udsættes for fuldkommen ensartede spørgestimuli i form af nøjagtig ens formulerede spørgsmål, oplæst i nøjagtig samme rækkefølge af interviewere, der opfører sig fuldkommen ens, hvorefter respondenteren uden anden påvirkning end spørgestimuli afgiver sin respons“ (Hansen & Andersen 2000: 107). Dette ideal svarer til det, man anvender ved kontrollerede medicinske forsøg. Men selv en læge ville næppe insistere på, at alle patienter skal have nøjagtig samme dosis bedøvende medicin for at opnå samme effekt – for eksempel fuld narkose – uanset patienternes alder, vægt eller foregående indtag af anden medicin. Her vil opnåelse af samme effekt eller 'output' hos forskellige patienter forudsætte, at stimuli eller 'input' tilpasses den enkelte patient.

Ligesom man næppe kan forvente at en indsprøjtning vil have samme effekt på en læge som på alle hans forskellige patienter, så kan man heller ikke forvente, at alle respondenter forstår det samme ved en spørgsmålsformulering, som en forsker gør. Dette er et grundvilkår i al forskning, der anvender sproglig kommunikation som måleinstrument. Og en insisteren på, at alle respondenter får den samme sprogstimuli, synes ikke helt meningsfuld, hvis man ikke også interesserer sig for kvaliteten af denne stimuli og for, hvordan den fungerer i praksis.

Ovenfor har jeg forsøgt at præsentere

nogle forskellige tilgange til at undersøge, hvordan spørgeskemaspørgsmål fungerer i praksis. Disse tilgange kan bringe indsigter i interviewprocessen, som kan bidrage til forbedringer af den sproglige kvalitet i spørgeskemainstrumenter. Fund fra empiriske studier af spørgsmålsformuleringer og interviewforløb kan også bidrage til en øget bevidsthed om den 'kvalitative' karakter af svar indsamlet i kvantitative interviews. En sådan øget bevidsthed kunne måske opløse en ellers temmelig fastlåst dikotomisk skelnen imellem kvalitative og kvantitative data i interviewundersøgelser. Som Olsen (1998) indikerer med sin titel: „Tallenes talende tavshed“, så er det altid sundt at reflektere over de sproglige, kognitive og interaktionelle detaljer, der skjuler sig bag de forførende entydige tal som spørgeskemaundersøgelser genererer.

Heldigvis synes der at være en stigende interesse for at teste og søge at sikre sproglig kvalitet i spørgeskemaundersøgelser. I hvert fald når det gælder udvikling af nye spørgeskemaspørgsmål. Det er imidlertid en udbredt praksis at genbruge spørgsmål, der er anvendt før i andre spørgeskemaundersøgelser. Og i sådanne tilfælde vil en spørgeskemaforsker hyppigt foretrække at fastholde en eksisterende formulering uanset, om den forekommer sprogligt problematisk – for at sikre sammenlignelige svar. Igen kan man næppe forestille sig, at en læge ville insistere på at fortsætte med at give forsøgspersoner et medikament, som havde vist sig at have utilsigtede eller diffuse effekter eller bivirkninger blot for at sikre sammenlignelighed over tid. Han ville nok hurtigst muligt forsøge at udskifte medikamentet med et bedre alternativ, med sigte på at opnå mere anvendelige målinger – fremover.

Som udgangsbøn er der blot tilbage at konstatere, at nok er sproget et uregerligt og til tider upålideligt måleinstrument. Men når det drejer sig om at undersøge menneskers holdninger eller livsvilkår, så er der jo ingen alternativer til det. Hverken i kvalitative eller kvantitative interviewundersøgelser – eller i registerbaserede undersøgelser – kommer vi udenom, at sproglig interaktion er en medkonstituerende faktor for den producerede viden. Og hårde data er aldrig hårdere end de sproglige formuleringer og den sproglige interaktion, som de bygger på.

Referencer

- ANDERSEN, BJARNE H. & MOGENS N. CHRISTOFFERSEN (1982). *Om spørgeskemaer – Problemer ved spørgsmålsformulering i interviewundersøgelser*, København: Socialforskningsinstituttet.
- BATESON, NICHOLAS (1984). *Data construction in social surveys*, London: Allen & Unwin.
- BEATTY, PAUL (1995). „Understanding the standardized/non-standardized controversy“, *Journal of Official Statistics*, Vol. 11, pp. 147-160.
- BISHOP, GEORGE F.; ROBERT W. OLDENDICK & ALFRED J. TUCHFARBER (1982). „Effects of Presenting One Versus Two Sides of an Issue in Survey Questions“, *Public Opinion Quarterly*, Vol. 46, No. 1, pp. 69-85.
- BISHOP, GEORGE F. (1992). „Qualitative analysis of Question-Order and Context Effects: The Use of Think-Aloud responses“, pp. 149-162 i Schwarz, Norbert & Seymour Sudman (red.) *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research*, San Francisco: Jossey-Bass Publishers.
- BRADBURN, NORMAN M. & SEYMOUR SUDMAN (1979). *Improving Interview Method and Questionnaire Design*, San Francisco: Jossey-Bass Publishers.
- CANNELL, CHARLES F.; PETER V. MILLER & LOIS OKSENBERG (1981). „Research on Interviewing Techniques“, i Leinhardt, Samuel (red.) *Sociological Methodology 1981*, San Francisco: Jossey-Bass Publishers.
- CATANIA, JOSEPH A.; DIANE BINSON; JESSE CANCHOLA; LANCE M. POLLACK; WALTER HAUCK & THOMAS J. COATES (1996). „Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behavior“, *Public Opinion Quarterly*, Vol. 60, pp. 345-375.
- CHRISTOFFERSEN, MOGENS N. (1984). *Brev, telefon, besøg*, København: Socialforskningsinstituttet.
- CONVERSE, JEAN M. & STANLEY PRESSER (1986). *Survey Questions – Handicrafting the Standardized Questionnaire*, Newbury Park: Sage Publications.
- DIPPO, CATHRYN S. & JANET L. NORWOOD (1992). „A Review of Research at the Bureau of Labor Statistics“, i Tanur, Judith M. (red.) *Questions about Questions*, New York: Russell Sage Foundation.
- ERICSSON, K. ANDERS & HERBERT A. SIMON (1980). „Verbal Protocols as Data“, *Psychological Review* Vol. 87, pp. 215-251.
- FODDY, WILLIAM (1993). *Constructing Questions for interviews and Questionnaires*, Cambridge: Cambridge University Press.
- FORSYTH, BARBARA H. & JUDITH T. LESSLER (1991). „Cognitive Laboratory Methods: A Taxonomy“, pp. 393-418 i Paul Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz & Seymour Sudman (red.) *Measurement Errors in Surveys*, New York: John Wiley & Sons.
- FOWLER, FLOYD J. & CHARLES F. CANNELL (1996). „Using Behavioral Coding to Identify Cognitive Problems with Survey Questions“, i Schwarz, Norbert & Seymour Sudman (red.) *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research*, San Francisco: Jossey-Bass Publishers.
- GENDALL, PHILIP; VICTORIA CARMICHAEL & JANET HOEK (1997). „A Test of the Conversational Logic Analysis Model of Question Order Effects“, *Marketing Bulletin*, Vol. 8, pp. 41-52.
- GUTERBOCK, THOMAS M. & RYAN A. HUBBARD (2000). „Anchors Away: The effect of verbal anchors on responses to a satisfaction scale“, Paper presented at AAPOR 2000, Portland, Oregon, May 20, 2000. http://www.virginia.edu/surveys/press/aapor/scales3_files/frame.htm
- HANSEN, ERIK J. & BJARNE H. ANDERSEN (2000). *Et sociologisk værktøj*, København: Hans Reitzels Forlag.
- HARRISON, DAVID A. & MARY E. McLAUGHLIN (1993). „Cognitive Processes in Self-report responses: Tests of Item Context Effects in Work Attitude Measures“, *Journal of Applied Psychology*, Vol. 78, No. 1, pp. 129-140.
- HIPPLER, HANS-J. & NORBERT SCHWARZ (1986). „Not forbidding isn't Allowing“, *Public Opinion Quarterly*, Vol. 50, pp. 87-96.
- HIPPLER, HANS-J.; NORBERT SCHWARZ & SEYMOUR SUDMAN (1987). *Social Information Processing and Survey Methodology*, New York: Springer-Verlag.
- HOLLEMAN, BREGJE (1999). „The nature of the forbid/allow asymmetry“, *Sociological Methods & Research*, Vol. 28, No. 2, pp. 209-245.

- HOUTKOOP-STEENSTRA, HANNEKE (2000). *Interaction and the Standardized Interview: The Living Questionnaire*, Cambridge: Cambridge University Press.
- HOUTKOOP STENSTRA, HANNEKE (2002). „Questioning Turn Format and Turn-taking Problems in Standardized Interviews“, pp 243-260 i Maynard, Douglas W et al. (red.), *Standardizations and Tacit Knowledge – Interaction and Practice in the Survey Interview*, New York: John Wiley & Sons.
- HOUTKOOP-STEENSTRA, HANNEKE & HANS VAN DER BERGH 2002: „Effects of Introductions in Large-Scale Telephone Survey Interviews“, i Maynard, Douglas W. et al. (red.), *Standardizations and Tacit Knowledge – Interaction and Practice in the Survey Interview*, New York: John Wiley & Sons.
- JABINE THOMAS B.; MIRON L. STRAF; JUDITH M. TANUR & ROGER TOURANGEAU (RED.) (1984). *Cognitive Aspects of Survey methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press.
- JOHNSON, TIMOTHY P.; D. O’ROURKE; N. CHAVEZ; SEYMOUR SUDMAN; RICHARD B. WARNECKE; L. LACEY & JOHN HORM (1997). „Social Cognition and responses to Survey Questions Among Culturally Diverse Populations“ i Lyberg, Lars E. et al. (red.) *Survey Measurement and Process Quality*, New York: Wiley & Sons.
- KJØLLER, METTE (1975). *Interaktionen mellem interviewer og respondent*, København: Socialforskningsinstituttet.
- KROSNICK, JON A. & DUANE F. ALWIN (1987). „An evaluation of a cognitive theory of response order effects in survey measurement“, *Public Opinion Quarterly*, Vol. 51, pp. 201-219.
- KROSNICK, JON A., & LEANDRE R. FABRIGAR 1997: „Designing rating scales for effective measurement in surveys“, pp 141-164 i Lyberg, Lars E. et al. (red.), *Survey Measurement and Process Quality*, New York: Wiley & Sons.
- KÖRMENDI, ESZTER, LISBETH EGSMOSE & JOHANNES NOORDHOEK (1986). *Datakvalitet ved telefoninterview*, København: Socialforskningsinstituttet.
- MAYNARD, DOUGLAS W. & NORA C. SCHAEFFER (2002A). „Standardization and Its Discontents“, pp 3-46 i Maynard, Douglas W. et al. (red.), *Standardizations and Tacit Knowledge – Interaction and Practice in the Survey Interview*, New York: John Wiley & Sons.
- MAYNARD, DOUGLAS W. & NORA C. SCHAEFFER (2002B). „Opening and Closing the Gate: the work of optimism in Recruiting Survey Respondents“, pp. 179-204 i Maynard, Douglas W et al. (red.), *Standardizations and Tacit Knowledge – Interaction and Practice in the Survey Interview*, New York: John Wiley & Sons.
- OKSENBERG, LOIS; CHARLES CANNELL & GRAHAM KALTON (1991). „New Strategies for Pretesting Survey Questions“, *Journal of Official Statistics*, Vol. 7, No. 3, pp. 349-365.
- OLSEN, HENNING (1998). *Tallenes talende tavshed. Måleproblemer i surveyundersøgelser*, København: Socialforskningsinstituttet.
- OLSEN, HENNING (2001). *Sprogforståelse og hukommelse i danske surveyundersøgelser I & II*, København: Socialforskningsinstituttet.
- PRESSER, STANLEY & JOHNNY BLAIR (1994). „Survey Pretesting: Do Different Methods Produce Different Results?“, *Sociological Methodology*, Vol. 24, pp. 73-104.
- REDLINE, CLEO; RICHARD SMILEY; MEREDITH LEE; THERESA DEMAYO & DON DILLMAN (1998). „Beyond concurrent interviews: An evaluation of cognitive interviewing techniques for self-administered questionnaires“, *Proceedings of the Section on Survey Research Methods*, Alexandria: American Statistical Association.
- ROPER, ELMO (1940). „Problems and Techniques“, *Public Opinion Quarterly*, Vol 4, No. 1, pp. 129-130.
- RUGG, DONALD (1941). „Experiments in wording questions: II“, *Public Opinion Quarterly*, Vol. 5, No. 1, pp. 91-92.
- RUSSO J. EDWARD; ERIC J. JOHNSON & DEBRA L. STEPHENS (1989). „The Validity of Verbal Protocols“, *Memory and Cognition*, Vol. 17, pp. 759-769.
- SCHAEFFER, NORA C. & DOUGLAS W. MAYNARD (1996). „From Paradigm to prototype and back Again“, i Schwarz, Norbert & Seymour Sudman (red.) *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research*. San Francisco: Jossey-Bass Publishers.

- SCHOBER, MICHAEL F. & FREDERIC G. CONRAD (1997). „Does conversational interviewing reduce Survey Measurement Error“, *Public Opinion Quarterly*, Vol. 61, pp. 576-602.
- SCHUMAN, HOWARD (1982). „Artifacts are in the mind of the beholder“, *The American Sociologist*, Vol. 17, pp. 21-28.
- SCHUMAN, HOWARD & STANLEY PRESSER (1981). *Questions and Answers in Attitude Surveys*, New York: Academic Press.
- SCHWARZ, NORBERT (1995). „What respondents learn from questionnaires“, *International Statistical Review*, Vol. 63, No. 2, pp. 153-177.
- SCHWARZ, NORBERT (1997). „Questionnaire Design: The Rocky Road from Concepts to Answers“, pp. 29-46 i Lyberg, Lars E. et al. (red.) *Survey Measurement and Process Quality*, New York: Wiley & Sons.
- SCHWARZ, NORBERT; BARBEL KNAUPER; HANS-J. HIPPLER; ELISABETH NOELLE-NEUMANN & LESLIE CLARK (1991A). „Rating scales: numeric values may change the meaning of scale labels“, *Public Opinion Quarterly*, Vol. 55, No. 4, pp. 570-582.
- SCHWARZ, NORBERT; FRITZ STRACK & HANS-PETER MAI (1991B). „Assimilation and contrast effects in part-whole question sequences: a conversational logic analysis“, *Public Opinion Quarterly*, No. 55, Vol. 1, pp. 3-23.
- SCHWARZ, NORBERT & SEYMOUR SUDMAN (1992). *Context Effects in Social and Psychological Research*, New York: Springer-Verlag.
- SFI / TOGEBY, LISE (1999). „Danskerne syn på samfundet“ – Ikke publiceret spørgeskema anvendt i undersøgelse under Magtudredningen, Socialforskningsinstituttet.
- SMIT, JOHANNES H.; WIL DIJKSTRA & JOHANNES VAN DER ZOUWEN (1997). „Suggestive Interviewer Behaviour in Surveys: An Experimental Study“, *Journal of Official Statistics*, Vol. 13, No. 1, pp. 19-28.
- STAX, HANNE-PERNILLE (2003). „Veje til viden – om hensigtsmæssig kommunikation i spørgeskemabaserede interview“, pp. 289-309 i Asmuß, Birte & Jakob Steensig (red.) *Samtalen på arbejde. Konversationsanalyse og kompetenceudvikling*, København: Samfundslitteratur.
- SUCHMAN, LUCY & BRIGITTE JORDAN (1990). „Interactional troubles in face-to-face survey interviews“, *Journal of the American Statistical Association*, Vol. 85, pp. 232-241.
- SUDMAN, SEYMOUR & NORMAN BRADBURN (1974). *Response Effects in Surveys: A Review and Synthesis*, Chicago: Aldine.
- SUDMAN, SEYMOUR & NORMAN BRADBURN (1982). *Asking Questions: A Practical Guide to Questionnaire Design*, San Francisco: Jossey-Bass Publishers.
- SUDMAN, SEYMOUR; NORMAN BRADBURN & NORBERT SCHWARZ 1996: *Thinking about Answers: The Application of Cognitive Processes to Survey methodology*, New York: Russell Sage Foundation.
- Tourangeau, Roger; Lance J. Rips & Kenneth A. Rasinski (1990). *The Psychology of Survey Response*, Cambridge: Cambridge University Press.
- VAN DER ZOUWEN, JOHANNES & WIL DIJKSTRA 2002: „Testing Questionnaires Using Interaction Coding“, i Maynard, Douglas W. et al. (red.) *Standardizations and Tacit Knowledge – Interaction and Practice in the Survey Interview*, New York: John Wiley & Sons.
- WILLIS, GORDON B.; PATRICIA ROYSTON & DEBORAH BERCINI (1991). „The Use of Verbal Report Methods in the Development and Testing of Survey Questionnaires“, *Applied Cognitive Psychology*, No. 5, pp. 251-267.

Bevaring af viden – Om oparbejdning af data i Dansk Data Arkiv¹

Af ANNE SOFIE FINK KJELDGAARD, SØREN PRIISHOLM OG BERNHARD HANSEN, DDA

Det er kendt for de fleste, at Dansk Data Arkiv (DDA) modtager, opbevarer og udleverer elektroniske forskningsdata. Det betyder imidlertid ikke, at DDA er et rent ekspeditionskontor. Der foregår faktisk en del arbejde med data mellem de modtages og registreres og de eventuelle efterfølgende udleveringer. Den væsentligste del af databearbejdningen i DDA kaldes 'oparbejdning', der er en ganske tidskrævende proces. Oparbejdningen er i høj grad usynlig for omverdenen. Ikke desto mindre er der store fordele forbundet med at oparbejde data i forbindelse med senere brug af data. I artiklen redegøres der for rationalet bag oparbejdning af data, og processen beskrives ret indgående, ligesom en del af slutproduktet præsenteres. Hensigten er at give alle, der er interesseret i arbejde med samfunds- og sundhedsvidenskabelige forskningsdata, en klar fornemmelse af fordelene ved opbevaring af data i standardiserede og systemuafhængige formater.

Bevaring af et datamateriales bestanddele

Data, der opbevares i DDA, stammer fra forskere eller fra organisationer, der indsamle data. Data opbevares i DDA, men rettighederne til data beholdes af den afleverende person/organisation. Det vil sige, at det alene er personen/organisationen, der har afleveret data, der bestemmer, om og i hvilket omfang data eventuelt må udleveres til tredjepersoner. Hovedparten af de datamaterialer, der indleveres til opbevaring i DDA, er samfundsvidenskabelige spørgeskemadata, en mindre del er sundhedsvidenskabelige og historiske data. Da hovedparten af data er spørgeskemadata, tager beskrivelsen af

oparbejdningsprocessen udgangspunkt i netop spørgeskemadata.

Den overvejende del af spørgeskemadatamaterialerne hidrører fra enkeltstående tværsnitundersøgelser. Det vil sige, at der kun findes én datafil med tilhørende dokumentationsfil. Praksis i DDA er, at datamaterialer, der i et eller andet omfang relaterer sig til hinanden eksempelvis materialer fra undersøgelser, der gentages over tid og eventuelt med de samme respondenter, betragtes som separate datamaterialer. Sådanne materialer kan siges at bestå af to dele, nemlig data og dokumentation. DDA har én opgave i forhold til bevaring af data og én opgave i forhold til bevaring af dokumentation.

De rene data er blot en talmatrice; denne udgør det kvantitative datamateriales substans. I et spørgeskemamateriale er talmatricen talkoderne for de enkelte spørgsmåls svarkategorier og for eventuelle andre kategorier, som dataindsamleren har tilføjet datasættet.

Dokumentation er information, data skal tilføres for at give mening. Man kan for eksempel ikke i datamatricen se, hvilken information et ettal, henholdsvis total i en tilfældig kolonne indeholder. Kodebogen fortæller eksempelvis at casen/respondenten er en mand (kategori 1) i tilfældet af ettallet, mens casen/respondenten er en kvinde (kategori 2) i tilfældet af totallet.

