

Rekkefølgeeffekter i spørreundersøkelser blant ungdom

Resultater fra et split-ballot-eksperiment

Response-order effects in youth surveys

Results from a split-ballot experiment

Anders Bakken

Forsker II

NOVA, OsloMet – storbyuniversitetet

anders.bakken@oslomet.no

Patrick Lie Andersen

Forsker III

NOVA, OsloMet – storbyuniversitetet

palian@oslomet.no

Lars Roar Frøyland

Stipendiat

NOVA, OsloMet – storbyuniversitetet

lars.r.froyland@oslomet.no

Dawit Shawel Abebe

Førsteamanuensis

Institutt for sykepleie og helsefremmende arbeid, OsloMet – storbyuniversitetet

dawit.abebe@oslomet.no

SAMMENDRAG

Et viktig spørsmål i metodelitteraturen er om svaralternativenes rekkefølge har betydning for hva respondentene svarer i spørreskjemaundersøkelser. Imidlertid pekes det på at svarrekkefølgeeffekter ikke er godt nok forstått, at få har undersøkt dette i skjermbaserte undersøkelser blant ungdom og at få studier undersøker om slike effekter varierer med

ungdoms bakgrunnskjenne­tegn. I denne artikkelen spør vi: Har rekkefølgen på svaralternativene betydning for omfanget av psykiske helseplager som måles – og varierer eventuelle rekkefølgeeffekter mellom respondenter med ulike kjennetegn, som kjønn, alder, skolekarakterer og hvor lang tid de bruker på undersøkelsen? For å besvare dette, benytter vi oss av et split-ballot eksperiment som var del av spørreundersøkelsen Ung i Oslo 2015, hvor vi analyserer svar fra 17 832 ungdommer i alderen 13–19 år. Analysene viser at når man måler psykiske helseplager blant ungdom i en elektronisk basert undersøkelse, har det en viss betydning hvilken rekkefølge svaralternativer presenteres i. Flere krysser av for svaralternativer som indikerer plager når disse står til venstre enn når de står til høyre. Resultatene gir dermed støtte til tidligere studier som har funnet lignende primacy-effekter. Målt ut fra effektstørrelsen er forskjellen i forekomst av psykiske symptomplager mellom de to eksperimentgruppene relativt begrenset, noe som tyder på at det først og fremst er andre forhold som påvirker hvordan ungdom svarer. Likevel er effekten av betydning, siden det er 3,3 prosentpoeng flere som rapporterer om mange symptomplager når plaget-alternativene står til venstre enn når de står til høyre. Analysene viser for øvrig at rekkefølgeeffekten er uavhengig av ungdommenes kjønn, alder og hvor lang tid de brukte på spørreskjemaet. Derimot varierer rekkefølgeeffekten tydelig etter ungdommenes karakterer. Effekten er størst blant de med dårlige karakterer, mens det ikke er noen effekt i det hele tatt blant de med aller best karakterer.

Nøkkelord

Kvantitative metoder, spørreundersøkelser, svarrekkefølge, ungdom, depressive symptomer

ABSTRACT

An important question in the method literature is whether the response-order in surveys affects respondents' answers. However, such response-order effects are not well understood. Few studies have examined results from screen-based surveys among adolescents, including whether such effects depend on the youth's background characteristics. In this article, we thus ask: Do the response-order effects have impacts on the observed prevalence of mental health problems (i.e., depressive symptoms) among adolescents – and, are such response-order effects dependent on adolescents' characteristics such as sex, age, academic performance and total survey response time? To answer these research questions, we utilize a split-ballot experiment that was part of the survey Young in Oslo 2015, in which we analyzed responses from 17,832 adolescents aged 13–19 years. The analyses show that the response order is of some importance when measuring mental health problems among adolescents in an electronic survey. Respondents more frequently report symptoms when the options indicating problems options are left-aligned, than right-aligned. The results thus provide support for previous studies that found similar primacy effects. Measured by the effect size, the difference in the occurrence of depressive symptoms between the two experimental groups is relatively limited, which indicates that it is primarily other factors that affect how young people respond. Nevertheless, the effect is important since the prevalence of

depressive symptoms is 3.3 percentage points higher if the bothered options are left-aligned, than if they are right-aligned. The analyses also show that the order effect is independent of gender, age and the amount of time spent to complete the questionnaire. On the other hand, this effect clearly varies according to the youth's academic performance. The effect is strongest among those with low school grades, while there is no effect at all among those with the highest grades.