Da datamaterialer basalt set består af to dele, har bevaringen også to aspekter. Dels skal selve talmaterialet bevares, og dels skal dokumentationen for tallenes oprindelse, indsamling og betydning bevares. For at sikre tallenes bevaring konverteres datamaterialer til et systemuafhængigt format OSIRIS (tegnformat), hvorfra data til enhver tid kan

1. Artiklen er baseret på en præsentation af procedurer for oparbejdning af data i Dansk Data Arkiv, som Søren Priisholm gav ved IASSIST – International Association for Social Science Information and Service and Technology – konferencen i Ottawa maj 2003.

konverteres til formater, der passer til diverse tidssvarende analysepakker som f.eks. SAS, STATA eller SPSS. Formatet OSIRIS III er valgt som arkivformat, idet det er et uhyre simpelt format udviklet specielt til at beskrive eksisterende datamaterialer. Når DDA modtager et datamateriale, arkiveres originaldata på en server. Er det muligt, foretages en konvertering af originaldata til OSIRIS-format med det samme, da der hermed fremstilles en fil, der er sikret langtidsbevaring. Den fysiske bevaring af data sikres ved placering på harddiske, der gennemgår en daglig back up-rutine.

Når forskere gør brug af tidligere indsamlede data, kræves der fuld indsigt i de praktiske og tekniske omstændigheder, undersøgelsen oprindeligt blev til under, derfor er det meget væsentligt, at der fremstilles pålidelig og brugbar dokumentation af de arkiverede datafiler. Den semantiske information bevarer ved, at datamaterialerne gennemgår en oparbejdningsproces, der indebærer, at al til rådighed stående information om data samles og på systematisk vis kobles til data. Dokumentationen har to hovedbestanddele: En overordnet beskrivelse af undersøgelsens ophav og produktion (studiebeskrivelse) og en nøgle til oversættelse af talkoderne i undersøgelsen til meningsfyldt information (kodebog).

Fordele ved oparbejdede data

Oparbejdning er en proces, et originalt datasæt gennemgår, når det dokumenteres fuldstændigt og omdannes til et langtidsholdbart arkivformat. Oparbejdning af datamaterialer er fordelagtigt både for forskeren/institutionen, der lader data deponere og for eventuelle efterfølgende brugere af data.

Fordelen for dataindsamleren ved at lade data deponere er først og fremmest, at vedkommende selv ikke behøver at holde styr på diverse filer og ikke behøver at omformatere filer i takt med udviklingen i software. Desuden er der åbenlyse samfundsmæssige fordele ved at have vigtige forskningsdata samlet ét centralt sted. For forskeren/institutionen, der vælger at lade data deponere ved DDA, betyder oparbejdning af data også, at al tilgængelig information om data integreres og gemmes. Drejer det sig om spørgeskemadata, er det en

væsentlig opgave at sikre, at kodebogen og spørgeskemaet bringes i takt. Er der koder, der ikke umiddelbart lader sig forklare af den indsendte dokumentation, vil DDA forsøge at indhente den relevante information. Dette er et arbejde, der naturligvis bliver vanskeligere jo længere tid siden, det er, forskeren selv har arbejdet med sit datamateriale. Alt i alt sikrer oparbejdning forskeren, at der findes en så fuldstændig og „langtidsholdbar“ version af data og dokumentation, som det er muligt at skabe.

Et oparbejdet datasæt giver i dets standardiserede format også fordele i forbindelse med efterfølgende brug af data både i forbindelse med søgning og identifikation af datasæt og i forbindelse med analyse. Med hensyn til søgning tilbyder DDA i dag brugeren at søge i informationen i studiebeskrivelser, men dette er faktisk kun en del af de søgemuligheder, der kunne udnyttes. Der arbejdes for tiden på et udviklingsprojekt, der skal gøre det muligt at søge i al informationen i kodebogen. Hermed vil det eksempelvis blive muligt at lade brugerne søge efter specifikke strenge/ord i variabelnavne.

Med hensyn til analyse af data betyder oparbejdningen især, at det bliver let at indhente information om datasættets tilblivelse og indhold, da al information er samlet i en såkaldt „datadokumentationspublikation“ (DDP). I kodebogen findes standardiserede koder for manglende data og kommentarer, hvoraf det eksempelvis vil fremgå, at data ikke er i overensstemmelse med dokumentationen, hvis dette er tilfældet. I forbindelse med efterfølgende brug af data betyder dette, at fejl i datasættet sandsynligvis vil være fundet på forhånd, hvorfor det er nemt at tage højde for. For DDA har oparbejdning den målbare effekt, at brugere af data uhyre sjældent får behov for vejledning i forbindelse med arbejde med data indhentet hos DDA; derfor er DDA's service over for brugerne helt overvejende reduceret til udlevering af datasæt.

Dokumentation af data

Dokumentationen af data, der fremkommer som et resultat af oparbejdningsprocessen, samles i „datadokumentationspublikationen“, der indeholder en studiebeskrivelse og en kodebog med frekvenser. Denne

publikation er det fysiske resultat af oparbejdningsprocessen. Studiebeskrivelsen består af en række oplysninger, der belyser undersøgelsen som helhed – dens tilblivelse og indhold. En studiebeskrivelse rummer for eksempel følgende oplysningskategorier:²

- Titel
- Nøgleord
- Dataindsamler
- Bibliografisk reference
- Kort sammenfatning af undersøgelsens emne
- Adgangskriterier

Ideelt set burde studiebeskrivelsen udarbejdes umiddelbart efter modtagelsen af et datamateriale, således at der straks efter modtagelsen kan søges på studiet i DDA's søgekatalog. Af ressourcemæssige årsager bliver studiebeskrivelsen dog ofte først udarbejdet i forbindelse med den senere oparbejdning af data.

Som nævnt er det andet element i datadokumentationspublikationen kodebogen, der er en oversættelse af datafilens tal, og den består af en variabelbeskrivelse med oplysninger om den enkelte variabls placering i datafilen, kodernes betydning, den fuldstændige spørgsmålstekst og andre væsentlige informationer – herunder frekvensfordelinger mm. En OSIRIS-kodebog er en tekstfil i kortbilledformat, hvilket vil sige, at de enkelte linjer maksimalt kan være 80 tegn. Det er dog muligt at fortsætte en tekststreng på en ny linje.

De to første variable i DDA's standard kodebog er variable indføjet af DDA, nemlig DDA's studienummer og DDA's løbenummer. Dernæst følger de originale variable fra den afleverede datafil. For alle numeriske variable indeholder kodebogen frekvensfordelinger både i antal og procent. OSIRIS-formatet er et opbevaringsformat og ikke et præsentrationsformat. Derfor præsenteres kodebogsinformationerne i trykte publikationer i et præsentrationsformat udviklet af DDA.

Trin i oparbejdningsprocessen

En forudsætning for en meningsfuld langtidsopbevaring af data er, at data og dokumentation opbevares i et format, som er tilstræk-

kelig simpelt til, at indholdet vil kunne læses og analyseres i fremtiden uanset, hvordan fremtidens computere og deres databehandlingssoftware kommer til at se ud. Enhver software- eller systemafhængighed skal derfor så vidt muligt fjernes, og datasættet skal dokumenteres så fuldstændigt, som det er muligt.

Oparbejdningsprocessen indebærer, at datamaterialet udsættes for forskellige standardiseringer og checks, der så vidt muligt bringer data og dokumentation i overensstemmelse. Blandt standardiseringerne er indføjes af DDA's studienummer og sekvensnummer og en tilpasning af materialet til nogle generelle koder for manglende data.

I spørgeskemaundersøgelser skelnes der mellem tre forskellige kategorier af manglende data:

- Irrelevant (10, 100, ... etc.) anvendes, hvor en respondent ikke skal besvare pågældende spørgsmål (filtrering).
- Uoplyst (9, 99, ... etc.) anvendes, hvor en respondent ikke har besvaret et spørgsmål, der burde være besvaret.
- Ikke afkrydset/bsvaret (11, 101, ... etc.) anvendes, hvor to eller flere undersøgelser er flettet sammen, og en respondent derfor ikke er blevet stillet alle spørgsmål

Ved at foretage en standardisering af koder for manglende data lettes analysearbejdet for kommende brugere. Fordelene ved standardiseringerne bliver naturligvis særligt iøjnefaldende ved arbejde med sammenligninger på tværs af datasæt. Materialet checkes desuden grundigt igennem, således at datadefinitionerne stemmer helt overens med de oplysninger, der gives i den originale dokumentation. Findes der fejl, manglende data eller uoverensstemmelser, dokumenteres de. Data ændres ikke, men DDA noterer i kodebogen unøjagtigheder eller uforståelige ting, således at efterfølgende brug af materialet gøres lettere. Er der tale om spørgeskemadata, gengives spørgeskemaet sidst i datadokumentationspublikationen; spørgeskemaerne indscannes og gemmes altid i pdf-format.

Mere præcist kan oparbejdningsprocessen

2. Bilag I er et eksempel på en studiebeskrivelse, der stammer fra den danske valgundersøgelse 1998.

indeledes i fem faser, hvis hovedelementer er beskrevet i det følgende. (For teknisk interesserede læsere er processen beskrevet i diagrammet i Bilag II. Heraf fremgår hvilke filtyper, der anvendes til hver enkelt kørsel. Ved sideskiftene er de enkelte trin gentaget for overskuelighedens skyld).

1. Modtagelse af data

Så snart der er afleveret et datamateriale, tildeles dette et unikt studiesagsnummer i DDA's journalsystem. Data afleveres i mange forskellige formater, derfor konverteres de modtagne data til en systemuafhængig „flad tekstfil“ samt til det format, der anvendes i oparbejdningen (p.t. SAS). Det er en god idé at deponere data ved DDA relativt kort tid efter indsamlingen, da sandsynligheden for at variable er slettet, overskrevet og rekodet endnu er lille. Da de færreste ønsker, at andre får adgang til deres data, førend deres eget analysearbejde er afsluttet og publiceret, er der typisk ikke adgang for tredjepersoner til data, førend indsamleren eksplicit giver meddelelse herom.

2. Formel kontrol

Den basale kontrol består i at efterprøve om variabel- og caseantal i datasættet faktisk er i overensstemmelse med det, der er angivet i den medsendte dokumentation, der er et 'lokaliseringskema', i hvilket personen, der aflevere data, fastsætter betingelser for andres eventuelle adgang til data. Endvidere kan den medsendte dokumentation være spørgeskema, kodebog etc. Der udskrives også frekvenstabeller på alle variable for at checke, at data er korrekt indlæst og svarer til dokumentationen. Denne kontrol er i et vist omfang sammenfaldende med det check af data, der foretages i forbindelse med arkivering ved modtagelse.

3. Omkodninger og dokumentation af data

En stor del af oparbejdningsjobbet består i omkodninger og dokumentation af data – udarbejdelse af den såkaldte kodebog. Ændringer og omkodninger foretages ved at redigere en SAS-scriptfil. I denne tilføjes DDA's eget unikke studienummer, samt sekvensnummer til materialet. Da labelnavne maksimalt må være 24 karakterer og skal være sigende, må disse ofte ændres.

Dernæst omkodes variable med henblik på at standardisere blandt andet filtre og formater (for eksempel længder og antallet af decimaler). Til sidst omnavngives variable til en fortløbende variabelrække. Samtidig med omkodningen af data, oprettes kodebogen fra bunden. Dette betyder for eksempel, at spørgeskematekst ofte indarbejdes ved, at man nøjagtigt afskriver spørgsmålstekster fra det originale skema.

4. Bearbejdning af data og kodebog

Via forskellige scripts kørt på data og kodebog, linkes data med kodebog og kodebogen formateres med henblik på skabelsen af den endelige datadokumentationspublikation. I denne proces tilføjes indledning og studiebeskrivelse til kodebogen. Slutproduktet for data er en fil ved navn „DATA.OSI“, som er en simpel talmatrice, mens „DDPLAS“ er kodebogen med printkoder for sideskift, sidenummer mm.

5. Lagring og udlevering

Det oparbejdede datamateriale lagres endelig i to lige vigtige dele – en datafil og en dokumentationsfil – ligesom de originale datafiler samt en fil indeholdende dokumentation af processen (en SAS-scriptfil) lagres. Ved senere udlevering af et datamateriale sammenføres datafil og dokumentationsfil. Dette gøres ved kørsel af programmer, som danner datafiler i det til enhver tid ønskede format. Ønsker man at „udgive“ studiet på internettet, sammenføres datafil og dokumentationsfil ved hjælp af et program, der konverterer til XML- og HTML-filer.

Som det kan udledes af ovenstående, er oparbejdning en ganske ressourcekrævende proces, som der nødvendigvis må økonomiseres med. I dag anvendes, hvad der svarer til gennemsnitligt 200 timer pr. datamateriale – dog med meget stor variation. DDA har derfor valgt at lade (den forventede) efterspørgsel efter de enkelte datamaterialer bestemme rækkefølgen, de oparbejdes i. Ses der på de enkelte trin i processen, finder man, at det at skrive selve kodebogen, omkodning af variable, omnavngivning af labels og variable, er de store tidsforbrugere, mens kørsel af de forskellige scripts tager meget lidt tid.

Fremtiden for oparbejdning

Oparbejdning er en kerneaktivitet i DDA og i høj grad den aktivitet, der gør DDA unik i forhold til alternativ data(op)bevaring i den danske forskningsverden. Samtidig er oparbejdning som nævnt også en ressourcekrævende aktivitet og den måde, der oparbejdes på i DDA, er ikke nødvendigvis den eneste rigtige. Derfor overvejes det også løbende at ændre processerne og udskifte programformaterne, som de forskellige anvendte scripts understøttes af. Eksempelvis kunne SAS-scripts erstattes med SQL-udtryk. Det afgørende for dataindsamlere, der lader data deponere ved DDA og for databrugere, er imidlertid ikke, hvilke processer data gennemgår og hvilke programmer, der anvendes for at sikre datas holdbarhed. Det afgørende er, at data bevares i et langtidshold-

bart format, og at data og dokumentation til enhver tid kan sammenkædes korrekt og læses. Dette vil der naturligvis ikke ændres ved uanset eventuelle ændringer i DDA's oparbejdningsprocedure. Selvom det ikke er strengt nødvendigt for dataindsamlere, der deponere data ved DDA, at kende til alle oparbejdningsprocessens detaljer, er det nok alligevel ganske rat at vide lidt om, hvilken behandling de deponerede data udsættes for. Artiklen har netop haft til hensigt at give en introduktion til denne behandling.

Læsere, der ønsker en mere detaljeret beskrivelse af dataarkivering og arbejde med datadokumentation – både generelt og sådan som det håndteres i DDA, kan henvises til bogen „*Datadokumentation – Metadata for samfundsvidenskabelige undersøgelser*“ af Karsten Boye Rasmussen, Odense Universitetsforlag 2000.

Bilag I. Eksempel på en studiebeskrivelse – Valgundersøgelsen 1998

Undersøgelsens nuværende og forventede status i DDA: data findes i DDA i OSIRIS format med maskinlæsbar dokumentation (klasse D).

Undersøgelsens placering i klynge(r): vælgerundersøgelser og politiske meningsmålinger.

Relevante nøgleord for undersøgelsen: danske nøgleord:
partipræference; folketingsvalg 1998; valgkamp 1998; politik, interesse for; diskussioner, deltagelse i; partipræference, alternativ; statsgælden, betydning for partivalg; arbejdsløshed, betydning for partivalg; EU, betydning for partivalg; miljøspørgsmål, betydning for partivalg; sundhedsvæsen, betydning for partivalg; skattespørgsmål, betydning for partivalg; flygtningespørgsmål, betydning for partivalg; efterlønsspørgsmål, betydning for partivalg; indvandrerspørgsmål, betydning for partivalg; ældrepolitik, betydning for partivalg; retssikkerhed, betydning for partivalg; uddannelsesspørgsmål, betydning for partivalg; lejeværdi, betydning for partivalg; børnefamiliers forhold, betydning for partivalg; meningsmålinger, betydning for partivalg; VKZ-flertal, betydning for partivalg; partitilhørsforhold; Rasmussen, Poul Nyrup; Nyrup Rasmussen, Poul; Jelved, Marianne; Møller, Per Stig; Jakobsen, Mimi; Nielsen, Holger K.; Ellemann-Jensen, Uffe; Kjærsgaard, Pia; Sjursen, Jann; Jacobsen, Kirsten; Aaen, Frank; folketingsvalg 1994; partimedlemskab; regeringens virke; regeringspræference; borgerlig versus socialdemokratisk regering; socialdemokratisk versus borgerlig regering; sociale udgifter, holdning til; levestandard; indtægtsforhold; statskontrol med erhvervslivet; erhvervslivet, statskontrol med; bankers stilling; storindustriens stilling; brugerbetaling; TV-valgkampe; „velfærdsstrategien“; „eksportstrategien“; samfundet, mål for; offentlige udgifter;

forsvaret, udgifter til; sundhedsvæsenet, udgifter til;
uddannelse, udgifter til; undervisning, udgifter til;
folkepension, udgifter til; forureningsbekæmpelse, udgifter til;
miljøproblemer, udgifter til; kulturelle formål, udgifter til;
børneinstitutioner, udgifter til; politiet, udgifter til;
understøttelse, udgifter til; arbejdsløshedsunderstøttelse,
udgifter til; dagpenge, udgifter til bistandshjælp, udgifter til;
broer, udgifter til; motorveje, udgifter til; u-landshjælp;
flygtninge, udgifter til; orlovsordninger, udgifter til;
hjemmehjælp, udgifter til; tilskudsordninger, udgifter til;
børnefamilier, udgifter til; kollektiv transport, udgifter til;
transport (kollektiv), udgifter til; bruttoskat; indkomstskat;
offentlig service; service, offentlig; skattesænkning versus
afdrag på statsgælden; afdrag på statsgælden versus
skattesænkning; statsgældafdrag versus skattesænkning;
Øresundsbroen; Storebæltsbroen; miljø versus økonomisk vækst;
økonomisk vækst versus miljøet; miljø versus erhvervslivets
konkurrenceevne; erhvervslivets konkurrenceevne versus miljøet;
konkurrenceevnen versus miljøet; NATO, holdning til;
indvandrerpolitikken, stramning af; flygtningepolitikken,
stramning af; demokratiets vilkår; EU, holdning til;
EU-folkeafstemningen 1993; folkeafstemningen 1993;
unionsafstemningen 1993; Edinburgh-folkeafstemningen 1993;
unionsafstemningen 1998; Amsterdam-folkeafstemningen 1998;
EU-folkeafstemningen 1998; folkeafstemningen 1998; politisk skala,
placering på; grøn politik; forbrydelser, straf for; strafudmåling
ved forbrydelser; kriminalitet, forebyggelse af; valg, afholdelse
af; folketingsmedlemmer, holdning til; stemmeafgivelse, betydning
af; flygtninge, holdning til; opholdstilladelse til flygtninge;
understøttelse, misbrug af; dagpenge, misbrug af;
arbejdsløshedsunderstøttelse, misbrug af; økonomisk situation,
landets og familiens; familiens økonomiske situation, udvikling i.

Sprog anvendt i denne undersøgelse: Dansk.

IDENTIFIKATIONER OG REFERENCER

Bibliografisk reference:

Datamateriale DDA-4189:

Valgundersøgelsen 1998.

Primærundersøgere: Jørgen Goul Andersen, Johannes Andersen, Ole
Borre og Hans Jørgen Nielsen.

DDA-4189, 1. udgave (ved Birgitte Grønlund Jensen, Jens Wagner og
Lene Wul).

Dansk Data Arkiv 1999.

1 datafil (2001 respondenter, 327 variable) med tilhørende
maskinlæsbar dokumentation (249 pp.).

Lokalarkiv hvor studiet forefindes: Studienummer (datasæt): 4189.

Donor (afleverende instans): Andersen, Jørgen Goul. AUC: Institut
for Økonomi, Politik og Forvaltning. Fibigerstræde 1, DK-9220
Aalborg Øst.

Afleveringstidspunkt: 1999-09-20.

Primær-undersøgelsesleder (forskningsinstitution):

- (A) Andersen, Jørgen Goul; Andersen, Johannes;
- (B) Borre, Ole;
- (C) Nielsen, Hans Jørgen.
- (A) AUC: Institut for Økonomi, Politik og Forvaltning;
- (B) AU: Institut for Statskundskab;
- (C) KU: Institut for Statskundskab.
- (A) Fibigerstræde 1, DK-9220 Aalborg Øst;
- (B) Universitetsparken, DK-8000 Århus C.;
- (C) Rosenborggade 15, DK-1130 København K.

Dataindsamler: ACNielsen AIM. Strandboulevarden 89, DK-2100 København Ø.

Økonomisk sponsor: Statens Samfundsvidenskabelige Forskningsråd.
Randersgade 60, DK-2100 København Ø.

ANALYSEBETINGELSER

Undersøgelsens emne (kort beskrivelse): Valgundersøgelsen 1998 indeholder data indsamlet af ACNielsen AIM i marts-juni 1998.

Valgundersøgelsen 1998 består i lighed med de forudgående valgundersøgelser af en hovedundersøgelse (denne undersøgelse) og en suppleringsundersøgelse (DDA-6301). Suppleringsundersøgelsen er en udbygning med ca. 2000 ekstra svarpersoner på nogle få variable, først og fremmest sociale baggrundsspørgsmål. I 1998 var det dog muligt at gå lidt videre, fordi suppleringsundersøgelsen blev gennemført som del af en anden politologisk undersøgelse (DDA-6300: Demokrati fra neden, 1998).

I undersøgelsen medtages der for første gang det fulde sæt af variable fra det internationale valgforskningssamarbejde CSES (Comparative Study of Electoral Systems), og undersøgelsens data oparbejdes endvidere som en del af dette datasæt.

I datamaterialet indgår følgende baggrundsspørgsmål: Respondentens køn, alder og civilstand; husstandens størrelse, og antal voksne (18 år eller derover) og børn i husstanden samt, hvis der er børn, disses alder; respondentens uddannelse, beskæftigelsesmæssige status samt privat eller offentlig ansættelse; arbejdsløshed inden for de sidste to år; medlemskab af en fagforening; respondentens og husstandens samlede årsindkomst; ejerforhold til boligen; samt bopælens geografiske placering og urbaniseringsgrad.

Desuden indgår spørgsmål om følgende emner:

Generel politisk holdning:

Hvilke problemer er de vigtigste, som politikerne skulle tage sig af; deltagelse i diskussioner om politik op til valget; generel interesse for politik; hvad respondenter stemte ved valget den 11. marts 1998; beslutningstidspunkt; om respondenter stemte med begejstring, med tilfredshed eller med blandede følelser; om respondenter stemte på et parti eller på en bestemt kandidat, og om det var en mand eller en kvinde; vigtigste grund til at stemme

som respondenterne gjorde; om respondenterne overvejede at stemme på et andet parti; betydning af en række emner i valgkampen for partivalg; vurdering af de enkelte partier og partilederne; om og hvad respondenterne stemte ved folketingsvalget i 1994; grad af partitilhørsforhold; medlemskab af et politisk parti; om en række problemer bedst løses af den nuværende socialdemokratisk ledede regering eller en borgerlig regering; holdning til en række politiske påstande ang. samfundets sociale udgifter, levestandard og indtægtsforhold, statens kontrol med erhvervslivet, bankernes og storindustriens stilling, brugerbetaling i den offentlige sektor samt udbytte af valgkampe i aviser, radio og TV; holdning til de to hovedstrategier („velfærdsstrategien“ og „eksportstrategien“) til at skabe flere arbejdspladser; vigtigste målsætninger for samfundet (opretholde lov og orden; give den enkelte større mulighed for at påvirke politiske beslutninger; bekæmpe prisstigninger; beskytte ytringsfriheden); holdning til det offentliges udgifter til forskellige formål; holdning til en forhøjelse af bruttoskatten samtidig med en nedsættelse af den almindelige indkomstskat; holdning til den offentlige service på en række områder (bibliotekerne; sygehusene; folkeskolen; børnehaver og vuggestuer; hjemmehjælpen); holdning til lavere skatter eller forbedring af den offentlige sektor; holdning til lavere skatter eller afdrag på statsgælden; holdning til Øresundsbroen og Storebæltsbroen; holdning til en række synspunkter fra den politiske debat; holdning til de foreslåede ændringer af flygtninge-/indvandrerpolitikken; holdning til en række spørgsmål om politik og politikere som helhed; tillid til danske politikere; tilfredshed med den måde demokratiet fungerer på; generel holdning til EU; hvad respondenterne stemte ved folkeafstemningen om Edinburgh-aftalen i 1993; stemmeintention ved den kommende folkeafstemning om Amsterdam-traktaten den 28. maj; grunde til at stemme ja eller nej; holdning til en række opfattelser af EU-samarbejdet; placering på en politisk venstre-højre skala; placering af respondenterne og partierne på en skala angående en række emner (hvor mange flygtninge vi kan tage imod; hvor stor den offentlige sektor bør være; miljøhensyn og „grøn“ politik; støtte til børnefamilierne; strengere straffe eller forebyggelse og human behandling af forbrydere; stramning af den økonomiske politik; forholdet til EU); om det sidste valg i Danmark gik rigtigt for sig; om de politiske partier tager hensyn til, hvad almindelige mennesker mener; om politiske partier er nødvendige; om folketingsmedlemmerne ved noget om, hvad almindelige mennesker mener; om det gør en forskel, hvem der har regeringsmagten; om det har nogen betydning, hvem folk stemmer på; om danskere almindeligvis siger deres politiske mening; om danskerne vælger parti ud fra politiske holdninger eller ud fra de partiledere, man bedst kan lide; om flygtninge skal sendes hjem hurtigst muligt eller skal integreres; om misbrug af dagpengesystemet er et stort problem; om respondenterne husker navnet på nogle af de kandidater, som stillede op i respondentens amt eller storkreds; landets økonomiske situation sammenlignet med for et år siden; respondentens udbytte af det økonomiske opsving; om respondenterne har haft kontakt med et medlem af Folketinget inden for de sidste 12 måneder.

Datatype: Spørgeskemaundersøgelse.

Observationsenheder: Individer.

Antal observationsenheder (respondenter): antal respondenter
(uvægtet): 2001.

Datasættets størrelse: 2001 respondenter; 327 variable pr.
respondent.

Det arkiverede studies fuldstændighed: 76
spørgeskemaspørgsmål/tests; 100% heraf medtaget i datasættet.

Periode dækket af undersøgelsen: Start på periode: 1998-03. Slut
på periode: 1998-06.

Dimensioner i tid: Tværgående (engangs studie).

Definition af det univers hvorfra stikprøven er udtaget:
aldersgrænser; boligforhold. Danske statsborgere på 18 år og
derover, bosiddende i private husstande.

Udvælgelsesprocedurer: Simpel tilfældig udvælgelse.

Undersøgelsens geografiske univers: national. DK.

Tidsrum for dataindsamling: Start: 1998-03. Slut: 1998-06.

Dataindsamlingsmetode: mundtligt interview.

Skema/test-type: struktureret spørgeskema.

Dataindsamlingspersonale: instituttets faste personale. ACNielsen
AIM.

BETINGELSER VED REANALYSE

Nuværende datarepræsentation: enkel hulning
(karakter-repræsentation).

Anvendelige analyse-pakker: SPSS. SAS.

Sprog anvendt i det skriftlige materiale: Studiebeskrivelse: Dansk
og engelsk. Skemaer: Dansk. Kodebog: Dansk.

Adgangsbetingelser: ingen restriktioner ved
videnskabelig/statistisk anvendelse.

Adgangstilladelse indhentes hos: Andersen, Jørgen Goul. AUC:
Institut for Økonomi, Politik og Forvaltning. Fibigerstræde 1,
DK-9220 Aalborg Øst.

REFERENCER TIL RELEVANTE PUBLIKATIONER/RESULTATER/STUDIER

Publikationer/rapporter fra primær-undersøgeren: (1) Andersen, Johannes; Borre, Ole; Andersen, Jørgen Goul; Nielsen, Hans Jørgen: Vælgere med omtanke - analyse af folketingsvalget 1998. Århus: Systime A/S, 304 pp. 1999.

Andre publikationer (sekundær analyse): (1) DrejØ, Kristine; Mathiasen, Kenneth Skovdall; Mogensen, Jens Hougaard; Møberg, Rasmus Juul: Økologi, en ny værdi? Aalborg: Aalborg Universitet, specialeopgave, 110 pp. + appendix; 1999; (2) Pedersen, Jonas Holm et al. Private sygeforsikringer - en sund beslutning... For hvem? Aalborg, specialeopgave, 82 pp. 2001.

Referencer til beslægtede studier:

DDA-0007: For- og eftervalgsundersøgelsen 1971.
DDA-0008: Valgundersøgelsen 1973.
DDA-0016: Valgundersøgelsen 1975.
DDA-0017: Vælgerundersøgelse januar og februar 1975.
DDA-0148: Panelundersøgelse af valgene 1971 og 1973.
DDA-0166: Valgundersøgelsen 1977.
DDA-0191: Vælgerundersøgelse februar og marts 1977.
DDA-0287: Valgundersøgelsen 1979.
DDA-0328: Vælgerundersøgelse oktober og november 1979.
DDA-0529: Valgundersøgelsen 1981.
DDA-0530: Vælgerundersøgelse december 1981 og januar 1982.
DDA-0658: Danske valgundersøgelser, gentagne spørgsmål 1971-1981.
DDA-0772: Valgundersøgelsen 1984.
DDA-1340: Valgundersøgelsen 1987.
DDA-1432: Valgundersøgelsen 1988.
DDA-1480: Vælgerundersøgelse september 1987 og maj 1988.
DDA-1564: Valgundersøgelsen 1990.
DDA-1565: Vælgerundersøgelse januar, februar og marts 1991.
DDA-2210: Valgundersøgelsen 1994.
DDA-3132: Vælgerundersøgelse januar og februar 1995.

MEDTAGNE BAGGRUNDSVARIABLE

Centrale personoplysninger: alder, kategoriseret; køn; civilstand.

Bopæl: områdets/statens navn; urbaniseringsgrad.

Boligforhold: ejendomsforhold til boligen.

Husstandens karakteristika: børn tilstede; husstandens totale størrelse (antal personer); antal unge (ca. 13-19 år); antal yngre børn (ca. 0-5 år); antal ældre børn (ca. 6-12 år); børnenes alder. Antal personer i husstanden på 18 år eller derover. Antal hjemmeboende børn og disses alder angivet som: 0-5 år, 6-17 år, 18 år og derover.

Beskæftigelse: respondent: stillingsbetegnelse (eller: kode); respondent: økonomisk sektor; respondent: beskæftigelsesmæssig status; respondent: arbejdsløshedens varighed (hvis arbejdsløs).

Indkomst: respondent: bruttoindkomst; hele husstanden: bruttoindkomst.

Uddannelse: respondenten: højest opnåede eksamen. Samt uddannelse udover skoleuddannelsen.

Politik: respondenten: udnyttelse af valgretten, valgdeltagelse; respondenten: politisk information; respondenten: politiske interesse; respondenten: identifikation med/præferencer for parti; respondenten: nuværende medlemsskab af partiet; respondenten: tidligere medlemsskab af partiet.

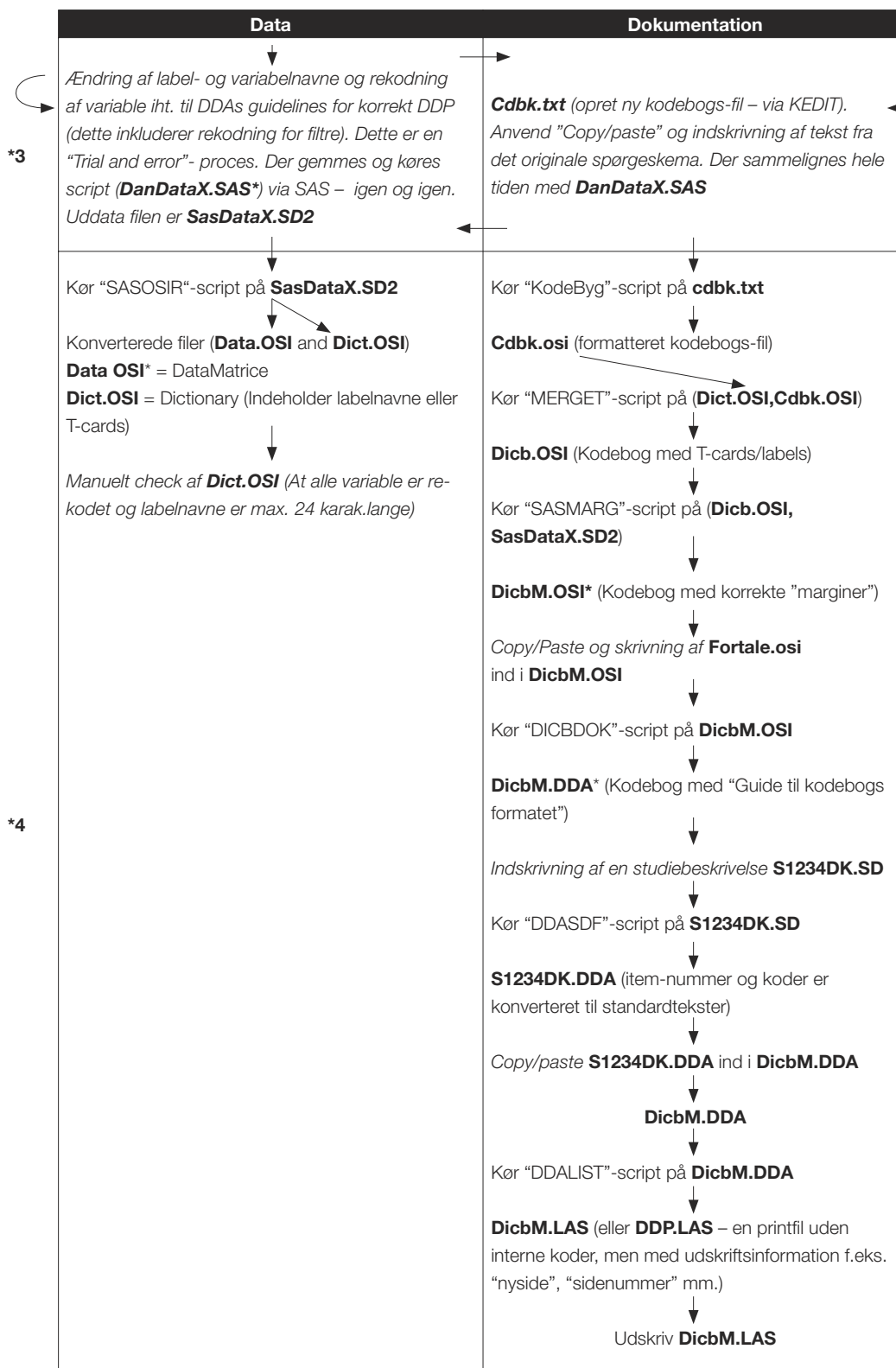
Læsning, massemedier og 'kulturel' påvirkning: aviser, dagblade; ugeaviser, blade ('seriøse'); blade (dameblade, kulørte hæfter, etc.).

Medlemsskab af foreninger: fagforeninger (nuværende, tidligere, varighed).

Andre baggrundsvARIABLE: Arbejdsløshed: respondenten og dennes ægtefælle/samlever.

Bilag II.
Diagrammatisk
beskrivelse af
oparbejdningsprocessen i DDA

	Data	Dokumentation
	Data0.ORG* (originale data fra donor, SPSS, SAS eller Statistical-files)* ↓ <i>Formel kontrol (checker om filerne er maskinlæsbare og at antallet af variable og respondenter er ok)</i> ↓	
1	Data1.ORG (checkede data i *.por or *.sav format) ↓	
*2	DanData.SAS (opret ny SAS-scriptfil ved brug af editoren KEDIT) ↓ Brug DanData.SAS til at læse Data1.ORG ind i SAS = Dan frekvenstabeller (SasData.SD2) ↓ <i>Udskriv (SasData.SD2) eller frekvenstabellerne til papir (på dette tidspunkt kan oparbejderen estimere hvor mange ressourcer der skal anvendes til studiet. F.eks. ved at checke hvor veldokumenteret studiet er).</i> ↓	



Forklaringer til diagrammet:

- Kursiverede tekster = manuelle sider af oparbejdningen.
- Alle filnavne er angivet med fed skrift, f.eks. **CreateData.SAS**.
- Alle anvendte scripts og programmer er angivet i anførselstegn som f.eks. „SASOSIR“.
- Diagrammets 5 punkter, svarer til de 5 trin i oparbejdningsprocessen.

Efter processens gennemførelse arkiveres følgende filer:

1. Data.OSI
2. Doku.OSI (=DicbM.DDA)
3. DanData.SAS (Består af alle SAS-programlinier = dokumentation af hele oparbejdningsprocessen)
4. Data1.org/Data0.org

Folketællingen 1801 en guldgrube

Af CAND. POLIT. OLE S.D. HANSEN

I METODE & DATA nr. 84 2001:1 præsenterede Dansk Data Arkiv (DDA) CD udgivelsen af folketællingen fra 1801. Ole Degn fra Landsarkivet for Nørrejylland gennemgik folketællingens arkivhistorie (proveniens, nuværende tilstand), Nanna Floor Clausen og Hans Jørgen Marker begge fra DDA beskrev KIP projektet samt procedurer og problemer omkring datarensning og kodning, og Hans Chr. Johansen fra Demografisk Forskningscenter, Syddansk Universitet Odense, belyste ud fra foreløbige tal nogle muligheder for tabuleringer og analyse i en delmængde af materialet. Nu har CD'en været på markedet i over to år, og jeg har ikke set tilløb til seriøse forskningsrapporter på baggrund af dette materiale, som må anses for at være lidt af en guldgrube og gave til forskning i fx historie, demografi, økonomi samt navneforskning. Det er derfor på høje tid, at der gøres en indsats for at præsentere materialet og nogle af de foreliggende muligheder i dette.¹

Jeg har bemærket, at den norske del af materialet er stillet til rådighed for offentligheden via Internettet. Det drejer sig i kort form om beskrivelse af alder, husstands- og erhvervsstruktur i Norge i 1801 totalt. Der er selvfølgelig i det norske materiale nogle f lakuner, som der også er i den danske del af tællingen, men de er så små, at de ikke anfægter optællingernes repræsentativitet. Hertil kommer en række fejlregistreringer, hvoraf nogle er nævnt af Hans Jørgen Marker, men der er flere, som dog alle er af mindre betydning. Det er klart, at ønsker forskere eller studerende at anvende materialet fx til større opgaver, så skal disse fejl og mangler kontrolleres. Historikere må i arkiverne for at forklare nogle af disse fejl og variationer. Men det er en ganske unik situation, at hele folketællingen findes på en CD, som er et moderne og effektivt index over, hvor pro-

blemerne skal søges afklaret, for at materialet kan anvendes.

Hans Chr. Johansen demonstrerer flere mulige analyser, som altså nu kan gennemføres på det totale materiale. I tilsvarende materialer samt i moderne internationale statistikker fra ulande finder man ligeledes det fænomen, at der er en ophobning på opgivelsen af alder omkring de runde aldre 40, 50, 60 osv. Problemet kan løses på forskellige måder, men det invaliderer ikke optalte aldersfordelinger og sammenligninger mellem landets områder og fx Norge eller Island. Et mere kildent problem er undersøgelsen af husstandenes sammensætning. Man burde kunne eftervise eller afvise påstanden om, at man i tidligere tider, her altså den 1. februar 1801, typisk eller bare ofte havde husstande, eller skal vi sige familier med 3 generationer eller mere. Materialet skal bearbejdes ret kraftigt for at muliggøre denne analyse. I Norge er man kommet frem til, at der i alt er lidt under 7,000 husstande efter en gennemarbejdet definition af, hvad en husstand er, og tallet for Danmark vil jeg antage bliver omkring 10,000. I Norge kan de ikke rigtig finde disse 3. generationsfamilier. Der er nogle på landet, men færre i byerne. Jeg tror, man vil komme til den samme konklusion i den danske 1801 folketælling.

Den typiske 3. generationsfamilie er efter min mening en drøm og et fatamorgana. Det burde lede frem til et tværfagligt projekt, hvor historikere kunne eftervise detaljer via arkiver for udvalgte typer. Matematikere eller statistikere kunne opstille og teste teoretiske modeller for sandsynligheden for, at en 3. generationsfamilie overhovedet kan opstå, når man tager den daværende dødelighed i betragtning.

Endelig kunne man gøre sig håb om at finde frekvensen af stedmødre og fædre.

1. Det her omtalte materiale er 1801 folketællingen, men betragtningerne er lige så relevante i forhold til den nyligt færdigindtastede 1845 folketælling. Faktisk indeholder dataene for 1845 folketællingen endnu flere analysemuligheder end folketællingsdataene fra 1801.

Materialet (og tidens litteratur i øvrigt) peger på, at hvis man i hvert fald som enke-mand sad med små moderløse børn, så var tilskyndelsen stor til at få en ny kone. Som en indskudt historie fra min egen familie er det nedskrevet, at en af enkemændene med 4 små børn var så heldig, at han mødte en ung pige med et uægte barn. Pigen sad og græd over tingenes tilstand, men det løste sig, da de giftede sig og fortsatte med reproduktionen, så de kom op på de sædvanlige 8 børn totalt. Dette mønster er måske i virkeligheden ikke så forskelligt fra nutidens, når man tænker på de mange skilsmisser, og hvad de fører med sig af papfædre og mødre. Enkerne med faderløse eller uforsørgede børn kom til at sidde alene tilbage. Det ses tydeligt i Norge og måske også i Danmark 1801 – måske endda også i dagens Danmark.