Keywords

Quantitative methods, survey, response order, youth, depressive symptoms

INTRODUKSJON

De siste 20 årene har ungdomsundersøkelser i flere vestlige land påvist en økning i selvrapporterte psykiske helseplager, særlig blant jenter (Bakken, 2017; Bor, Dean, Najman & Hayatbakhsh, 2014; Collishaw, 2015; von Soest & Wichstrøm, 2014). Et standardinstrument som har vært mye brukt er Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth & Covi, 1974). Instrumentet skal fange opp tegn på depresjon og angst gjennom å be respondenter om å rapportere hvor mye plaget de er av ulike symptomer. Svaralternativene er en firedelt skala som varierer fra «ikke plaget i det hele tatt» til «veldig mye plaget». Instrumentet har vært brukt i flere norske ungdomsundersøkelser, som for eksempel Ungdata, Ung i Norge og Ung i Oslo, og originalversjonen har blitt validert både i populasjonsstudier og i undersøkelser med kliniske pasienter (Siqueland, Moum & Leiknes, 2016).

I den opprinnelige versjonen av instrumentet er svaralternativet som indikerer at respondentene ikke er plaget, plassert på venstre side i spørreskjemaet. Dette vil normalt være det første svaralternativet respondentene leser. «Veldig mye plaget» er plassert lengst til høyre. I undersøkelsen Ung i Oslo, som er gjennomført flere ganger siden 1996 (Andersen & Bakken, 2015; Bakken, 2007; Øia, 2007, 2012), har svaralternativene derimot vært presentert for respondentene i motsatt rekkefølge. Tidligere forskning har vist at rekkefølgen av svaralternativer i spørreskjemaer påvirker resultatene, og at spørreskjemaer som utfylles av respondentene selv oftere får svar på alternativene som står til venstre (f.eks. Bishop & Smith, 2001; Chan, 1991; Friedman, Herskovitz & Pollack, 1994; Fuchs, 2005). Vi spør oss i den forbindelse om rekkefølgen på svaralternativene i Ung i Oslo-undersøkelsen kan ha bidratt til å overestimere nivået av depressive plager blant ungdommer i Oslo, og dermed også ha innvirkning på vår oppfatning av dagens ungdomsgenerasjon. Analysene har samtidig allmenn interesse når det gjelder spørreskjemametodikk, da rekkefølgeeffektene vi analyserer antakelig også vil gjelde for andre typer spørsmål enn de som måles her. Studien belyser med dette spørreundersøkelsers reliabilitet når det gjelder måling av psykiske helseplager blant ungdom.

Selv om rekkefølgeeffekter i spørreundersøkelser har blitt dokumentert i flere undersøkelser, er det fortsatt ubesvarte spørsmål. Blant annet er måten rekkefølgeeffektene virker på, ikke godt nok forstått (Schwarz, Hippler & Noelle-Neumann, 1992). Dessuten har et relativt begrenset antall studier undersøkt slike effekter i elektroniske spørreundersøkelser

blant ungdom (men se bl.a. Chang & Krosnick, 2010; Sanjeev & Balyan, 2014). Det er også ønskelig med flere analyser som undersøker om rekkefølgeeffekter varierer blant ungdom med ulike kjennetegn, for eksempel etter kjønn og alder. Så langt vi har oversikt finnes det heller ingen norske studier om temaet. Derfor er det behov for forskning som kan si noe om hvor stor betydning slike effekter har i norsk sammenheng. At temaet psykiske helseplager blant ungdom har vært så høyt på dagsordenen de siste årene, gjør analyser av akkurat denne tematikken særlig aktuell.