En anden mulig undersøgelse, som den totale folketælling på CD kan give anledning til, er mere dybtgående analyser, hvor man anvender data til at udvælge stratificerede stikprøver repræsentativt. Herefter kan man gå videre og søge flere detaljer i arkiverne. Man kan fx vælge 200-300 husstande eller personer og undersøge forholdene i dybden i allehånde arkiver, og resultatet vil være repræsentativt ganske vist med variation. Man har ikke tidligere haft mulighed for at udvælge de mulige emner ud fra de faktiske totaler i hele populationen.

Et fristende emne for en analyse er spørgsmålet om hvor mange personer, der fx i København egentlig havde hof, militær og flåde som eksistensgrundlag. Det er ikke nok at tælle antallet af grenaderer og matroser, der findes i erhvervsangivelserne. Nej, man skal have hele den „militære husstand“ med i en sammenligning i forhold til byens øvrige husstande. Nogle ganske overfladiske optællinger i materialet tyder på, at man vil opnå en bedre forståelse af begrebet „Kongens København“. København levede simpelthen af „Kongen“.

For at forstå hvad materialet indeholder, vil jeg også efterlyse en egentlig historisk behandling af situationen omkring 1801 folketællingen. De norske historikere, Einar Lie og Kjartan Soltvedt, har i nogle artikler

beskrevet den politiske baggrund og de tanker, man gjorde sig i den enevældige centraladministration. Man burde egentlig også undersøge, hvorfor man lagde vægt på speciel registrering af personer med handicap og personer, der ikke kunne forsørge sig selv. Der er sikkert sammenhæng med „Den københavnske fattigplan af 1799“, som pålagde magistraten at forsørge, om end mådeligt, de borgere, der p. gr.a. handicap eller arbejdsløshed ikke var i stand hertil. Denne „sociallov“ blev udstrakt til hele landet med „Fattigordningerne 1802-03“. Var man interesseret i størrelsesorden af den forsørgerbyrde, man hermed pålagde byernes magistrater og landområdernes sogne? Dette kan have været en medvirkende årsag til indførelsen af sogneråd i starten af 1800-tallet. Nogen skulle forvalte forsørgelsen. Det kan vel heller ikke udelukkes, at den franske revolution i 1789 var en af de trusler, som Søren Mørch mener, driver udviklingen i bl.a. socialpolitikken. En historisk undersøgelse kan således placere folketællingen 1801 i et større billede af datidens Danmark-Norge.

Afslutningsvist må materialet siges at være en sand guldgrube for navneforskere. Jeg var med, da man med stort besvær for første gang kunne tælle alle navne i CPR efter 1968. Nu udgives årets hitliste for henholdsvis pige- og drengenavne som en ren rutine. Nu kan navneforskerne med megen lille indsats tælle og fordele for- som efternavne på den eksisterende administrative inddeling fra sogn til amt. Og med de seneste tal fra KIP projektet om kommende folketællinger i samme fine kvalitet for 1787 og 1845 vil fornavnenes vandring gennem det ganske land kunne analyseres. Måske var det vrd at tage et lille kig på Norge i sidespejlet i en sådan forskning.

DDA stiller selvfølgelig folketællings-dataene til rådighed, og derfor er alle muligheder åbne for at påbegynde arbejdet med de her omtalte mulige undersøgelser. Både studerende og forskere, der ønsker at arbejde med materialerne, bør i højere grad præsenteres for de foreliggende muligheder, og her har alle interesserede en personlig opgave.

Enkestanden i Danmark 1801

En analyse af en specifik social gruppes leveforhold,
baseret på den komplette 1801 folketælling¹

Af Nanna Floor Clausen, DDA

Baggrund

Danmark råder nu over den komplette indtastning af 1801 folketællingen. I Metode & Data, nr. 84, 2001:1 er der blevet redegjort for arbejdet med at nå dette resultat. Afslutningen af indtastningen har ikke betydet, at arbejdet med at forbedre indtastningerne er ophørt. Arbejdet med korrekturlæsning er fortsat, og i særdeleshed skal nævnes arbejdet med at forbedre indtastningen af København. Disse data var af en særlig karakter, da indtastningen af København blev foretaget, før der var udviklet et standardprogram til indtastninger af folketællinger. Væsentlig information var udeladt, og felter var slået sammen. Indholdet fra de to felter 'Stilling i husstanden' og 'Erhverv' var blevet skrevet i det samme felt, og ydermere var der af pladshensyn indført mange selvopfundne forkortelser. Desuden var det ikke muligt at identificere de enkelte husstande.

Med udgangspunkt i den komplette folketælling var det interessant at undersøge forholdene omkring en udvalgt gruppe. Alle de personer, der i feltet 'Civilstand' er kategoriseret som 'Enke/Enkemand', er blevet udtrukket. Enkestanden er blevet valgt, fordi det forekom interessant at se, hvordan den klarede sig uden de sociale sikkerhedsnet, der findes i dag. Det var min opfattelse før undersøgelsen, at dette var en sårbar gruppe, der ofte havde meget begrænsede midler til rådighed. Et mål for undersøgelsen var at se, hvor mange fra denne gruppe, der boede alene og hvor mange, der boede i en familie eller husstand med mere end én person. Et andet mål var at undersøge, om der var forskelle mellem de to køns leveforhold, og om der var regionale forskelle. Efter

færdiggørelsen af denne folketælling er der udført et stort arbejde med at standardisere erhvervsfeltet. Dette arbejde har blandt andet baseret sig på det standardiseringsarbejde, der er gjort i HISCO². Men da dette arbejde ikke er detaljeret nok til et landbrugssamfund som det danske på denne tid, blev der tillige udarbejdet en særskilt erhvervskodning. Mange af de erhverv, der optræder hos denne gruppe, findes ikke i HISCO, da HISCO bygger på aktive erhverv indenfor landbrug og især industri. De erhvervskoder, der er blevet udarbejdet af Hans Jørgen Marker, DDA, er blevet anvendt i denne undersøgelse.

Undersøgelsens emner er:

- Aldersfordeling
- Husstandsstørrelse
- Erhverv
- Hvor det har været muligt, er der blevet foretaget mere detaljerede analyser, der inddrager køn og bopælsområde

Metode

I 2001 blev Folketællingen 1801 færdig-indtastet, og det blev muligt at søge i den. Først da alle data var indført i en database, begyndte arbejdet med at se nærmere på selve datakvaliteten. Der var stadig mange felter, der ikke var udfyldt, for eksempel, 'Køn', og oplysningen om alder er skrevet på mange meget forskellige måder. Da disse to felter var blevet standardiseret og kodet, blev det muligt at lave ny statistik for hele befolkningen. Efterfølgende er der som nævnt blevet foretaget yderligere korrekturlæsning, og især er København med omkring 1 mio. poster blevet rettet og forbedret.

En indtastning afleveres fra indtasterne for

1. Artiklen er en dansk bearbejdet version af et paper fremlagt ved Association of History and Computing's conference i Tromsø i august 2003.
2. HISCO – historical international classification of occupations – er et internationalt samarbejde om kodning af historiske erhverv.

hvert enkelt sogn som en 'Paradox-fil'. Disse filer lægges derefter ind i MSSQL-database. Denne database opdateres regelmæssigt, når der kommer nye data ind, og når eksisterende data er blevet korrekturlæst. Der findes to udgaver af denne database: en til forskningsbrug mm. og en udgave, der anvendes til søgninger via Internettet. Folketællingen 1801 er derudover blevet gjort til en selvstændig database, her benævnt '1. Version', der er blevet anvendt, når data skulle standardiseres og kodes. Da denne database ikke er blevet opdateret med de mindre sogne, men dog med København, blev der til brug for denne undersøgelse lavet endnu en database med Folketællingen 1801, her benævnt '2. Version'. Dette betød, at det arbejde, der var udført med standardiseringer og kodninger, kunne anvendes på noget af materialet, men ikke på det, der siden dette arbejde var blevet korrekturlæst.

Den nye 1801-tabel har 940.341 poster. Fra denne tabel blev alle med Civilstand 'enke/enkemand' udtrukket i en ny tabel. Resultatet blev en tabel med 49.447 poster.

Feltet 'Køn'

Det første, der skulle gøres, var at tilføje en kode for køn. I 25.246 tilfælde var dette felt ikke udfyldt. For at kunne udfylde dette felt blev den information, der var i felterne 'Navn', 'Stilling i husstanden' og undertiden også i feltet 'Erhverv', udnyttet. For ca. 10.000 af mændene kunne det tidligere arbejde udnyttes ved at linke til den anden tabel vha. felterne indtastningsnr og løbenr. Slutresultatet var, at kun for 10 poster var det ikke muligt at afgøre køn. Feltet blev derefter kodet til en numerisk værdi.

Feltet 'Alder'

Det næste vigtige felt var feltet 'Alder'. Ved at benytte det arbejde, der allerede var udført, lykkedes det at opdatere de fleste felter. I indtastningerne er aldersfeltet et tekstfelt, hvor alder kan være skrevet på mange måder, jævnfør kilderne, og desuden indeholde '?', hvor teksten ikke kan læses. Heltalsværdien blev udtrukket fra feltet til et nyt felt, således at det er muligt at lave statistiske analyser.

Sammenkædning af poster til en husstand

Det vanskeligste arbejde var at tildele et

husstandsnummer til en husstand. Først blev der tildelt unikke husstandsnumre til hver person i tabellen med enker. Arbejdet med tildeling af husstandsnummer var begyndt før denne undersøgelse, men er ikke afsluttet. Under arbejdet med indskrivning af folketællingerne er indtasterne blevet bedt om at udfylde feltet 'Husstands/familienr'. Mange gør dette korrekt, og det udnyttes på Internettet til funktionen 'Vis husstand'. Men det bliver samtidig gjort forskelligt i de forskellige indtastninger. For at kunne lave generelle analyser er det nødvendigt at have en standard for familienummer, der er unik for hver husstand i hele folketællingen. Det arbejde, indtasterne har gjort med at tilføje et husstandsnummer, kan udnyttes, men desværre ikke altid. Undertiden er der indenfor hver husstand tilføjet et løbenr til hver person i husstanden. Det medfører, at hver person mister identifikationen med den øvrige husstand. I disse tilfælde er det nødvendigt at se nærmere på data for at afgøre, hvornår en ny husstand begynder. Dette afgøres oftest ved at se på indholdet i feltet 'Stilling i husstanden', da husstandsoverhovedet altid nævnes først. Desværre findes der ikke en betegnelse 'Husstandsoverhoved', der kan anvendes til dette. Der blev skrevet et program, der kunne søge efter nogle af de mange måder, denne stilling kan angives på. Det har som sagt været meget tidskrævende, og i mange situationer kan det være vanskeligt at afgøre om noget er en eller flere husstande. For eksempel: tilhører en husmoders søster den samme husstand som husmoderen eller har hun sin egen husstand, når hun er angivet som 'indsidder'? Især tilfældene med indsiddere er vanskelige at afgøre og må undersøges individuelt. Men på grund af datamængden er dette kun sjældent overkommeligt. Det, der programmæssigt knytter en husstand sammen, er felterne: stedsnavn, husstands/familienummer og matrikel i tilknytning til feltet 'Stilling i husstanden'.

Ved at sammenkæde til '1. Version' lykkedes det at få tildelt husstandsnummer til 37.000 poster i enketabellen. Fra denne tabel lykkedes det i at tildele husstandsnummer til 195.000 poster i '2.version-tabellen'. De resterende 11.000 poster blev fundet og ved hjælp af et program, blev disse personer tildelt et unikt husstandsnummer, der ikke var brugt i nogen tabeller. Da dette var gjort, fik

de husstande, der kunne sammenkædes med enketabellen, tildelt et husstandsnummer. Det drejede sig om 212.000 poster. De resterende 300 poster blev fundet manuelt ved at sammenholde antallet af poster i hver indtastning. Det viste sig så, at nogle poster var duplikeret. Disse poster blev fjernet, og processen gentaget endnu mere omhyggeligt.

En ny tabel blev oprettet ved at udvælge alle de poster fra '2. Version', der havde fået tildelt et husstandsnummer. Ved at sammenholde antallet af husstandsnumre i denne tabel med enketabellen blev flere fejl, som for store husstande og for få numre, opdaget. Disse fejl blev rettet manuelt. Desuden havde sammenkædningen mellem tabellerne medført nogle fejl, der skulle rettes – primært manuelt/enkeltvist.

Ved at sammenholde med '1. Version' viste en nærmere undersøgelse, at i nogle byer var der foretaget en yderligere opsplitning i husstande, end det umiddelbart fremgik. Disse nye og mindre husstande blev derpå

tildelt '2.versionen'. I denne version blev alle enker/enkemænd tildelt en kode for at adskille dem fra de øvrige i husstanden og for at kunne opdage evt. andre fejl ved sammenkædningen. Antallet af husstandsnumre og af enker var ikke det samme, hvad det burde have været. Antallet af husstande, der havde mere end en enke, blev derefter undersøgt. Der viste sig at være 3251 af denne type husstande.

Resultat af ovenstående arbejde

44.154 husstande der indeholder 223.547

poster.

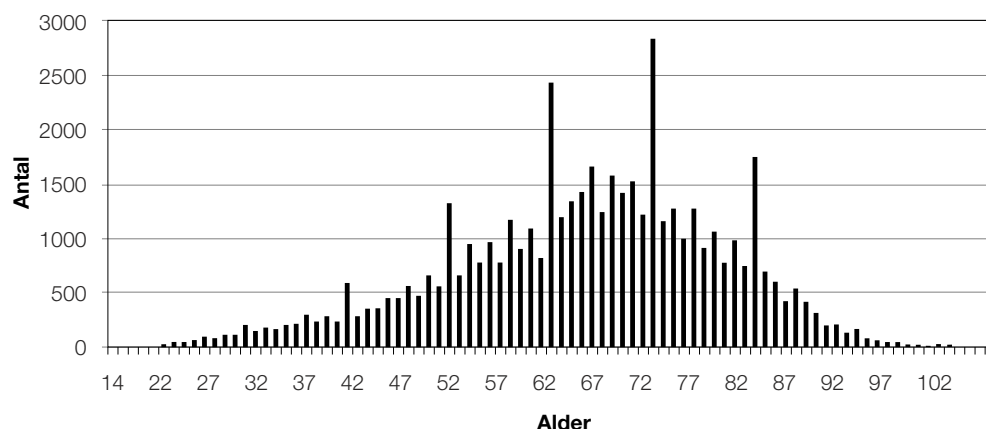
Enker: 36.942 = 75%

Enkemænd: 12.194 = 25 pct.

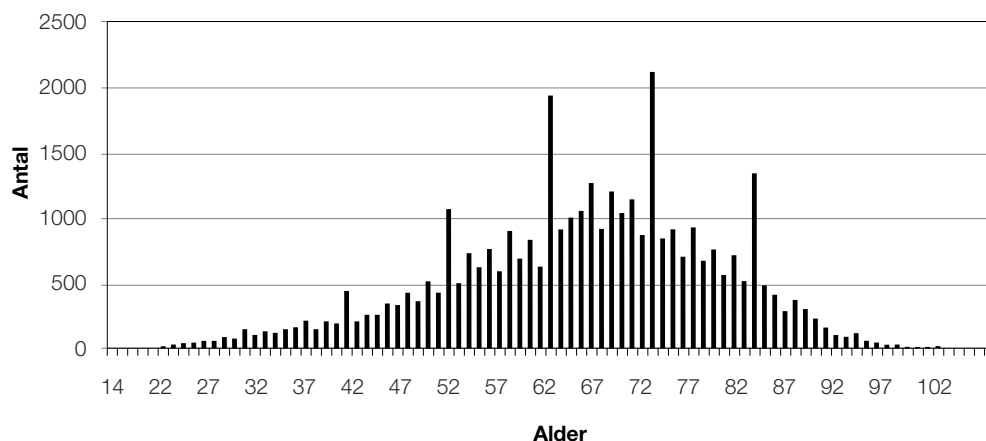
Analyse af alder

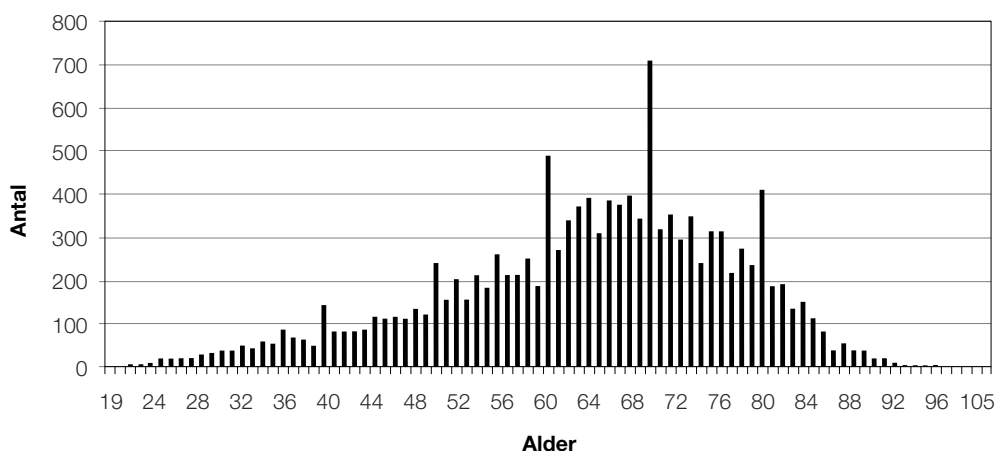
En nærmere analyse af aldersfordelingen afslørede, at der var 13 poster, der havde en alder mellem 1 og 13 år. Disse poster blev nærmere undersøgt og sammenlignet med

Figur 1.
Aldersfordeling



Figur 2.
Aldersfordeling
Kvinder





Figur 3.
Aldersfordeling Mænd

de oprindelige afleverede indtastninger, der havde de samme oplysninger. Enkelte havde dog kommenteret den påfaldende lave alder. I nogle tilfælde viste undersøgelsen, at det var i feltet 'civilstand', at der var skrevet forkert. Disse poster blev fjernet.

Det fremgår tydeligt af Figur 1, 2 og 3, at der er en overrepræsentation af de 'runde' aldre. Den basale statistik i Tabel 1 er fremkommet ved at beregne det gennemsnitlige antal personer per år i tiårsintervallerne omkring de runde år fra 40 til 80. Det vil sige, det omhandler det gennemsnitlige antal personer for hvert år i gruppen af 36-45 årige, 46-55 årige og 56-65 årige osv. Gennemsnittallet er derpå sammenlignet med det faktiske antal personer, der har fået tildelt det runde år. Ses der på 'begge køn', er der for eksempel i alt 3446 personer i intervallet 36-45 år, 583 af disse er angivet til 40 år. Det vil altså sige, at de 40 årige udgør 169 pct. af gennemsnittet på 345 personer. For hvert tiårsinterval angives således det runde års procent af gennemsnittet. I de sidste kolonner er opgørelsen opdelt på mænd og kvinder.

En sammenligning af denne alders-

fordeling med de tal, der er fundet af Hans Christian Johansen³ viser, at til trods for, at han også har fundet en overrepræsentation af aldre omkring 5- og 10-år, er der en hel anderledes overrepræsentation for enkestændens vedkommende. Der er tale om meget lidt nøjagtighed og/eller interesse for denne gruppe. Hans Christian Johansen har vist, at overrepræsentationen omkring 'runde' aldre stiger med højere alder, og det bekræftes af denne undersøgelse. Det kan desuden vises, at tendensen er endnu stærkere udtalt for kvinder end for mænd.

Alle personer, der havde en alder over 100 år, er blevet undersøgt nøjere. Mange viste sig at være fejlskrivninger og fejllæsninger, men der er dog folk med en høj alder, hvilket fremgår af husstanden på næste side (boksen er kopi fra Dansk Demografisk Database – <http://ddd.dda.dk>)

En analyse af aldersfordelingen viser, at 65 pct. af kvinderne er fra 60 år og 69 pct. af mændene er fra 60 år og frem. Dette understreger selvfølgelig, at risikoen for at blive enke/enkemand stiger med alderen, men afspejler desuden det faktum, at når en ægte-

Begge køn		Mænd		Kvinder	
Alder	Pct. af gennemsnittet	Alder	Pct. af gennemsnittet	Alder	Pct. af gennemsnittet
40	169%	40	162%	40	172%
50	173%	50	147%	50	180%
60	184%	60	159%	60	192%
70	200%	70	189%	70	205%
80	220%	80	186%	80	233%

Tabel 1.
Overrepræsentation af personer med runde år i forhold gennemsnittet i tiårsintervaller

3. I blandt andet Metode & Data, nr. 84, p. 17.

Samtlige personer i husstanden

Holbæk, Ods, Faarevejle, Faareveile Bye, , , 20, FT-1801

Der vises flg. felter:

Navn, Alder, Civilstand, Stilling i husstanden, Erhverv , Fødested
Niels Nielsen, 45, Gift, Mand, Huusmand med jord,
Marie Pedersdatter, 56, Gift, Hans kone,
Bodil Nielsdatter, 8, Ugift, Deres barn,
Ewen Larsen, 105, Enkemand, Gl mand, Lever af almisse af sognet,

fælle til en yngre person døde, giftede denne person sig hurtigt igen. Især hvis de havde en gård eller lignende, der skulle passes.