Hovedmålet med artikkelen er altså å undersøke om svaralternativenes rekkefølge er av betydning for hvor stort omfang av psykiske helseplager man måler. Hypotesen er at man oppnår en høyere forekomst av plager når alternativet «veldig mye plaget» presenteres til venstre i spørreskjemaet enn når det plasseres til høyre. Hypotesen blir undersøkt gjennom et eksperimentelt elektronisk spørreskjemadesign der et stort utvalg tenåringer deltok, hvor halvparten av deltakerne fikk spørreskjemaet presentert på én måte, mens den andre halvparten fikk spørreskjemaet presentert med alternativene i motsatt rekkefølge (se figur 1). Vi spør: Har rekkefølgen på svaralternativene betydning for omfanget av psykiske helseplager som måles – og varierer eventuelle rekkefølgeeffekter mellom respondenter med ulike kjennetegn, som kjønn, alder, skolekarakterer og hvor lang tid de bruker på undersøkelsen?

Vi vil først gå igjennom tidligere empiriske studier av rekkefølgeeffekter i spørreskjemaundersøkelser, før vi presenterer metode, data og resultater fra vår undersøkelse. Et viktig funn i denne studien er at vi finner rekkefølgeeffekter, og i avslutningen diskuterer vi dette i lys av noen forklaringsmodeller for hvorfor man finner slike effekter, og særlig hvorfor de spesielt kan gjøre seg gjeldende blant ungdom. Avslutningsvis løfter vi fram noen implikasjoner av vår studie.

TIDLIGERE STUDIER AV REKKEFØLGEEFFEKTER

I metodelitteraturen er det lang tradisjon for å undersøke om måten spørreskjemaer utformes på har betydning for resultatene fra undersøkelsen. Allerede på 1920-tallet ble mulige rekkefølgeeffekter studert, og det ble tidlig påvist at rekkefølgen har betydning (Mathews, 1927, 1929). I senere studier har en også vært opptatt av å undersøke om rekkefølgeeffekter varierer etter blant annet respondentenes alder, kjønn og kognitive ferdigheter (Krosnick & Alwin, 1987), men også om de varierer ut fra spørreskjemaets utforming, tematikk og antall spørsmål i spørsmålsbatterier (Tourangeau, Couper & Conrad, 2004). Forskningen har i hovedsak pekt på to ulike typer rekkefølgeeffekter. Den ene omtales som «primacy»-effekt, det vil si en tendens til å velge svaralternativer som blir presentert på venstre side av spørreskjemaet (f.eks. Bishop & Smith, 2001; Chan, 1991; Friedman, Herskovitz & Pollack 1994; Fuchs, 2005). Den andre effekten, «recency»-effekten, er tendensen til det motsatte: å velge svaralternativer lengst til høyre eller nederst i en rekke (f.eks. Bishop & Smith, 2001; Holbrook, Krosnick, Moore & Tourangeau, 2007; Schuman & Presser, 1996). Studier viser at «primacy»-effekten er mest framtrædende i spørreundersøkelser hvor respondentene ser selve spørreskjemaet, mens «recency»-effekten er mest framtrædende i intervjubaserte spørreundersøkelser hvor spørsmålene leses opp for respondenten, men også i tilfeller der svaralternativene er dikotome (McClendon, 1986).

Det er gjennomført flere studier av rekkefølgeeffekter blant barn og unge, og disse konkluderer som oftest med at respondenter har en tendens til å velge svaralternativer til venstre eller øverst i spørreskjemaet. I studien «The Youth Tobacco Survey» gjennomførte for eksempel O'Halloran og kolleger (2014) et eksperiment i to amerikanske delstater, der de fant signifikante primacy-effekter for omtrent en tredel av alle spørsmålene de undersøkte. Chan (1991) fant også signifikante primacy-effekter på en holdningsskala blant 15–17-åringer i Taiwan. Fuchs (2005) gjennomførte to undersøkelser blant tyske barn og ungdommer hvor han avdekket primacy-effekter for de yngste deltakerne (13–15 år), men ingen effekter for de eldste deltakerne. Derimot var rekkefølgeeffektene mer markante blant ungdommer med svake skoleprestasjoner enn blant elevene som gjorde det godt på skolen.

Det finnes flere studier som i nyere tid har testet om rekkefølgeeffekter gjør seg gjeldende når undersøkelsene blir gjennomført elektronisk. Blant annet gjennomførte Chang og Krosnick (2010) en studie blant amerikanske studenter, hvor rekkefølgeeffekter i en elektronisk spørreundersøkelse ble sammenliknet med effekter fra en intervjuundersøkelse. De fant en signifikant tendens til at ungdommene valgte det siste svaralternativet når spørsmålene og alternativene ble lest opp, mens de ikke kunne dokumentere rekkefølgeeffekter når spørsmålene ble stilt elektronisk. I en elektronisk spørreundersøkelse blant indiske studenter fant derimot Sanjeev og Balyan (2014) en tydelig primacy-effekt.