Analyse af husstandsfordelingen

Formålet med denne analyse var at undersøge hvilken type husstand, enkestanden levede i. Boede de alene eller sammen med andre? Da vi endnu ikke er færdige med arbejdet med Folketællingen 1801, er det ikke muligt at sammenligne husstandsfordelingen for enkestanden med husstandsfordelingen for befolkningen under et. I denne artikel foretages alene undersøgelser af enkestandens husstandsfordeling i hele landet og fordelt på de tre regioner: København, købstæderne og landdistrikterne.

Når man ser på Figur 4, kan det ses, at den mest udbredte husstandsstørrelse bestod af 3 eller 4 personer: 5.920 husstande bestod af tre personer, 5.808 bestod af 4 personer. 51 pct. af husstandene omfattede 2-5 personer. For hele landet gælder, at 4.589 husstande er husstande, hvor en enke/enkemand levede alene. Det svarer til 11 pct. af husstandene i denne undersøgelse.

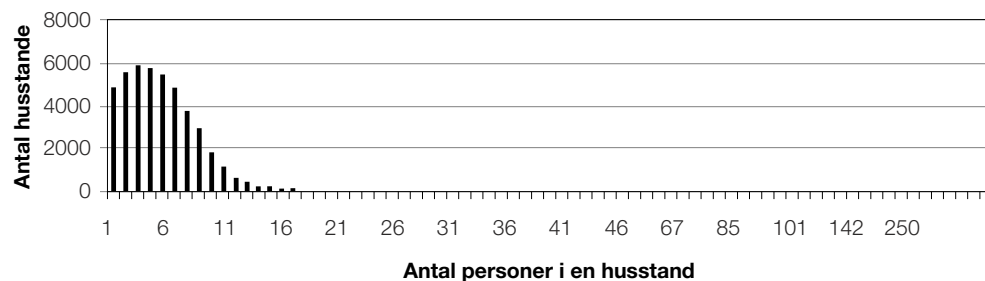
Et stort spørgsmål indenfor dette emne er fortolkningen af husstandsbegrebet i institu-

tioner. Nogle husstande omfatter mere end 200 personer, men en nærmere granskning af disse husstande viser, at de er hospitaler. Ud fra folketællingens oplysninger kan det ikke afgøres, om en person i et hospital var del af en husstand et andet sted udenfor hospitalet, eller om vedkommende boede fast på hospitalet og på den måde blev forsørgt. Hvis det sidste er tilfældet: er de så alle del af den samme husholdning, eller er hver person at betragte som sin egen husstand? I denne undersøgelse bliver alle fra det samme hospital betragtet som en husstand, da de alle bor og spiser det samme sted.

Udover at se på husstandsfordelingen indenfor enkestanden generelt er der lavet analyser for de tre tidligere omtalte geografiske opdelinger. Dette har kunnet lade sig gøre, fordi alle koderne for hvert sogn er tilføjet en kode for en af de tre regioner. Ved at sammenkæde til denne tabel har det været muligt at lave disse analyser.

Antallet af husstande i København med en enke/enkemand var 6.131. Resultatet for København viser tydeligt (jævnfør Figur 5), at den mest udbredte husstandsstørrelse var på en person. 1561 husstande svarende til 25 pct. af husstandene var kun på en person. 909 husstande svarende til 15 pct. bestod af

Figur 4.
Husstandsfordeling
hele landet



to personer: en enke/enkemand og endnu en person. Men det er også i København, at man kan finde de største husstande. Som nævnt ovenfor er disse alle hospitaler (eller tugthuse) med 200-300 personer.

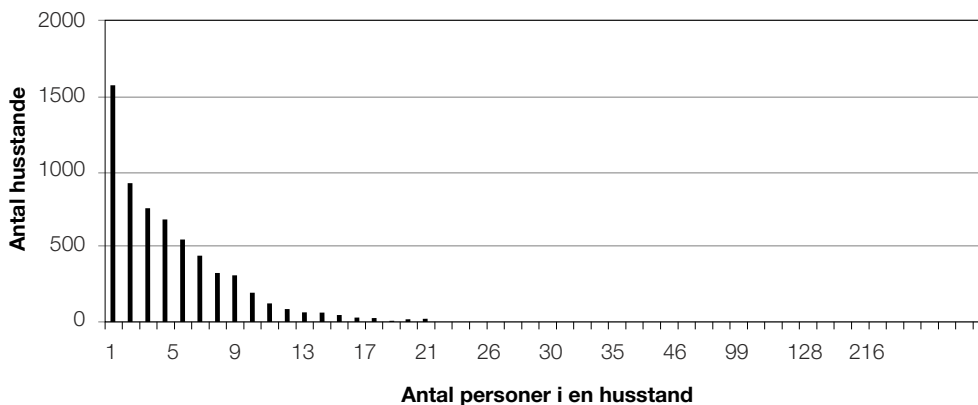
Antallet af husstande i denne 'region' var 5.014. I købstæderne var den mest udbredte husstandsstørrelse 3 personer (17 pct.), men næsten lige så mange var på to personer, nemlig 16 pct. I købstæderne var der også store husstande på omkring 100 personer, men også her drejer det sig om hospitaler. Antallet af enker, der boede alene, var i købstæderne 464 svarende til 9 pct.

Antallet af husstande i landdistrikterne er langt det største: 33.007 (jævnfør Figur 7), da langt den største del af befolkningen levede i landdistrikterne. Generelt er husstandene her større, og der er en mere jævn fordeling. 4.420 af husstandene omfattede 4 personer svarende til 13 pct. og 4.387 husstande havde 5 personer – også 13 pct. Der var tillige et stort antal husstande på 10-33 personer; det skyldes de mange godser og store landejendomme. I landdistrikterne boede 2.834 enker/enkemænd alene. Det svarer til 9 pct. ligesom i købstæderne.

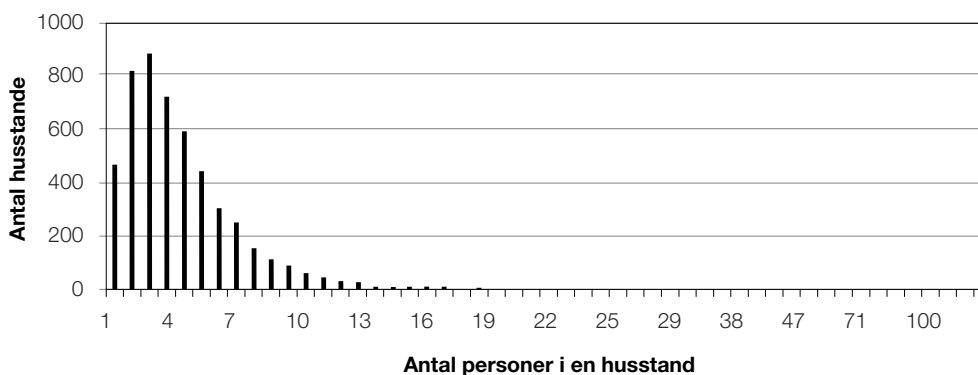
Kort om Fattigdomslovgivningen

For at forstå vilkårene for enkestanden vil der her blive givet en kort orientering om lovgivningen vedrørende fattige og fattighjælp. I 1708 kom der en lov om fattighjælp. Denne lov blev revideret i 1730 og var i kraft, indtil en ny revideret lov blev vedtaget i 1802/03.

I loven fra 1730 blev der skelnet mellem værdigt og uværdigt trængende. De værdigt trængende var folk, der ikke var i stand til at arbejde på grund af alder eller invaliditet. Hvert sogn skulle tage sig af sine egne fattige, og tiggeri fra folk, der ikke hørte til sognet, var forbudt. Der var indført en særlig skat, der skulle betale for hjælpen til de fattige. Understøttelsen kunne blive givet som mad, idet de fattige cirkulerede mellem gårdene i sognet ifølge en tilrettelagt plan. Den lokale præst skulle overvåge, at skatten til de fattige blev betalt, og at der blev sørget for de fattige. Dette var en vanskelig opgave, da der ofte var stor modstand mod, at der skulle lovgives særligt for de fattige. Man foretrak at hjælpe på frivillig basis – som velgørenhed. Tiggeri



Figur 5.
Husstandsfordeling
– København



Figur 6.
Husstandsfordeling
– Købstæder

Figur 7.
Husstandsfordeling
– landdistrikter



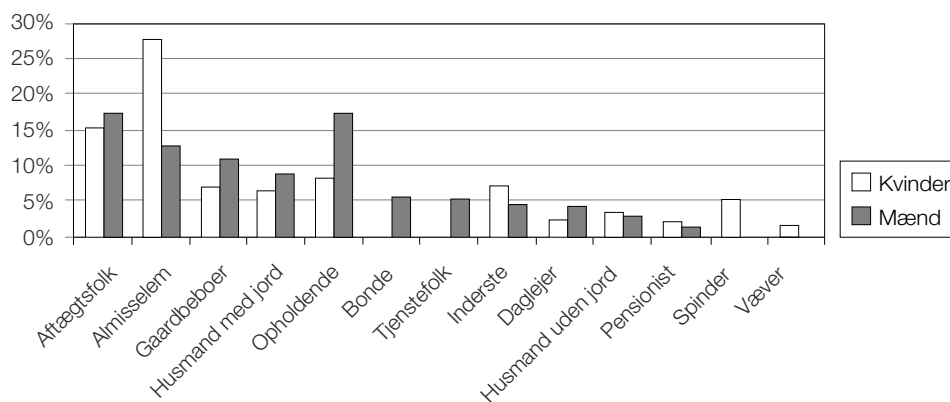
fra uværdigt trængende ophørte dog ikke med vedtagelsen af lovene. Gentagne gange blev tiggeri forbudt, uden at det hjalp. Værdigt trængende kunne derimod få tilladelse til at tigge, og i undersøgelsen af erhvervsfordelingen er der 16 personer, der er opført som værende tiggere. Nogle af de gamle og svagelige personer blev anbragt på hospitaler og de, der var i stand til at arbejde men ikke gjorde det (sandsynligvis på grund af arbejdsløshed), blev anbragt i 'tugt-, rasp- og forbedringshuse'. Disse institutioner fandtes dog ikke i landdistrikterne og der var også i

byerne for få af dem til at løse problemerne med arbejdsløshed.

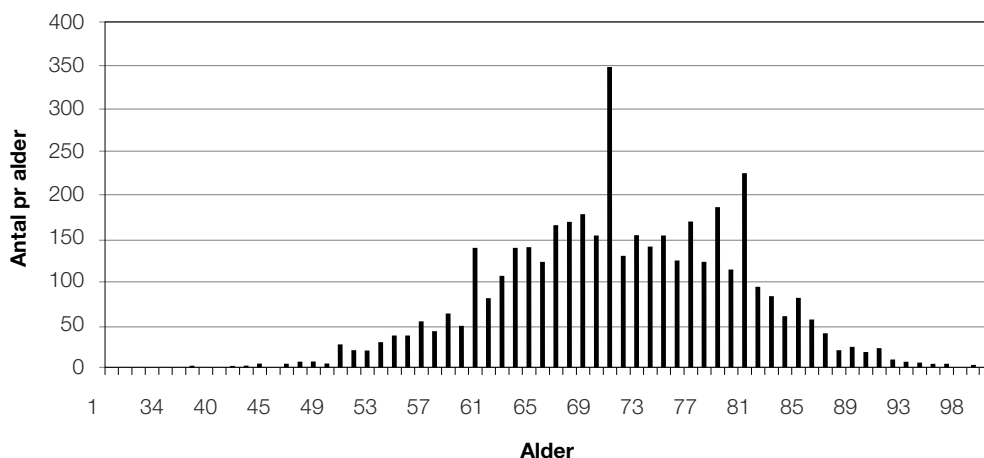
Gårdmænd havde ofte bedre muligheder for at sikre sig i alderdommen. De kunne lave aftægtsaftaler med de nye ejere af gården og få aftægt på deres egen tidligere gård. De nye ejere var ofte enten et af deres egne børn, men det kunne også være helt nye ejere, der var villige til at indgå en aftægtsaftale med de tidligere ejere. Disse aftægtsaftaler specificerer til mindste detalje hvor meget og hvad, der skulle ydes i aftægt.

Tabel 2.
De 20 mest
almindelige erhverv
for enker

Antal	Kode	Erhverv	Andel
8288	37200	Almisselem	25,2%
4163	37900	Aftægtsfolk	12,5%
2209	38100	Opholdende	6,7%
2063	600	Gaardbeboer	6,3%
1789	900	Husmand med Jord	5,4%
1750	14500	Spinder	5,3%
1658	38000	Inderste	5,0%
1301	38600	Pensionist	4,0%
916	1500	Daglejer	2,8%
874	1000	Husmand uden Jord	2,7%
776	1800	Tjenestefolk	2,4%
653	37300	Fattig	2,0%
607	38200	Logerende	1,8%
545	1600	Bonde	1,7%
367	13000	Væver	1,1%
336	14600	Syerske	1,0%
335	15200	Lever af Haandarbejde	1,0%
294	38800	Lever af sine Midler	0,9%
217	15300	Skrædder	0,7%
171	3900	Husholderske	0,5%
164	1700	Husmoder	0,5%



Figur 8.
Andelen af de mest udbredte
erhverv fordelt på mænd og
kvinder



Figur 9.
Aldersfordeling
– Aftægtsfolk

Beskæftigelse

Angivelsen af en persons beskæftigelse kan udtrykkes på mange måder. I enkestanden var der alene tale om 14.467 forskellige typer beskæftigelse – eller rettere, der var benyttet dette antal måder at angive en beskæftigelse på. De fleste er dog variationer af stavemåde eller udtryksmåde. Blot det at oplyse, at en person fik almisse, kunne udtrykkes på 1.550 måder. At være en form for pensionist kunne angives på 322 måder (og der er tale om en gruppe på blot 720 personer). Som omtalt i begyndelsen kan HISCO koderne ikke anvendes på enkestanden, og der er i stedet for anvendt de koder, der er introduceret i '1. version' af Folketællingen 1801.

Oplysningen om en persons erhverv kan findes både i feltet 'Erhverv', men også i feltet 'Stilling i husstanden'. Arbejdet med at standardisere disse felter har også været anvendt ved kodningen af feltet 'Køn'.

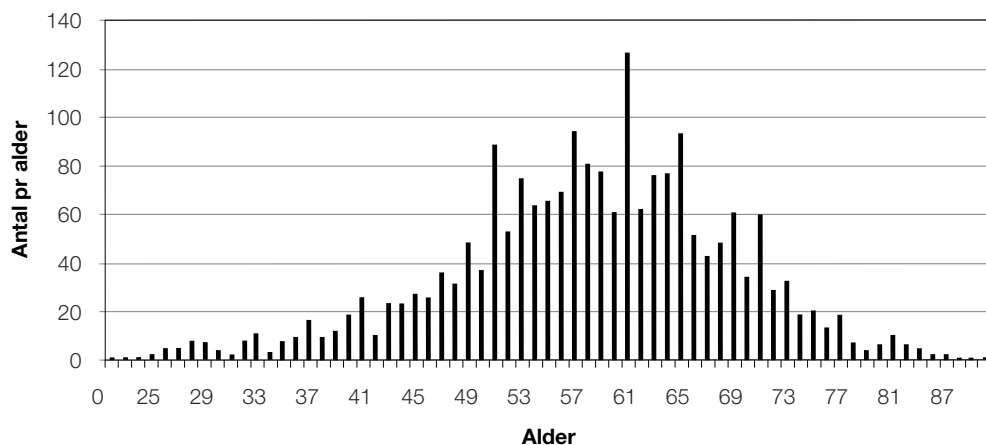
Da arbejdet med at bestemme og kode de

to felter var afsluttet, manglede 7.913 personer at få tildelt en erhvervskode. Det skyldtes oftest, at der ikke var angivet noget overhovedet om deres erhverv eller indtægtsforhold. Kodningen og standardiseringen resulterede i en reduktion af erhvervstyper til blot 220 forskellige. Nedenfor kan resultatet ses i Tabel 2 for hele gruppen og i Figur 8, der opgjort for hvert køn (hvor dog kun for de 10 hyppigste erhverv er medtaget).

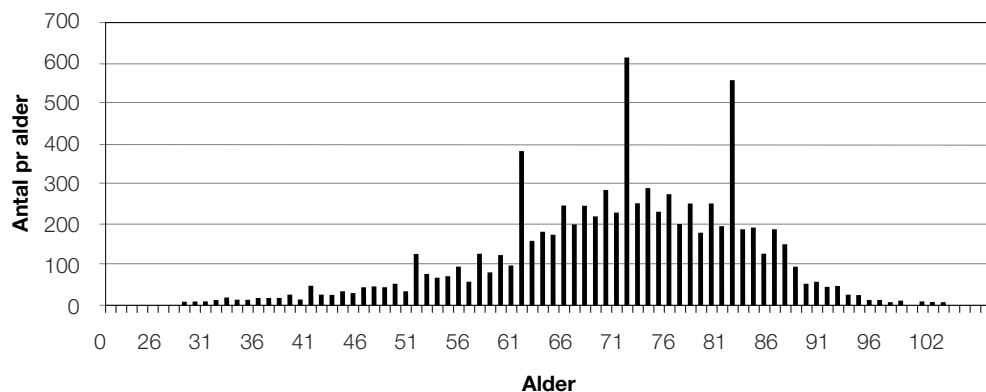
Uden standardiseringen og kodningen ville det ikke have været muligt at lave disse tabeller, selvom det medfører, at der mistes noget information. Det er især, hvor det drejer sig om en population som enkestanden, hvor en meget stor andel er uden almindelig beskæftigelse, men overlever på anden vis. Men ved ikke at have brugt vore egne koder og kun anvendt HISCO ville de fleste fra denne gruppe være blevet anbragt i den samme kategori.

Ved at se nærmere på tallene fremgår det, at de 10 mest almindelige erhverv dækkede

Figur 10.
Aldersfordeling
– Gårdbeboere

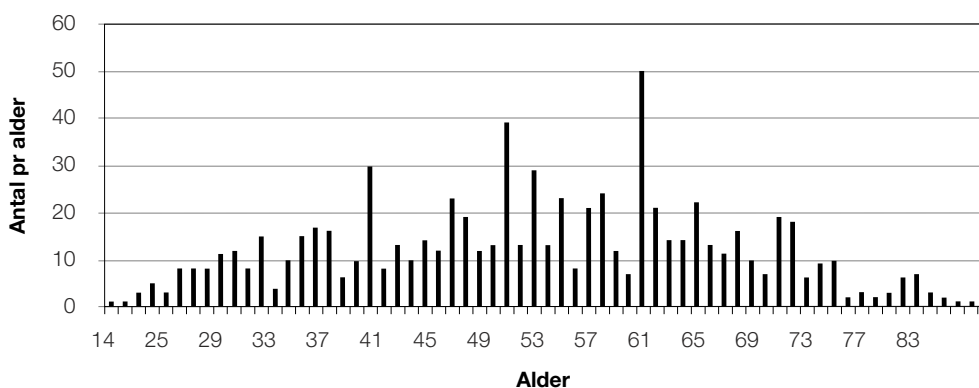


Figur 11.
Aldersfordeling
for almissemødlemmer



Tabel 3.
Erhvervsfordeling
København

Alle			Mænd			Kvinder		
Antal	Erhverv	Procent-andel	Antal	Erhverv	Procent-andel	Antal	Erhverv	Procent-andel
1299	Almisselem	43,2	169	Almisselem	31,8	1130	Almisselem	45,7
402	Pensionist	13,4	87	Arbejdsmand	16,4	327	Pensionist	13,2
215	Spinder	7,2	75	Pensionist	14,1	207	Spinder	8,4
158	Fattig	5,3	69	Tømrer	13,0	149	Fattig	6,0
143	Syerske	4,8	48	Skomager	9,0	142	Syerske	5,7
106	Tjenestefolk	3,5	45	Skrædder	8,5	106	Tjenestefolk	4,3
101	Vaskekone	3,4	9	Fattig	1,7	101	Vaskekone	4,1
96	Lever af sine Midler	3,2	9	Lever af sine Midler	1,7	87	Lever af sine Midler	3,5
87	Arbejdsmand	2,9	8	Spinder	1,5	53	Opvarningskone	2,1
69	Skomager	2,3	7	Logerende	1,3	52	Husholderske	2,1
69	Tømrer	2,3	5	Kromand	0,9	49	Logerende	2,0
56	Logerende	1,9	1	Syerske	0,2	40	Kromand	1,6
54	Skrædder	1,8				21	Skomager	0,9
53	Opvarningskone	1,8				9	Skrædder	0,4
52	Husholderske	1,7						
45	Kromand	1,5						



Figur 12.
Aldersfordeling
tjenestefolk

76 pct. af gruppens erhvervstyper og 20 pct. dækkede 89,5 pct..