Primacy-effekter har også vært undersøkt med dataprogrammer som følger respondentenes øyebevegelser. Analysene viste tydelig at deltakerne i studien fokuserte mer på alternativene til venstre enn alternativene til høyre (Galesic, Tourangeau, Couper & Conrad, 2008). Det finnes også en studie som har undersøkt hvorvidt det er forskjeller i svaratferd mellom respondenter med høy og lav utdanning, og hvorvidt kombinasjonen av tidsbruk på å besvare skjemaet og utdanningsnivå forsterker tendensen til primacy-effekter (Malhotra, 2008). Analysene viste at blant respondenter med lav utdanning var det en tendens til å velge det første alternativet dersom man besvarte skjemaet raskt, mens tendensen forsvant dersom respondentene tok seg bedre tid til å besvare skjemaet. Blant respondenter med høy utdanning fant han derimot ingen primacy-effekter, uavhengig av hvor lang tid de brukte på å besvare undersøkelsen.

Målet med artikkelen er for det første å undersøke om det finnes svarrekkefølgeeffekter når vi måler psykiske helseplager blant ungdom gjennom et elektronisk, nettbasert spørreskjema. For det andre ønsker vi å undersøke om eventuelle svarrekkefølgeeffekter slår ulikt ut blant ungdom med ulike kjennetegn. Basert på tidligere studier forventer vi at rekkefølgeeffektene er sterkere desto yngre respondentene er, desto svakere skolekarakterer de har og desto kortere tid de bruker på utfyllingen av spørreskjemaet. Vi vil også undersøke kjønnsforskjeller i svarrekkefølgeeffekter. Siden gutter i gjennomsnitt modnes senere enn jenter (Grøgaard & Arnesen, 2016), er antakelsen at rekkefølgeeffekter er mer uttalt blant gutter enn blant jenter.

METODE OG DATA

Forskningsspørsmålene undersøkes gjennom et såkalt split-ballot eksperiment, der deltakerne i ungdomsundersøkelsen Ung i Oslo 2015 besvarte seks spørsmål fra det validerte spørsmålsbatteriet Hopkins Symptom Checklist (HSCL; Derogatis et al., 1974).¹ En tilfeldig halvpart fikk spørsmålene presentert med svaralternativene ikke plaget og lite plaget lengst til venstre, og alternativene ganske mye og veldig mye plaget lengst til høyre. Den andre halvparten fikk svaralternativene i motsatt rekkefølge, med veldig mye plaget lengst til venstre. Se figur 1 for et bilde av de to skjemaene.

Opprinnelig versjon av HSCL

Har du i løpet av den siste uka vært plaget av noe av dette:	Ikke plaget i det hele tatt	Lite plaget	Ganske mye plaget	Veldig mye plaget
Følt at alt er et slit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hatt søvnproblemer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg ulykkelig, trist eller deprimert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt håpløshet med tanke på framtida	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg stiv eller anspent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bekymret deg for mye om ting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg ensom	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Ung i Oslo-versjon av HSCL

Har du i løpet av den siste uka vært plaget av noe av dette:	Veldig mye plaget	Ganske mye plaget	Lite plaget	Ikke plaget i det hele tatt
Følt at alt er et slit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hatt søvnproblemer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg ulykkelig, trist eller deprimert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt håpløshet med tanke på framtida	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg stiv eller anspent	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bekymret deg for mye om ting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Følt deg ensom	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figur 1. To versjoner av HSCL med ulik svarrekkefølge.

1. Bakgrunnen for at eksperimentet ble gjennomført var at Ung i Oslo-studien ble del av en større ungdomsundersøkelse, Ungdata, som gjennomføres i kommuner over hele landet (se Bakken, 2017). I Ungdata har instrumentet HSCL siden 2011 blitt brukt slik det originalt ble satt opp av Derogatis et al. (1974), mens Ung i Oslo-undersøkelsen har hatt svaralternativene i motsatt rekkefølge. For både å kunne sammenlikne med andre Ungdata-undersøkelser og med de tidligere Ung i Oslo-undersøkelsene, ble det i Ung i Oslo 2015 besluttet å ha med begge varianter av svarrekkefølgene – men da splittet i to tilfeldige deler av utvalget.

For å undersøke mulige effekter av svarrekkefølgen, er man avhengig av at det er helt tilfeldig hvem som får den ene og den andre varianten. Dette sikret vi på følgende måte: Ved innlogging fikk hver respondent tildelt et tilfeldig nisifret respondentnummer. Nummeret ble gitt anonymt, og det var unikt for hver respondent. Når respondentene kom til den delen av spørreskjemaet der de ble bedt om å besvare HSCL, ble alle med partallsnummer loset videre til den validerte varianten av instrumentet fra Derogatis et al. (1974), med svaralternativet «ikke plaget i det hele tatt» lengst til venstre, mens alle med oddetallsnummer fikk varianten fra Ung i Oslo med svaralternativet «veldig mye plaget» lengst til venstre.

Analysene viste at eksperimentet var vellykket. Vi fant ingen signifikante forskjeller mellom gruppene som fikk de to ulike variantene av HSCL, verken når det gjaldt kjønn, alder, sosioøkonomisk bakgrunn eller svartid. Dette innebærer at forskjeller i svarfordeling mellom de to HSCL-variantene utelukkende kan tilbakeføres til svaralternativenes rekkefølge, og dermed representere rekkefølgeeffekter.

Formålet med Ung i Oslo var å gi en bred oversikt over ulike sider ved osloundoms levekår og oppvekstsituasjon. Undersøkelsen i 2015 var nettbasert, og ble gjennomført i skoletiden blant mer enn 24 000 elever fra de aller fleste ungdomsskoler og videregående skoler i byen (se Andersen & Bakken 2015 for detaljer). Elevene fikk én skoletime til å fylle ut skjemaet, og spørsmålet om psykiske helseplager kom omtrent midtveis i undersøkelsen. Ungdommene ble der bedt om å oppgi i hvor stor grad de har vært plaget i løpet av den siste uka av følgende seks symptomer: «Følt at alt er et slit», «hatt søvnproblemer», «følt deg ulykkelig, trist eller deprimert», «følt håpløshet med tanke på framtida», «følt deg stiv eller anspent» og «bekymret deg for mye om ting». Svaralternativene «ikke plaget i det hele tatt», «lite plaget», «ganske mye plaget», og «veldig mye plaget» ble kodet fra 1 til 4. Cronbach's Alpha for HSCL er 0,88, og er likt uansett hvilken retning svaralternativene ble presentert i. Basert på svarene beregnet vi den enkeltes gjennomsnittsskåre. I tillegg benyttet vi et dikotomt mål, som skiller mellom de med et spesielt høyt nivå av plager og resten. Grensen for mye plaget ble satt ved gjennomsnittsskåren 3,0 poeng, det vil si der hvor respondentene i «gjennomsnitt» har krysset av for at de er «ganske mye plaget» av alle de seks symptomene.

Studien var først og fremst rettet mot skoleungdom under 20 år. 320 respondenter som var eldre enn 19 år deltok likevel i undersøkelsen, så disse er tatt ut av analysene. Vi ekskluderte også 2143 respondenter som har manglende svar (missing) på minst ett av spørsmålene som brukes i regresjonsanalysene, og 3086 respondenter som har brukt svært kort (<12 minutter) eller svært lang tid (>69 minutter) på undersøkelsen. De med ekstremt kort svartid ($N=15$) som blir ekskludert har som regel hoppet over svært mange spørsmål, har falt fra tidlig i undersøkelsen eller har brukt usannsynlig kort tid på hvert spørsmål. Hovedgrunnen til at noen har ekstremt lang svartid ($n=3071$) er at reelt sluttidspunkt ikke ble registrert, altså av datatekniske årsaker. Det analytiske utvalget består av 17 832 respondenter. Dette utvalget skiller seg ikke fra de som ble ekskludert, verken på måten de har svart på HSCL eller hvilken versjon av HSCL de fikk.