Som det kan ses af oversigten, er mange af erhvervene ikke egentlige erhverv, men mere at opfatte som en måde at bo på. At være f.eks. 'indsidder' eller 'logerende' kan ikke betegnes som værende et erhverv. Men det er den eneste oplysning, der findes i folketællingen om mange personers erhverv for denne gruppes vedkommende. Denne

'beskæftigelse' er angivet for 12 pct. af personerne.

Den store proportion af kvinder i materialet afspejler sig tydeligt i erhvervsfordelingen.. Det er for eksempel forklaringen på, at 'erhvervet' almisselem er det mest udbredte erhverv. 29 pct. af kvinderne var almisselemmer mod kun 14 pct. af mændene. For kvindernes vedkommende var kun 12 pct. på aftægt mod mændenes 15 pct.

Alle			Mænd			Kvinder		
Antal	Erhverv	Procent-andel	Antal	Erhverv	Procent-andel	Antal	Erhverv	Procent-andel
898	Almisselem	20,7	90	Almisselem	9,7	763	Almisselem	23,0
527	Spinder	12,2	61	Pensionist	6,6	525	Spinder	15,8
386	Pensionist	8,9	59	Købmand	6,4	313	Pensionist	9,4
291	Logerende	6,7	58	Daglejer	6,3	236	Logerende	7,1
265	Fattig	6,1	55	Logerende	5,9	228	Fattig	6,9
220	Lever af Haand-arbejde	5,0	35	Skomager	3,8	219	Lever af Haand-arbejde	6,6
162	Daglejer	3,7	33	Fattig	3,6	117	Syerske	3,5
144	Lever af sine Midler	3,3	27	Lever af sine Midler	2,9	115	Lever af sine Midler	3,5
118	Syerske	2,7	27	Skrædder	2,9	102	Daglejer	3,1
103	Aftægtsfolk	2,4	23	Snedker	2,5	93	Aftægtsfolk	2,8
103	Opholdende	2,4	21	Tjenestefolk	2,3	81	Opholdende	2,4
93	Tjenestefolk	2,2	18	Jordbruger	1,9	70	Tjenestefolk	2,1
67	Købmand	1,6	18	Opholdende	1,9	40	Betler	1,2
51	Jordbruger	1,2	16	Tømrer	1,7	33	Jordbruger	0,1
46	Betler	1,0	15	Murer	1,6	33	Vaskekone	0,91
46	Skomager	1,0	12	Tjenestekarl	1,3	32	Væver	1,0
39	Skrædder	0,9	11	Matros	1,2	28	Kromand	0,8
39	Væver	0,9	10	Arbejdsmand	1,0	24	Jordemoder	0,7
37	Kromand	0,9	10	Fisker	1,1	22	Husholderske	0,7
36	Vaskekone	0,8	9	Aftægtsfolk	0,1	21	Bager	0,6
28	Uforståelig	0,7	9	Brændevins-brænder	0,9	20	Uforståelig	0,6

Tabel 4.
Erhvervsfordeling
Købstæder

Tabel 5.
Erhvervsfordeling
– Landsogne

Alle			Mænd			Kvinder		
Antal	Erhverv	Procent- andel	Antal	Erhverv	Procent- andel	Antal	Erhverv	Procent- andel
6091	Almisselem	24	1236	Aftægtsfolk	17	5123	Almisselem	28
4059	Aftægtsfolk	16	921	Almisselem	13	2820	Aftægtsfolk	15
2106	Opholdende	8	785	Gaardbeboer	11	1544	Opholdende	8
2063	Gaardbeboer	8	622	Husmand med Jord	9	1316	Inderste	7
1789	Husmand med Jord	7	548	Opholdende	8	1272	Gaardbeboer	7
1658	Inderste	6	403	Bonde	6	1163	Husmand med Jord	6
1008	Spinder	4	377	Tjenestefolk	5	1003	Spinder	5
874	Husmand uden Jord	3	321	Inderste	5	656	Husmand uden Jord	4
754	Daglejer	3	312	Daglejer	4	435	Daglejer	2
577	Tjenestefolk	2	217	Husmand uden Jord	3	413	Pensionist	2
544	Bonde	2	94	Pensionist	1	290	Væver	2

For kvinderne dækker de 10 hyppigste erhverv 80 pct. af de anvendte erhvervstyper. Dette betyder, at kvinderne havde færre slags erhverv: 118 forskellige typer erhverv sammenlignet med mændenes 191 erhvervstyper. For mændene dækker de 10 hyppigste erhverv kun 15 pct. af de anvendte erhvervstyper. Figur 9, 10, 11 og 12 angiver aldersfordelingen for de mest udbredte erhvervstyper.

For yderligere at belyse forholdene for enkestanden er der udarbejdet analyser af beskæftigelsestyperne i de tre geografiske områder og desuden er resultatet sammenholdt med køn. Resultaterne af disse analyser fremgår af Tabel 3, 4 og 5.

For Københavns vedkommende viste det sig, at der i alt kun var 16 forskellige erhvervskategorier, hvoraf mændene endda var fordelt på kun 12 kategorier. I købstæderne var der tale om 157 forskellige erhvervskategorier. Mændene var fordelt på 128 forskellige erhverv og kvinderne på 89 forskellige erhvervstyper.

Figurene viser, at erhvervsfordelingen for mænd er mere jævn end for kvinder: de 10 hyppigste erhverv dækker for 51 pct. af mændene hvorimod de 10 hyppigste erhverv dækker 81 pct. af kvinderne.

I landdistrikterne dækker de 10 mest udbredte kategorier 82 pct. af personerne, men her gælder det igen, at mændene har større variation i deres erhvervsmuligheder.

Da flertallet i denne undersøgelse udgøres af kvinder, er deres erhvervsvalg/erhvervsmulighed bestemmende for det generelle mønster, og det har derfor været vigtigt at skelne mellem kønnene ved beskrivelsen af erhverv og øvrige forhold.

Konklusion

Ovenstående gennemgang viser, at det er muligt at finde mange detaljer og megen information om forskellige grupper i folketællingen. Der er stadig meget tilbage at analysere, og det vil forhåbentligt også ske. I det foregående er vist nogle af de hidtil opnåede resultater, og nogle af disse kan stadig undersøges endnu mere detaljeret. De væsentligste resultater af, hvad der nu er opnået, kan sammenfattes i nedenstående fem punkter.

- Andelen af enker og enkemænd udgjorde 49.441 personer svarende til 5,25 pct. af hele befolkningen. Flertallet i denne gruppe udgjordes af enker: 75 pct. En mulig forklaring kan være, at det var sværere for enker at blive gift igen, end det var for enkemænd.
- Resultatet fra analysen af aldersfordelingen viser, at majoriteten var fra 60 år og opefter. Dette er næsten det samme for begge køn. Men når det drejer sig om fordelingen mellem kønnene for alderen

40 år og derunder, er der store kønsforskelle. I denne gruppe var der dobbelt så mange mænd som kvinder: 12 pct. mænd i forhold til 6 pct. kvinder. En sandsynlig årsag er de mange kvinder, der døde i barselsseng.

- Analysen af husstandsfordelingen viser, at 11 pct. af enker/enkemænd boede alene. Det viser også, at flertallet ikke boede alene, men enten sammen med slægtningen eller i store husstande, som det for eksempel gælder for tjenestefolks vedkommende.
- Det er måske ikke overraskende, at det største antal personer i husstande på 1 fandtes i København: 25 pct. mod 11 pct. i resten af landet. Det har endnu ikke været muligt at lave en undersøgelse af erhvervstypen for husstandsoverhovedet i disse husstande, hvor der er enker/enkemænd, og det har endnu ikke været muligt at undersøge, hvor ofte en enke/enkemand var husstandsoverhoved – når der var mere end 1 person i husstanden.

- Krydstabuleringerne mellem alder og erhverv og mellem køn og erhverv og geografisk område viser store forskelle geografisk og mellem kønnene. Som nævnt tidligere har overvægten af kvinder stor indflydelse på undersøgelsesresultatet. I byerne levede størstedelen af både mænd og kvinder af almisser, men i landdistrikterne, hvor flertallet boede, var der flere mænd på aftægt, end der var mænd, der levede af almisse.

Den generelle konklusion på undersøgelsen er, at der var store forskelle imellem kønnene: mænd giftede sig hurtigere igen, når en ægtefælle var død, og de var generelt bedre stillede end kvinderne. Men ikke overraskende var en stor del i denne gruppe fattige og havde svært ved at klare sig. På den anden side beholdt mange i denne gruppe deres erhverv til en meget høj alder – og det var nødvendigt, da der ikke fandtes en generel pensionsordning eller andet til deres forsørgelse.

Færdiganmeldelser

I det følgende beskrives 14 datamaterialer, der i perioden juni 2003 til november 2003 er blevet færdigoparbejdet i DDA og dermed er fuldt dokumenteret og sikret langtidsoptbevaring i det systemuafhængige OSIRIS-format.

Fritidsundersøgelsen 1993

Af JETTE STRAND WERMUTH, DDA

DDA-studienummer:

1828

Materialetype:

Spørgeskemadata, besøgsinterviews

År for dataindsamling:

1993

Størrelse:

1 datafil med 647 respondenter og 767 variable

Indhold:

Undersøgelsens formål var at tilvejebringe oplysninger med henblik på en statistisk og videnskabelig belysning af børns og voksnes kultur- og fritidsvaner og deres sammenhæng med arbejde, uddannelse og familieforhold. Datamaterialet er baseret på interviews med 647 personer på mellem 7 og 16 år. Forældrene, der blev interviewet ved samme lejlighed, blev spurgt om tilladelse til, at deres barn efterfølgende kunne interviewes. Nærværende datamateriale indbefatter svarene fra de interviewede børn. Endvidere er der oplysninger om deres forældre. Data om samtlige voksne respondenter (med og uden børn) fra samme stikprøve på i alt 1843 personer er oparbejdet som et særskilt datamateriale, DDA-3906. Endvidere findes svarene fra børnene alene også i et særskilt oparbejdet datamateriale, DDA-3905.

Datamaterialet indeholder nedenstående oplysninger om børn mellem 7 og 16 år, begge år inkluderet:

Samvær og aktiviteter med forældre, samvær og aktiviteter med venner, tidsforbrug i skolen og på lektier, pasningsordning efter skoletid, fritidsklub, skemalagte fritidsinteresser (art og tidsforbrug), egne (selvor-

ganiserede) fritidsinteresser, fritidsarbejde, læsning af skønlitteratur, læsning af faglitteratur, læsning af tegneserier, læsning af underholdende blade, aviser og stofområder i avisen, biblioteksbenyttelse, TV-vaner og interesseområder, radiovaner og interesseområder, biografbesøg, spillefilm på video og TV, koncertbesøg, brug af elektroniske medier, egen udøvelse af musik og sang, teaterbesøg, museumsbesøg, sport.

Datamaterialet indeholder nedenstående oplysninger om de interviewede børns forældre:

Baggrundsoplysninger om respondenter og dennes ægtefælle/samlever: køn, alder, statsborgerskab, skoleuddannelse, erhvervsuddannelse, hovedbeskæftigelse, privat/offentlig ansættelse, arbejdstid, husstandens sammensætning, herunder børnenes alder.

Materialet indeholder endvidere oplysninger om: Teaterbesøg – herunder type af teater (permanente, turnerende, amatør) og art af forestilling (skuespil, ballet, opera, musical, revy, børneteater), amatørteatervirksomhed, musik, koncertbesøg – herunder musikgenre, brug af elektroniske medier, egen udøvelse af musik og sang, biografbesøg, spillefilm på video og TV, besøg på kunstmuseer og udstillinger, besøg på andre museer, andre aktiviteter – herunder foredragsarrangementer, dyreparker, bingo, foreningsmedlemskaber – aktiviteter – tillidshverv, læsning af skønlitteratur, læsning af faglitteratur, læsning af tegneserier, læsning af underholdende blade, læsning af distriktsblade, læsning af aviser og stofområder i avisen, biblioteksbenyttelse, biblioteket som informationscenter, TV-vaner og interesseområder, radiovaner og interesseområder, sport og motion herunder egne og organiserede aktiviteter og tilskueraktiviteter, folkekirkemedlemskab, kirkegang, medlem-

skab af andre trossamfund, fritidsundervisning, deltagelse i aftenskole mv.

Dataindsamler:

Socialforskningsinstituttet

Donor og adgangskriterier:

Socialforskningsinstituttet. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Odense Kommune og borgerne 1990

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

1579

Materialetype:

Spørgeskemadata, postenquete

År for dataindsamling:

1990

Størrelse:

1 datafil med 1828 respondenter og 195 variable

Indhold:

Datamaterialet er indsamlet af ECO-Analyse for Odense Kommune med det formål at tilvejebringe et grundlag for en analyse af borgernes tilfredshed med og holdninger til kommunens service.

Undersøgelsen omfattede følgende fem generelle emner:

1. Borgernes tilfredshed med forskellige kommunale servicetilbud
2. Borgernes holdning til kommunens skatte- og servicepolitik
3. Borgernes erfaringer med kontakten til de kommunale kontorer
4. Borgernes erfaringer med de kommunale åbningstider
5. Kommunens information over for borgerne

Desuden blev fire områder gjort til genstand for en dybere undersøgelse. Det drejede sig om bytrafikken, biblioteksvæsenet, genbrugsordningerne samt grønne områder.

Undersøgelsen blev gennemført i maj og

juni 1990. Der blev udsendt spørgeskemaer til 3000 tilfældigt udvalgte borgere på 18 år og derover. Stikprøven på de 3000 personer blev delt op i to grupper på hver 1500 personer (stikprøve A og stikprøve B). Herved kunne spørgeskemaet holdes nede på en rimelig størrelse, da der i et vist omfang blev stillet forskellige spørgsmål til de to grupper. Spørgsmål, som krævede et højt respondentantal, blev stillet i begge spørgeskemaer. Endvidere blev formuleringen for de samme spørgsmål i enkelte tilfælde varieret en smule. Denne metode er anvendt på områder, hvor det undersøges, om det havde nogen betydning for holdningen blandt borgerne, om de kendte de faktiske omkostninger mm. ved kommunale serviceforanstaltninger.

I alt returnerede 1828 respondenter brugbare spørgeskemaer. Dette giver en samlet svarprocent på 61. Svarprocenten var næsten ens i de to stikprøver, nemlig 60,3 i stikprøve A og 61,5 i stikprøve B.

De returnerede skemaer svarer med hensyn til køn og alder meget fint til den samlede befolkning i Odense Kommune. Datamaterialet er med andre ord repræsentativt og kan som sådant bruges til at udtale sig med rimelig stor sikkerhed om, hvad samtlige borgere i Odense Kommune mente om kommunens service i starten af 90'erne.

Dataindsamler:

ECO-Analyse

Donor og adgangskriterier:

ECO-Analyse. Ingen adgangsrestriktioner overhovedet. Henvendelse bedes rettet til DDA.

Fysisk handicappedes vilkår 1994-1995

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

2085

Materialetype:

Spørgeskemadata, telefoninterviews

År for dataindsamling:

1995

Størrelse:

1 datafil med 1655 respondenter og 418 variable

Indhold:

Datamaterialet blev indsamlet med det formål at tilvejebringe oplysninger til en statistisk og videnskabelig belysning af handicappedes leveforhold på en række områder.

Datamaterialet beskriver handicappedes grad af integration i samfundslivet og afdækker barriererne herfor. Barriererne søges ikke kun i individet, men også i samfundets reaktion på individets funktionstab.

Data blev indsamlet i to faser. Den første fase, der gennemførtes som interviews af op til 25.000 personer i 1994, afklarede, om respondenterne havde fysiske handicaps. I den anden fase blev de, der i første fase oplyste, at de havde fysiske handicaps, interviewet om deres vilkår.

Datamaterialet indeholder følgende oplysninger:

Fase 1:

Baggrundsoplysninger: Køn, alder, husstandssammensætning, uddannelse, stilling,

erhverv, forsørgelsesgrundlag. Endvidere er der screeningsoplysninger til fase 2 om funktionsevne og helbred.

Fase 2:

Baggrundsoplysninger: Samme oplysninger som indsamlet i fase 1, indkomstkilder og indkomstens størrelse, boligens indretning og adgangsforhold.

Funktionsevne- og helbredsoplysninger: Samme oplysninger som indsamlet i fase 1, diagnoser, brug af tekniske hjælpemidler.

Handicappets baggrund: Varighed og udvikling, selvtillid/oplevet indflydelse på egen situation.

Relation til arbejdsmarkedet: Nuværende og tidligere tilknytning, arbejdstid, jobforløb (syge og arbejdsløshedsperioder), revalideringsforløb.

Arbejds miljø: Arbejdspladsens indretning, erfaringer med holdninger hos arbejdsgiver og kolleger, egne holdninger til arbejde.

Relation til de offentlige systemer: Former for modtaget hjælp/service, erfaringer og tilfredshed med den offentlige service.

Sociale kontakter: Deltagelse i samfundslivet, herunder i handicapforeninger.

Dataindsamler:

Socialforskningsinstituttet

Donor og adgangskriterier:

Socialforskningsinstituttet. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Odense Kommune og borgerne 1995

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

2736

Materialetype:

Spørgeskemadata, postenquete

År for dataindsamling:

1995

Størrelse:

1 datafil med 1787 respondenter og 200 variable

Indhold:

Datamaterialet er indsamlet af ECO-Analyse for Odense Kommune med det formål at tilvejebringe et grundlag for en analyse af borgernes tilfredshed med og holdninger til kommunens service. Endvidere har formålet været at sammenligne resultaterne med resultater fra en lignende undersøgelse i 1990 (DDA-1579: Odense Kommune og borgerne 1990).

Respondenterne blev spurgt om deres brug af og tilfredshed med forskellige kommunale servicetilbud. Undersøgelsen omfatter desuden borgernes holdning til kommunens skatte- og servicepolitik, tilfredsheden med de kommunale åbningstider samt kommunens information over for borgerne. Endvidere er der særlig fokus på borgernes syn på og/eller tilfredshed med udvalgte serviceområder, nemlig hjemmehjælp, børnepasning, folkeskoler, svømmehaller og indsatsen på energi- og miljøområdet. Endelig er der i forlængelse af det aktuelle kommuneplanarbejde en række spørgsmål omkring Odense bymidte.

Der blev udsendt spørgeskemaer til 3000 tilfældigt udvalgte borgere på 18 år og derover. Respondentgruppen blev delt op i to undergrupper på hver 1500 personer (stikprøve A og stikprøve B). Formuleringen for et enkelt spørgsmål angående en passende størrelse for kommuneskatten er varieret en smule. Denne metode er anvendt for at undersøge, hvorvidt det havde nogen betydning for svarene, om skattens størrelse blev koblet sammen med det kommunale serviceniveau i spørgsmålsformuleringen.

I alt returnerede 1787 respondenter brugbare spørgeskemaer. Dette giver en samlet svarprocent på 60. Svarprocenten er ens i de to stikprøver.

Dataindsamler:

ECO-Analyse

Donor og adgangskriterier:

ECO-Analyse. Ingen restriktioner ved videnskabelig/statistisk brug. Henvendelse bedes rettet til DDA.

Unge andengenerations indvandrere 1999

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

4956

Materialetype:

Spørgeskemadata, besøgsinterviews

År for dataindsamling:

1999

Størrelse:

1 datafil med 693 respondenter og 325 variable

Indhold:

Datamaterialet er indsamlet med det formål at belyse, i hvilken udstrækning personer med

baggrund i Tyrkiet, Pakistan og Jugoslavien er integreret i det danske samfund. Materialet indeholder informationer om respondenternes uddannelse, arbejde, bolig, familiedannelse, fritid og sociale kontakter samt civilstand. Endvidere indeholder materialet spørgsmål om: deltagelse i organisationsarbejde, politisk deltagelse og politiske holdninger.