Kjønn er målt som en dikotom variabel. Som indikator på alder bruker vi informasjon om hvilket klassetrinn eleven gikk på – fra 8. trinn til Vg3. Karakterer ble målt gjennom spørsmål om hvilken karakter eleven fikk ved siste karakteroppgjør i fagene norsk skriftlig, engelsk skriftlig og matematikk (skala 1–6). Basert på svarene beregnet vi den enkeltes gjennomsnittskarakter (gjennomsnitt 4,05, standardavvik 0,84). Tidligere studier har vist

at selvrappporterte karakterer er et godt mål på faktiske skoleprestasjoner (Kuncel, Credé & Thomas, 2005). Til sist inkluderte vi informasjon om hvor lang tid ungdommene brukte på å besvare undersøkelsen. Dette ble målt ved differansen mellom tidspunktet for innlogging og det registrerte tidspunktet for utlogging. Respondentene ble delt inn i fem like store grupper, med svartider på henholdsvis 12–23 minutter, 24–28 minutter, 29–32 minutter, 33–38 minutter og 39–69 minutter. Gjennomsnittstiden er beregnet til 32 minutter, standardavvik 8,7.

ANALYSER

For å sjekke om rekkefølgen på svaralternativene påvirker svarene på HSCL-instrumentet, undersøkte vi om det var forskjeller i gjennomsnittlige skårer mellom de som fikk de to ulike variantene. Vi beregnet gjennomsnittsskårer for hvert enkelt ledd og på tvers av alle seks ledd. Vi undersøkte gjennom krysstabellanalyse om det var forskjeller mellom de to gruppene i andelen med høy skåre på HSCL. Signifikansnivåer ble undersøkt gjennom t-tester og kjikvadrat-test. Effektstørrelser ble undersøkt gjennom Cohens *d*. Hvorvidt rekkefølgeeffekter varierer med ulike kjennetegn ved ungdommene, ble undersøkt gjennom regresjonsanalyser med den dikotome varianten av HSCL som avhengig variabel, hvor vi skilte mellom høye skårer ($HSCL > 3$) og lavere skårer ($HSCL \leq 3$). Når man anvender OLS regresjon på denne typen kategorisk variabel med to verdier, betegnes analysene som lineære sannsynlighetsmodeller. Ved å bruke slike modeller får man koeffisienter som er lettere å fortolke (som endringer i prosentpoeng) og økt sammenlignbarhet mellom modeller og undergrupper enn hva man oppnår med ikke-lineære regresjonsteknikker (Hellevik, 2009; Mood, 2010).² I regresjonsanalysene inkluderte vi fire samspillsledd mellom HSCL-varianten på den ene siden og kjønn, klasstrinn, karakterer og svartid på den andre.

RESULTATER

Tabell 1 viser at det er en effekt av svarrekkefølge. Ungdommene som fikk spørreskjema-varianten med «veldig mye plaget» til venstre, skårer i gjennomsnitt høyere på HSCL enn ungdommene som fikk varianten med «ikke plaget» til venstre. Dette gjelder både for instrumentet som helhet og for alle enkeltledd, selv om størrelsen på forskjellen varierer noe mellom de ulike leddene. Gjennomsnittsforskjellen på samlemålet er 0,11, noe som gir en Cohens *d* på 0,14. Dette karakteriseres vanligvis som en liten effekt (Cohen, 1988). At rekkefølgen svaralternativene presenteres på likevel har en viss betydning, ser vi når vi sammenlikner hvor stor andel som har et høyt nivå på symptomplagene. Blant ungdommene som fikk varianten med «mye plaget» til venstre, er det 18,5 prosent som skårer høyt. Dette er 3,3 prosentpoeng høyere enn blant de som fikk varianten med «ikke plaget» til venstre.