Dataindsamler:

Socialforskningsinstituttet

Donor og adgangskriterier:

Socialforskningsinstituttet. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Borgerne og lovene 1997

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

10416

Materialetype:

Spørgeskemadata, besøgsinterviews

År for dataindsamling:

1997

Størrelse:

1 datafil med 3001 respondenter og 338 variable

Indhold:

Baggrunden for indsamlingen af dette data-materiale er Rockwool Fondens Forskningsenheds forskning i den sorte økonomi i Danmark. Dette arbejde har før indsamlingen af nærværende datamateriale været koncentreret om årsagerne til den sorte sektors vækst ud fra en økonomisk synsvinkel. Ikke-økonomiske forhold som for eksempel moral blev ikke inddraget. For at kunne belyse moralen på skatteområdet i et bredere perspektiv besluttede bestyrelsen for Rockwool Fonden i 1996 at iværksætte et forskningsprojekt, der kunne skabe et overblik over moralen omkring forskellige dele af lovgivningen som for eksempel skatteområdet, det sociale område, trafik- og straffelove, forsikringsområdet m.v.

Den væsentligste del af undersøgelsens

materiale består af svarene på interviews med ca. 3000 respondenter. Disse interviews blev lagt tilrette og gennemført under Jørgen Goul Andersens ansvar. Hans Jørgen Nielsen supplerede denne del af undersøgelsen med en række dybdeinterviews, der skulle nuancere og uddybe resultaterne af hovedundersøgelsen.

Følgende baggrundsvariable indgår i undersøgelsen: Respondentens køn, alder, skoleuddannelse, erhvervsuddannelse, personlig indkomst, husstandsindkomst, stilling, ansættelsessektor samt respondentens forældres stilling.

Informationen om borgernes moral omfatter blandt andet: Spørgsmål om holdninger til og erfaring med skattesnyd, socialt bedrag, forsikringssvindler, trafikovertrædelser, sort arbejde og butikstyveri. Endvidere er der spørgsmål om politiske holdninger og tilfredshed med offentlige myndigheder.

Dataindsamler:

Gallup

Donor og adgangskriterier:

Jørgen Goul Andersen, Gunnar Viby Mogensen, Hans Jørgen Nielsen og Marie Louise Hultin. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Trafik og levevilkår i Kgs. Enghave 2001

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

10829

Materialetype:

Spørgeskemadata, postenquete

År for dataindsamling:

2001

Størrelse:

1 datafil med 375 respondenter og 178 variable

Indhold:

Baggrunden for denne dataindsamling er paradokset mellem Kgs. Enghaves store trafikbelastning og de relativt få bilejere i kvarteret, der er en af de „bilfattige“ bydele i København. I kvarterløftplanen for Kgs. Eng-

have hedder det: „Kgs. Enghave er blevet en bydel, andre bare kører igennem. Trafikmængden er så massiv, at grænseværdierne for støj og forurening i forhold til boligområder langt fra overholdes. Trafikårener tværs gennem bydelen besværliggør færdslen og kan være en barsk oplevelse især for svage trafikanter som børn, ældre og handicappede. Desuden skaber trafikårener både markante barrierer mellem de enkelte kvarterer i bydelen, og de indkapsler bydelen i en trafikfælde og spærrer for en sammenhæng med omgivende grønne arealer og havn.“

(Kgs. Enghave Kvarterplan 1998, side 25). På denne baggrund var projektets formål: For det første at analysere sammenhængen mellem trafikbelastning, velfærd og levevilkår i den københavnske bydel Kgs. Enghave. For det andet at analysere erfaringerne med borgerinddragelse og belyse, hvilken rolle trafikpolitikken spiller i den samlede kvar-

terløftpolitik. For det tredje at udvikle nye metoder til empirisk belysning og måling af trafikbelastning.

Datamaterialet indeholder udover nogle få baggrundsspørgsmål en række spørgsmål om respondenternes tilfredshed med at bo i Kgs. Enghave og i særdeleshed spørgsmål om respondentens trafikforbrug syn på trafiksituationen i Kgs. Enghave.

Dataindsamler:

Gallup for Roskilde Universitetscenter og Transportrådet

Donor og adgangskriterier:

John Andersen, Lars Bohn og Trine Nordgaard. Ingen restriktioner overhovedet for anvendelse, men enhver form for publikation kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Valgdeltagelse ved kommunalvalget i 1997 i København og Århus

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

11062

Materialetype:

Registersamkørsel

År for dataindsamling:

1997

Størrelse:

1 datafil med 628.659 cases og 16 variable

Indhold:

En af de bedste indikatorer på et velfungerende demokrati er en høj valgdeltagelse. I Danmark har valgdeltagelsen længe været høj. Hvad folketingsvalg angår, skal vi tilbage til 2. verdenskrig for at finde en valgdeltagelse på under 80 procent. Ved kommunalvalgene er valgdeltagelsen dog noget lavere, idet den i de seneste 50 år har svinget omkring 70 procent med en svagt nedadgående tendens.

Der findes i Danmark en vis viden om, hvem der stemmer, og hvem der ikke gør det. Eksempelvis stemmer de yngste og de ældste aldersgrupper mindre end andre. Dertil findes

der viden om, at arbejdsløse og ufaglærte arbejdere stemmer mindre end andre, samt at nytilflyttede stemmer mindre ved kommunalvalgene end folk, der har boet længere tid i kommunen. Meget tyder således på, at der er forbundet en social og erhvervsmæssig marginalisering med en lav valgdeltagelse. Grundlaget for denne viden er imidlertid ikke særligt solidt. Dette skyldes, at viden om valgdeltagelsen i hovedsagen bygger på spørgeskemaundersøgelser. Problemet med spørgeskemaundersøgelser som kilde er, at de mennesker, der ikke deltager i valgene, er de samme som dem, man ikke kan få til at deltage i spørgeskemaundersøgelserne.

Dertil kommer, at folk er tilbøjelige til at skjule deres manglende valgdeltagelse over for interviewer. Resultatet er, at valgdeltagelsen i spørgeskemaundersøgelser altid fremgår som højere, end den i virkeligheden er. Oplysninger fra spørgeskemaundersøgelserne er således behæftet med en betydelig usikkerhed.

Dertil kommer, at der stort set ingen viden er om indvandrernes brug af den kommunale valgret. Ved kommunalvalget i 1981 kunne

indvandrere uden dansk statsborgerskab, men med mindst tre års ophold i Danmark for første gang stemme ved et valg i Danmark. Ved kommunalvalget i 1997 var det således femte gang, at indvandrere uden statsborgerskab kunne stemme. Bortset fra enkelte summariske oplysninger fra 1981 og 1985 var der ingen viden om, i hvilket omfang de udnyttede denne rettighed og heller ingen viden om, hvorvidt forskellige grupper af indvandrere anvendte valget i forskelligt omfang.

Med ovenstående udgangspunkt gennemførte Institut for Statskundskab i Århus i samarbejde med Københavns og Århus kommune en undersøgelse af valgdelta-

gelsen ved kommunalvalget i 1997 blandt samtlige valgberettigede i de to kommuner. Det empiriske grundlag er den foreliggende registersamkørsel fra de to kommuner.

Dataindsamler:

Lise Togeby, Jørgen Elklit, Palle Svensson, Birgit Møller, Århus og Københavns kommuner

Donor og adgangskriterier:

Lise Togeby, Jørgen Elklit, Palle Svensson og Birgit Møller. Enhver adgang kræver et specielt arrangement med donor, i almindelighed ingen adgang overhovedet. Henvendelse bedes rettet til DDA.

Oplevet diskrimination 1998

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

11127

Materialetype:

Spørgeskemadata, telefoninterviews og post-enquete

År for dataindsamling:

1998

Størrelse:

1 datafil med 1132 respondenter og 143 variable

Indhold:

Datamaterialet er indsamlet i forbindelse med en undersøgelse, som har til formål at dokumentere omfanget af den oplevede diskrimination blandt etniske minoriteter i Danmark. Nævnet for Etnisk Ligestilling besluttede i maj 1998 at gennemføre en landsdækkende spørgeskemaundersøgelse blandt etniske minoriteter med oprindelse i Bosnien, Libanon, Somalia og Tyrkiet for at undersøge omfanget af den oplevede diskrimination i hverdagslivet, på arbejdsmarkedet og ved etniske minoriteters møde med offentlige institutioner.

Der er almindelig enighed om, at der i begrebet diskrimination indgår to hovedele-

menter. For det første indgår forskelsbehandling af personer eller grupper, og for det andet indgår, at denne forskelsbehandling opfattes som moralsk uacceptabel eller som ulovlig. Diskrimination, der måles i denne undersøgelse, er de etniske minoriteters oplevelse af diskriminationen, hvilket stiller spørgsmålet, om de etniske minoriteter oplever, at de er udsat for en forskelsbehandling, der bygger på etniske tilhørsforhold. De forskelle, der optræder mellem forskellige gruppers diskrimination, kan skyldes, at nogle grupper udsættes for mere diskrimination end andre, men det kan også skyldes, at nogle grupper er mere tilbøjelige til at tolke oplevelser i det danske samfund som udslag af diskrimination, og endelig kan det skyldes, at nogle grupper er mere tilbøjelige til at fortælle om den oplevede diskrimination end andre.

Interviewene er gennemført af tosprogede interviewere. Interviewpersonerne kunne selv vælge, om interviewene skulle gennemføres på dansk eller på deres modersmål. Da der viste sig at være store problemer med at fremskaffe telefonnumre på de udvalgte personer, er telefoninterviewene suppleret med postspørgeskemaer. I undersøgelsen indgår der i alt 1132 personer mellem 18 og 66 år, som har haft ophold i Danmark i mindst 3 år, heraf 342 fra Bosnien, 226 fra Somalia, 279

fra Libanon og 285 fra Tyrkiet. Svarprocenten er 58 procent for bosnierne, 48 procent for tyrkerne, 47 procent for libanenserne og 40 procent for somalierne. Den lave svarprocent skyldes ikke modvilje mod at deltage i undersøgelsen, men at det har været svært at få kontakt med de udvalgte personer.

Undersøgelsen indeholder et stort antal spørgsmål om oplevet diskrimination inden for en række områder af samfundslivet. Det første område omfatter arbejdsmarkedet og erfaringer på arbejdspladsen, mens det andet område omfatter det, der sker på gaden, i offentlige transportmidler, forretninger og forlystelseslivet. Der er her tale om den anonyme og tilfældige kontakt mellem danskere og etniske minoriteter. Det tredje område omfatter problemer i forbindelse mellem religionsudøvelsen, og det sidste område omfatter kontakt med offentlige myndigheder.

At der er store forskelle på, hvor stor diskrimination de forskellige grupper oplever, betyder ikke, at alle medlemmer af en etnisk gruppe oplever samme grad af diskrimination. Faktorer som køn, alder eller danskkundskaber kan spille en rolle for oplevelsen af diskrimination. En anden faktor, som påvirker oplevelsen af diskrimination, er tilliden til andre mennesker, hvor det blandt bosniere, tyrkere og libanesiske mænd er dem, som har lille tillid til andre mennesker, som føler sig mest diskrimineret.

Dataindsamler:

Danmarks Statistik

Donor og adgangskriterier:

Lise Togeby og Birgit Møller. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Trivsel, livsvilkår og livskvalitet under krisen 1982-1994

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

11521

Materialetype:

Spørgeskemadata, telefoninterviews

År for dataindsamling:

1982, 1986, 1988, 1990, 1994

Størrelse:

1 datafil med 9156 respondenter og 211 variable

Indhold:

Data er indsamlet i forbindelse med projektet: „Danskernes livskvalitet og trivsel under krise og opsving“ („Kriseprojekt“), hvori der indgår fem undersøgelser med dataindsamlinger i maj 1982 (DDA-1610), i april/maj 1986 (DDA-1433) i november 1988 (DDA-1611, i november 1990 (DDA-1612) og i april/maj 1994 (DDA-1898)).

Projektets formål er på basis af en række psykologiske „krisevariable“ at undersøge, om der kan konstateres psykologiske udviklingstendenser i den danske befolkning, hvad angår

krisebevidsthed, livskår, livskvalitet, livsstil, handlinger og holdninger under krisen i den undersøgte periode.

Undersøgelsernes variable kan opdeles efter fem temaer:

Befolkningens krisebevidsthed: Placering af ansvaret for krisen – internt og eksternt, vurdering af udviklingen i de kommende år (fremtidssyn), samt offervilje med hensyn til afhjælpning af krisen.

Befolkningens livsvilkår under krisen: Oplevede livsvilkårsændringer i de senere år af økonomiske forhold, boligforhold, arbejdsforhold samt forbrugsmuligheder og „livsvilkår i øvrigt“.

Befolkningens psykiske eller subjektive livskvalitet: Samfundstrivsel, almen trivsel, psykisk stress og oplevelsen af at være kriseramt.

Befolkningens livsstil og handling under krisen: Ressourcebevidst livsstil ydre, offensiv handling samt aktiv og adaptiv resignation.

Befolkningens holdninger under krisen måles ved: Holdning til statens krisebekæmpende rolle samt holdning til arbej-

de og arbejdsløshed. Endvidere er en ny variabel medtaget fra 1990-materialet, nemlig: Befolkningens holdning til selvsvarlighed.

Undersøgelsesmetoden samt spørgsmålssættet har i alle delundersøgelser været det samme, men gennem årene er der suppleret med specifikke spørgsmål om blandt andet kriseårsager, overarbejde, offentlige styringsmidler samt holdning til alkohol.

Dataindsamler:

Socialforskningsinstituttet

Donor og adgangskriterier:

Bo Sommerlund, Eggert Pedersen, Jørgen Goul Andersen, Knud Erik Sabroe, Jørgen Dalberg-Larsen og Ole Steen Kristensen. Ingen restriktioner for anvendelse, men enhver form for publikation kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

De danske partimedlemmer 2000

Af BIRGITTE GRØNLUND JENSEN, DDA

Materialetype:

Spørgeskemadata, postenquete

DDA-studienummer:

14069

År for dataindsamling:

2000-2001

Størrelse:

1 datafil med 5266 respondenter og 334 variable

Indhold:

Datamaterialets variable kan inddeles i syv grupper:

Baggrundsspørgsmål: Køn, fødselsår, bopælskommune, statsborgerskab, etniske baggrund, husstandsindkomst, skoleuddannelse, erhvervsuddannelse, hovedbeskæftigelse samt ansættelsessektor.

Indmeldelse i politisk parti: År for indmeldelse, begrundelse for indmeldelse, udmeldelses overvejelse.

Kontakt med partiet: Omfang af kontakt med partiet, tilfredshed med omfang af kontakt med partiet, kontingent størrelse, tilfredshed med kontingent størrelse, omfang af elektronisk kontakt med partiet, omfang af mødedeltagelse, tilfredshed med møder.

Aktiviteter i partiet: Respondenternes nuværende og tidligere aktivitetsniveau i partiet, omfang af deltagelse i forskellige typer af partiaktiviteter, begrundelser for eventuelt faldende aktivitetsniveau, ønske om opstilling til valg, tillidsposter i partiet samt repræsentation i forskellige offentlige organer, medlemskab af andre organisationer.

Kommunalvalget 1997 og folketingsvalget 1998: Respondenternes stemmeafgivning ved de to valg, aktivitetsniveau i forbindelse med valgkampene samt holdninger til partiernes valgkampagner.

Politiske holdninger: Respondenternes holdninger til 12 forskellige politiske emner, holdninger til størrelsen af offentlige udgifter inden for 10 politik områder, selvplacering og partiplacering på venstre/højre skalaen samt på flygtninge politiske, miljø politiske, velfærds politiske, rets politiske og EU politiske dimensioner, holdninger til 10 politiske partier, tilfredshed med det danske demokrati, holdning til forskellige demokrati modeller, holdninger til politisk repræsentation.

Holdninger til respondentens eget parti: Generelt indtryk af partiet, holdninger til i hvor høj grad partiet bekymrer sig om særlige befolkningsgrupper, holdninger til 13 udsagn om partiets organisation, holdninger til 13 udsagn om medlemmernes rolle i partierne.

Dataindsamler:

Lars Bille, Roger Buch, Jørgen Elklit, Bernhard Hansen, Hans Jørgen Nielsen og Karina Pedersen.

Donor og adgangskriterier:

Lars Bille, Roger Buch, Jørgen Elklit, Bernhard Hansen, Hans Jørgen Nielsen og Karina Pedersen. Ingen restriktioner ved videnskabelig/statistisk brug. Henvendelse bedes rettet til DDA.

Fritidsundersøgelsen 1998, børn

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

14701

Materialetype:

Spørgeskemadata, besøgsinterviews

År for dataindsamling:

1998

Størrelse:

1 datafil med 660 respondenter og 266 variable

Indhold:

Undersøgelsens formål var at tilvejebringe oplysninger med henblik på en statistisk og videnskabelig belysning af børns kultur- og fritidsvaner. Datamaterialet er baseret på interviews med børn mellem 7 og 15 år, idet deres forældre forudgående blev spurgt, om de ville give tilladelse til, at deres barn blev interviewet.

Følgende oplysninger om børn mellem 7 og 15 år, begge år inkluderet, indgår i undersøgelsen:

Samvær og aktiviteter med forældre, samvær og aktiviteter med venner, tidsforbrug i skolen og på lektier, pasningsordning efter skoletid, fritidsklub, skemalagte fritidsinteresser (art og tidsforbrug), egne (selvorganiserede) fritidsinteresser, fritidsarbejde, læsning af skønlitteratur, læsning af faglitteratur, læsning af tegneserier, læsning af underholdende blade, aviser og stofområder i avisen, biblioteksbenyttelse, TV-vaner og interesseområder, radiovaner og interesseområder, biografbesøg, spillefilm på video og TV, koncertbesøg, brug af elektroniske medier, egen udøvelse af musik og sang, teaterbesøg, museumsbesøg, sport.

Dataindsamler:

Socialforskningsinstituttet.

Donor og adgangskriterier:

Socialforskningsinstituttet. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Fritidsundersøgelsen 1998, voksne

Af BIRGITTE GRØNLUND JENSEN, DDA

DDA-studienummer:

14702

Materialetype:

Spørgeskemadata, besøgsinterviews telefoninterviews

År for dataindsamling:

1998

Størrelse:

1 datafil med 1364 respondenter og 524 variable

Indhold:

Undersøgelsens formål var at tilvejebringe oplysninger med henblik på en statistisk og videnskabelig belysning af den voksne

befolknings kultur- og fritidsvaner og disses sammenhænge med arbejde, uddannelse og familie. Voksne uden børn er interviewet over telefon, mens voksne, hvis børn også er blevet interviewet (jævnfør DDA-14701), er interviewet ved besøg.

Følgende oplysninger om voksne på 16 år og derover indgår i undersøgelsen:

Baggrundoplysninger om respondenter og dennes ægtefælle/samlever: køn, alder, statsborgerskab, skoleuddannelse, erhvervsuddannelse, hovedbeskæftigelse, privat/offentlig ansættelse, arbejdstid, husstandens sammensætning, herunder børnenes alder.

Materialet indeholder endvidere oplysninger om: Teaterbesøg – herunder art af forestilling (skuespil, ballet, opera, musical, revy, børneteater), musik, koncertbesøg

– herunder musikgenre, brug af elektroniske medier, egen udøvelse af musik og sang, biografbesøg, spillefilm på video og TV, besøg på kunstmuseer og udstillinger, besøg på andre museer, andre aktiviteter – herunder foredragsarrangementer, dyreparker, bingo, foreningsmedlemskaber – aktiviteter – tilidshverv, læsning af skønlitteratur, læsning af faglitteratur, læsning af tegneserier, læsning af underholdende blade, læsning af distriktsblade, læsning af aviser og stofområder i avisen, biblioteksbenyttelse, biblioteket som informationscenter, TV vaner og interesseområder, radiovaner og interesseområder, sport

og motion herunder egne og organiserede aktiviteter og tilskueraktiviteter, folkekirke-medlemskab, kirkegang, medlemskab af andre trossamfund, fritidsundervisning, deltagelse i aftenskole mv.

Dataindsamler:

Socialforskningsinstituttet

Donor og adgangskriterier:

Socialforskningsinstituttet. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Arbejde og hjemmeliv før og nu 1950-1988

Af JETTE STRAND WERMUTH, DDA

DDA-studienummer:

13776

Materialetype:

Spørgeskemadata, postenquete

År for dataindsamling:

1988.

Størrelse:

1 datafil med 4627 respondenter og 92 variable

Indhold:

Dette datamateriale er indsamlet blandt et udsnit af den voksne befolkning som et led i et forsøg på at måle underdeklarationens

omfang i perioden fra 1950'erne til 1980'erne. Fokus er særligt på nogle af tiårene efter 2. Verdenskrig. Interessen går tidsmæssigt op til omkring eller lige efter 1980. Herefter blev de danske målinger af 'sort arbejde' relativt systematiske. Et andet led i projektet er forskning baseret på arkivalier i Danmarks Statistik, i en stor dansk provinsbys skatteforvaltning samt i Rigsarkivet, Dansk Data Arkiv m.v.

Dataindsamler:

Danmarks Statistik

Donor og adgangskriterier:

Rockwool Fondens Forskningsenhed. Enhver adgang kræver donors skriftlige tilladelse. Henvendelse bedes rettet til DDA.

Beretning for Dansk Data Arkiv 2002

Af HANS JØRGEN MARKER, DDA

Generelt

Generel ledelse og administration

I 2002 tidsregistrerede DDA-medarbejderne 26510 timer (2001: 31108) svarende til 16,2 årsværk (2001: 19,1). Disse årsværk blev præsteret af i alt 21 personer. Ved årets begyndelse var der 21 personer ansat. I løbet af året blev Ole Oluf Andersen ansat som akademisk medarbejder. Følgende fratrådte i løbet af året: informatikassistent Carsten Ernst, akademisk medarbejder Heidi Witendorff Sørensen, informatikassistent Jan Jensen, programmør Knud Petersen, akademisk medarbejder Martin Hother Nielsen og dataassistent Ulla Andersen. I løbet af året var kontorassistent Lilly Schmidt ansat i jobtræning. Akademisk medarbejder Anne Sofie Fink blev i løbet af året ansat som forsker. Ved årets udgang var der 14 medarbejdere.

Økonomi

Da der i løbet af året slet ikke blev anskaffet arbejdsstationer, kom DDA ud af året med en ret betydelig opsparing.

IT-arkivalier

Sagsbehandling

Tilgangen af materialer fra Kildeindtastningsprojektet udviste igen en stigning. Det må ventes, at der på et tidspunkt vil indtræde en stagnation.

Øvrige lokaliseringer var noget mindre end året før, men dog stadig på et højt niveau. ERAS-projektet udgjorde omkring

en trediedel af tilvæksten. For samfundsvidenskabelige forskningsprojekter var der tale om den største tilvækst siden 1994. Der var således også for disse projekter tale om et godt resultat.

Antallet af arkiverede datamaterialer fra andet end KIP var lavere end året før. Fordelingen på fag var 70 fra samfundsvidenskab, 3 fra historie og 11 fra sundhedsvidenskab. Resultatet er meget højt for samfundsvidenskab.

Antallet af færdigoparbejdede studier er betydeligt lavere end året før. Det skyldes især den store nedgang i timerressourcer, som DDA oplevede i løbet af året. Det medførte, at der i 2001 kunne anvendes 5242 timer på oparbejdning mod 7893 året før. Det skyldes dog også, at de 28 færdiggjorte studier i gennemsnit har været meget arbejdskrævende. Der er gennem årene anvendt 6680 timer på de studier, som blev færdige i 2002.

ERAS

Forhåbningerne om, at ERAS projektet-umiddelbart ville blive fortsat som en indlejring i DDA under finansloven 2003, gik ikke i opfyldelse. Den sidste del af året blev derfor anvendt på at foretage en oprydning i de sundhedsvidenskabelige aktiviteter, hvorefter ERAS medarbejderne forlod DDA med årets udgang.

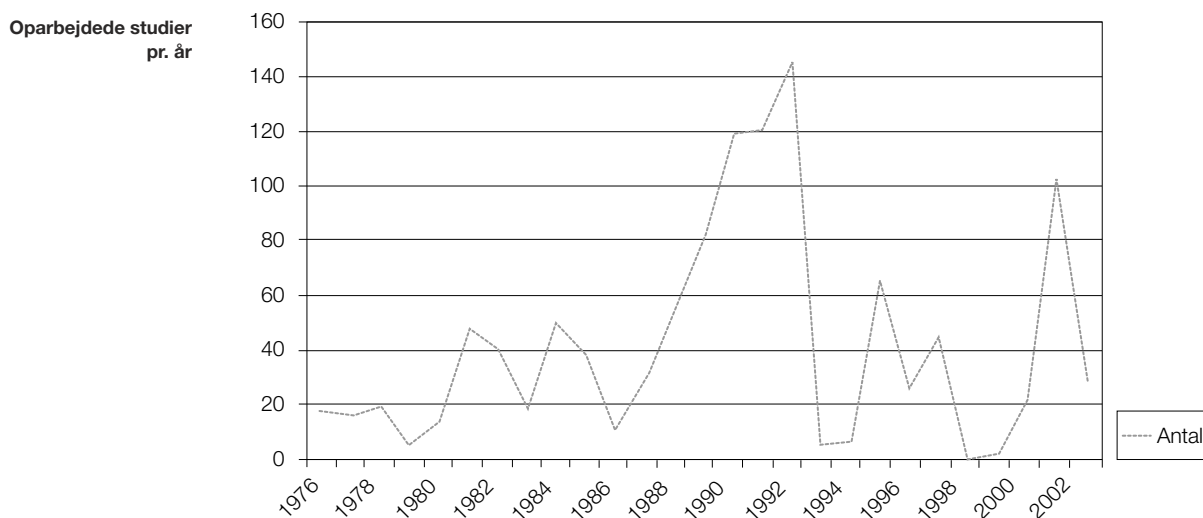
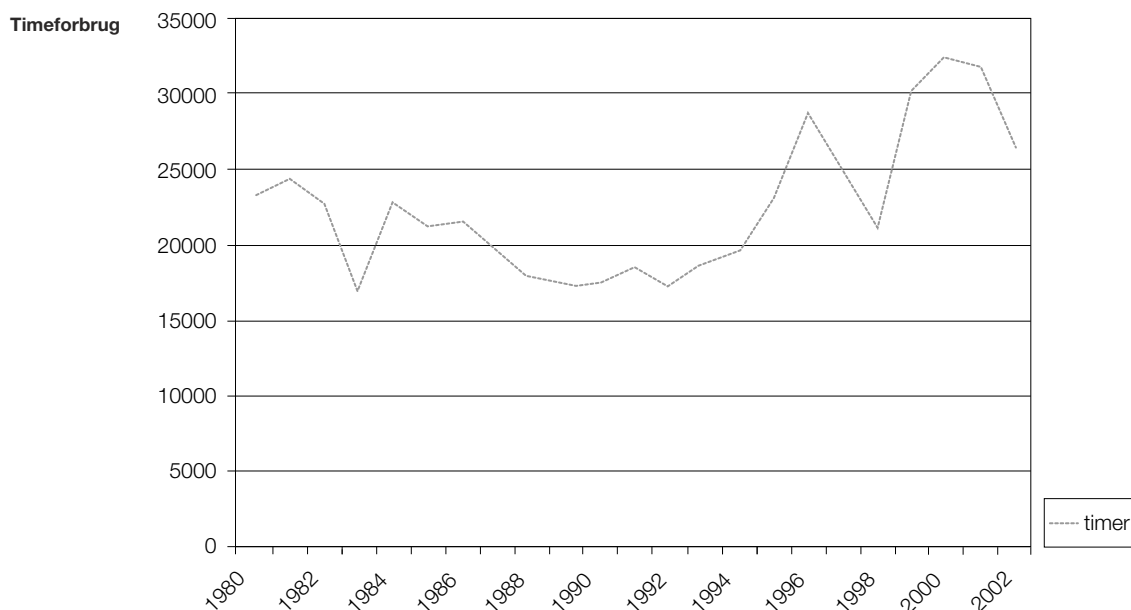
Kildeindtastningsprojektet

Følgegruppen for Kildeindtastningsprojektet, KIK, afholdt møder i januar, maj og oktober i DDA.

Der blev produceret en cd-rom, DDD-4,

År	1998	1999	2000	2001	2002
Lokaliseringer	1127	1438	2079	2188	2549
Lokaliseringer uden KIP	48	62	109	159	125
Arkivering	1116	1472	1998	2151	2687
Arkivering uden KIP	27	96	25	122	84
Oparbejdning	0	1	21	102	28

Sagsbehandling



med data fra alle andre folketællinger end 1801.

I oktober blev det sidste kursus i Kildeindtastning afholdt i DDA. Kurserne, som ellers var velbesøgte og udbytterige, var også en betydelig omkostning for DDA. Særlig problematisk var det, at DDA vedligeholdte et kursuslokale, som stort set kun anvendtes til de to årlige KIP-kurser. Maskinerne i lokalet var efterhånden temmelig forældede, og der ville derfor kræves en ikke ubetydelig investering for at bringe lokalet i tidssvarende stand.

IT-systemer

NESSTAR

FASTER-projektet fortsatte tre måneder ind i 2002.

StudJour

Et program til registrering af studierelevant information blev udviklet i løbet af året. I programmet StudJour gemmes oplysninger om sager, akter, status, involverede personer, servicetype og lignende vedrørende DDA's studier. Til akterne gemmes endvidere maskinlæsbare dokumenter. Til visse funktioner under lokalisering og arkivering

The screenshot shows the StudJour application window. At the top, there's a menu bar with options like 'Sag', 'Aft', 'Adressat', 'Journalplan', 'Opbej', 'Statistik', 'Funktioner', 'System', and 'Hjælp'. Below the menu, there are several input fields and dropdown menus for 'Journalnummer', 'Sagbeholder', 'Projekt', 'Titel', 'StudeStatus', 'Adgangskategori', 'Opbej', and a checkbox for 'Har data på ERAS maskinen'. There are also fields for 'Kontaktperson' and a table for 'Personundersøgelser' with columns 'Id' and 'Navn'. Below this is a table for 'Data' and another for 'Adgangsgivende myndighed'. A 'Viser 12 af 584' indicator and navigation buttons are present. At the bottom, there are two tables: 'Arkiver' and 'Sagsreferencer'.

Id	Navn
N.00043	Gallup A/S

Id	Navn
N.00043	Gallup A/S

Id	Navn
N.00043	Gallup A/S

Arkiver	Erhvervsdato	Journaldato	Type	Beskrivelse
1	2 19. juli 2001	1. august 2001	N	Rettelser til kodebogen
2	1 17. juni 1997	4. september 1997	U	Benyttet vægtningsmetode i Gallup Omnibus årgang 1975-1977

Journalnr	Titel
1395-64/0036	AU Børn, Ole Gallup data til AU
1395-64/0077	DK Jensen, Kelly: Oversigt over Gallups politiske index
1395-64/0435	KU Wiese, Tordem: Tidlige Gallup-årgange som led i SHF's projekt, 'Nyere dansk politik ->

kan der udskrives standardbreve fra StudJour. Endvidere indeholder StudJour nogle standardstatistikker. Funktionaliteten fra programmet StudStyr, som blev udviklet i 2001, er indkorporeret i StudJour.

StudJours database har arvet informationen fra StudStyr og Doculive. Efter ibrugtagningen af StudJour blev programmet KIPAdmin udviklet i 2000 også omlagt i en ny version, således at tabellerne fra KIPAdmins database blev integreret i StudJour databasen. Efter denne omlægning er programmerne StudJour og KIPAdmin altså forskellige brugergrænseflader til den samme database.

SDSKriv

Der blev indledt arbejde med et program til redigering af studiebeskrivelses information. Traditionelt har studiebeskrivelsesinformation været vedligeholdet i tekstfiler. Ved færdiggørelsen af SDSKriv skal denne information være placeret i StudJour databasen, ligesom SDSKriv i størst muligt omfang trækker på den information, som allerede er at finde i StudJours tabeller.

Øvrige systemer

Der blev foretaget vedligeholdelse af flere systemer herunder Bibliotekssystemet, DDABib, og filhåndteringssystemet, FStud.

Formidling

Tilgængeliggørelse – Metode & Data

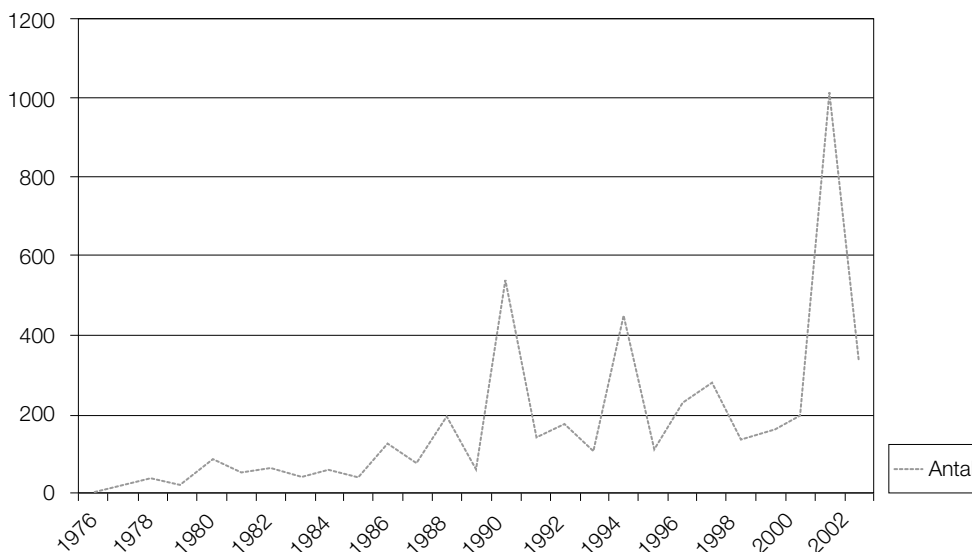
I løbet af året udkom planmæssigt to numre af Metode & Data. Det var numrene 86 og 87. Nummer 86 var et særnummer om arkivering af sundhedsvidenskabelige data. Metode&Data 87 indeholdt to artikler og DDA's beretning for år 2000.

Kurser, foredrag og rundvisninger

I april havde DDA besøg af et hold metode-studerende fra økonomistudiet ved Syddansk Universitet. Disse fik en introduktion til DDA's virke, materialer og ydelser. Ligeledes i april havde DDA besøg af et hold studerende fra Moesgaard, som drøftede forskningsmæssige muligheder og resultater på grundlag af folketællingen 1801.

Der blev afholdt kurser i kildeindtastning i april og oktober.

År	1998	1999	2000	2001	2001
Servicesager	286	190	160	138	123
Servicesager uden KIP	93	104	143	127	122
Udleverede studier	195158	465732	347720	2.403	1.460
Udleverede studier uden KIP	138	154	197	998	340
cd-rom'er	924	919	1295	1.382	1.149
Søgeopgaver	23	21	22	26	40
Søgeopgaver uden KIP	22	21	22	26	40

Udleverede studier**Anden formidling**

Antallet af servicesager uden KIP ser ud til at variere noget omkring et niveau lidt over 100. Antallet af udleverede studier er lavere end året før, men dog stadig højere end de nærmest foregående år. Tallet af udleverede studier i 2001 var meget ekstremt, det er så langt det største i arkivets historie.

Salget af cd-rommer fra kildeindtastningsprojektet er lavere end de to foregående år. Hvis de følgende år udviser tilsvarende fald vil de give anledning til bekymring om Kildeindtastningsprojektets overlevelse.

Mængden af søgeopgaver er noget højere end året før, men de fleste brugere søger selv i vores kataloger. Egentlige søgeopgaver forekommer kun undtagelsesvis.

Forskning

Anne Sofie Fink blev ansat som forsker i DDA 1. februar 2002.

DDA's akademiske personale bidrog med papers ved forskellige konferencer.

Ved ESSHC konferencen i Den Haag gav Nanna Floor Clausen og Hans Jørgen Marker paperet: „Making census data useful“.

Ved IASSIST Conferencen i Storrs, Connecticut gav Anne Sofie Fink paperet: „Towards a Danish research portal“.

Hans Jørgen Marker var førsteopponent på Frode Ulvunds disputats: „Ein studie af karrierer som insette i tvangsarbeid og straffeanstalten i Bergen ca. 1850-1975“ ved universitetet i Bergen.

Mogens Dam, Anne Sofie Fink og Heidi Wittendorff Sørensen udgav bogen „Statistik i praksis – genveje til data og eksempler på statistiske analyser“ på Forlaget Politiske Studier.

Andet**Cessda**

Cessda havde sit årsmøde i Ljubljana. Der blev optaget to nye medlemmer: Luxembourg og Rumænien.

English Summary

By BERNHARD HANSEN, DDA

In the previous issue of *Metode & Data* Dorte Gyrd-Hansen and Richard Jones presented results of a survey about the population's attitudes towards private health insurance. The authors demonstrated how answers to questionnaire items with the same substantial meaning varied a lot according to the context in which the questions were posed. In this issue Hanne-Pernille Stax deals with a related theme in the article „Language as a tool of measurement“. She demonstrates how questions with the same substantial meaning can result in very different patterns of answers solely due to small variations in language and design. Hanne-Pernille Stax does not only discuss problems associated with survey questionnaire design and analysis of survey data, she does also discuss several different methods, which may improve the design of questionnaires and the linguistic quality of the specific questions. These are methods, which everyone preoccupied with designing questionnaires and with analysing survey data should pay attention to. In Danish Data Archives we do not design survey questionnaires. However, we work a lot with survey data. In particular much time is spent on processing the data which researchers and data collectors have deposited with the Danish Data Archives. The purpose of processing is to document the data according to the guidelines of Danish Data Archives and to secure durability of data in system independent formats. The data processing procedures, as applied in the Danish Data Archives, are described in the second article „Data Preservation – Data processing in Danish Data Archives“ by Anne Sofie Fink Kjeldgaard et al. Conducting data analyses and processing social science data is one of two main activities in the Danish Data Archives. The other main activity concerns the organisation of the transcription of the Danish census records. Regarding the finished transcription of the 1801 census, Ole S. D. Hansen makes a call for careful analyses based on the data of this census, and he suggests several starting points

for such analyses. In the fourth article Nanna Floor Clausen presents some results of preliminary analyses of the living conditions of widows and widowers based on the complete dataset of the 1801 census. Among other things, Nanna Floor Clausen investigates distributional variations conditioned by the widows' and widowers' sex, age, occupation, and place of living. In addition to these articles the issue contains descriptions of fourteen datasets which have now been processed according to Danish Data Archives standards. The datasets are: DDA-1828 The Danish Leisure Study 1993; DDA-1579 Odense Municipality and the Citizens 1993; DDA-2085 Living Conditions of Physically Disabled Persons 1994-1995; DDA-2736 Odense Municipality and the Citizens 1995; DDA-4956 Young Second Generation Immigrants 1999; DDA-10416 The Citizens and the Laws 1997; DDA-10829 Traffic and Living Conditions in Kongens Enghave, Copenhagen, 2001; DDA-11062 Turnout at the Local Elections in Aarhus and Copenhagen 1997; DDA-11127 Self-experienced Discrimination 1998; DDA-11521 Quality of Life during the Crisis 1982-1994; DDA-14069 The Danish Party Members 2000; DDA-14701 The Danish Leisure Study 1998, Children; DDA-14702 The Danish Leisure Study 1998, Adults; DDA-13776 Work and Home Life 1950-1988. Moreover, this issue contains the annual report for the Danish Data Archives 2002. Finally, we ought to mention that in the future *Metode & Data* will only be published once a year. Substitute newsletters and special issues will be published as occasions require. Relevant news will also be published on our website more frequently. Although the next issue of *Metode & Data* will not be published before the autumn 2004 we still warmly welcome papers for publication (at a maximum of 10,000 words) containing methodological considerations and/or results of empirical studies with relevance for the social, health and historical sciences.

Dansk Data Arkiv (DDA):

DDA er landsdækkende arkiv for edb-læsbare forskningsmaterialer med relevans for en bred vifte af politiske, sociale, økonomiske, historiske og medicinske fagområder. Sammen med Rigsarkivet, landsarkiverne i København, Odense, Viborg og Åbenrå samt Erhvervsarkivet i Århus indgår DDA i gruppen Statens Arkiver, der ledes af rigsarkivar, dr. phil. Johan Peter Noack.

Medarbejdere i DDA:

Nanna Floor Clausen
Solveig Gjaldbæk
Anette Gulbæk
Bernard Hansen
John Gregers Hvidkjær Hansen
Birgitte Grønlund Jensen
Anne Sofie Fink Kjeldgaard
Henning Lauritsen
Hans Jørgen Marker
Steen Ousager
Anders Pedersen
Søren Priisholm
Jens Wagner
Jette Strand Wermuth
Jens Peter Nielsen