2. Alle analysene er også gjort med logistisk regresjon, noe som gir tilsvarende resultater.

Tabell 1. Skåre på enkeltledd og samlemål på Hopkins Symptom Checklist blant ungdom som fikk to ulike varianter av rekkefølgen på svaralternativene. Gjennomsnitt, prosent og standardavvik

	HSCL original (Ikke plaget til venstre)			HSCL Ung i Oslo (Veldig mye plaget til venstre)			Diff.	<i>d</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>			
HSCL total (1–4)	2,04	0,79	8727	2,15	0,80	9105	0,11	0,14	<,001
Følt at alt er et slit (1–4)	2,23	1,01	8672	2,38	1,01	9057	0,15	0,15	<,001
Hatt søvnproblemer (1–4)	2,07	0,98	8682	2,16	0,98	9066	0,09	0,09	<,001
Følt deg ulykkelig, trist eller deprimert (1–4)	1,87	0,97	8673	1,96	0,99	9054	0,09	0,09	<,001
Følt håpløshet med tanke på framtida (1–4)	1,91	1,01	8670	1,99	1,03	9047	0,08	0,08	<,001
Følt deg stiv eller anspent (1–4)	1,83	0,94	8652	1,97	0,97	9019	0,13	0,14	<,001
Bekymret deg for mye om ting (1–4)	2,30	1,05	8666	2,42	1,05	9046	0,12	0,11	<,001
Prosentandel	%		<i>n</i>	%		<i>n</i>	Diff.	χ^2	<i>p</i>
Mange symptomplager (>3,0)	15,2		8727	18,5		9105	3,3	33,86	<,001
Ikke besvart HSCL-spørsmålene	9,5		11 378	9,3		12 003	-0,2	0,23	>,05

Note. *M* = gjennomsnitt; *SD* = standardavvik; *n* = antall svar; *d* = Cohens *d*.

Antall respondenter som har skåre på samlemålet er høyere enn på hvert enkeltspørsmål. Dette skyldes at noen respondenter har manglende svar på enkelte ledd i HSCL, men likevel kan få totalskåre basert færre enn seks ledd. 88% av de som brukes i analysene har imidlertid svart på samtlige ledd.

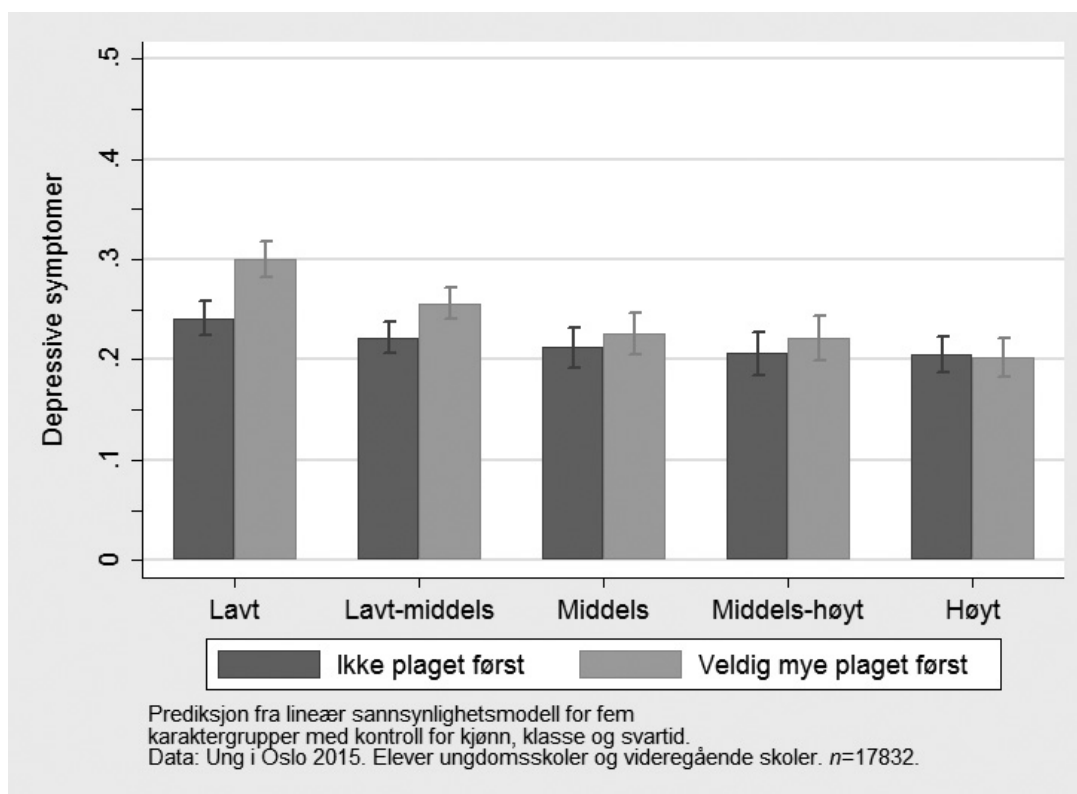
Tabell 2 viser regresjonsanalyser av sannsynligheten for å rapportere et høyt nivå av symptomplager, der vi undersøker mulige samspillseffekter mellom variant av spørreskjemaet og faktorene kjønn, klassetrinn, skolekarakterer og svartid. I modell 1–4 inkluderte vi én og en variabel, mens alle variablene ble tatt med i modell 5. Resultatene viser at den generelle rekkefølgeeffekten som vi tidligere dokumenterte, er mer eller mindre av samme størrelse blant gutter og jenter, blant ungdom på ulike klassetrinn og blant ungdom som brukte ulik tid på spørreskjemaet. Derimot finner vi et nokså tydelig samspill mellom karakterer og spørreskjemavariant. Dette er illustrert i figur 2, som viser at rekkefølgeeffekten er størst blant ungdom som får de dårligste karakterene. Jo høyere karaktersnitt ungdom har, dess mindre forskjell blir det mellom de som fikk de to variantene av HSCL. Blant de med de aller beste karakterene har derimot ikke rekkefølgen på svaralternativene noe å si i det hele tatt.

Tabell 2. Regresjonsanalyse av sannsynligheten for å ha høy skåre på psykiske symptomer (HSCL>3) etter svarrekkefølge og ulike kjennetegn ved ungdom. Lineære sannsynlighetsmodeller ($n = 17\,832$).

	Modell 1 Kjønn <i>b</i> (SE)	Modell 2 Klassetrinn <i>b</i> (SE)	Modell 3 Karakterer <i>b</i> (SE)	Modell 4 Svartid <i>b</i> (SE)	Modell 5 Alle variabler <i>b</i> (SE)
Svarrekkefølge (plaget lengst til venstre=1)	0,037*** (0,008)	0,025** (0,010)	0,119*** (0,021)	0,047*** (0,010)	0,134*** (0,025)
Kjønn (jente=1)	0,156*** (0,008)				0,156*** (0,008)
Rekkefølge*kjønn	-0,009 (0,011)				-0,006 (0,011)
Klassetrinn (0–5)		0,026*** (0,002)			0,023*** (0,002)
Rekkefølge*klassetrinn		0,003 (0,003)			0,002 (0,003)
Karakterer (1–6)			-0,009 (0,005)		-0,018*** (0,005)
Rekkefølge*karakter			-0,028*** (0,007)		-0,029*** (0,007)
Svartid (0–4)				-0,023*** (0,003)	-0,022*** (0,003)
Rekkefølge*svartid				-0,007 (0,004)	-0,007 (0,004)
Konstant	0,069*** (0,006)	0,093*** (0,007)	0,180*** (0,015)	0,199*** (0,007)	0,115*** (0,018)
Justert R^2	,042	,016	,005	,011	,069

Note. *b* = standardiserte koeffisienter; SE = standardfeil.

* $p < ,05$; ** $p < ,01$; *** $p < ,001$



Figur 2. Beregnet sannsynlighet for å rapportere høye psykiske symptomplager etter spørreskjemavariant og fem karakternivå. Konfidensintervall (95%).

Vi undersøkte også samspillsleddene i tabell 2 på tvers av kjønn. Analysene, som ikke er tatt med her, viser at ingen av samspillsleddene er signifikant forskjellige for gutter og jenter. Dette betyr at klassetrinn, skolekarakterer og svartid betyr det samme for hvordan gutter og jenter besvarer de to spørreskjemavariantene. Funnene i analysene er dermed gyldige for begge kjønn.

DISKUSJON

Analysene viser at resultatene i spørreundersøkelser påvirkes av i hvilken rekkefølge svaralternativer presenteres. At det er flere som krysser av for svaralternativer som indikerer plager når disse står til venstre enn når de står til høyre, viser at vi har med primacy-effekter å gjøre. Resultatene gir dermed støtte til tidligere studier som har funnet en tendens til at folk krysser av til venstre når de selv rapporterer gjennom et spørreskjema (f.eks. Bishop & Smith, 2001; Chan, 1991; Friedman et al., 1994; Fuchs, 2005). Målt ut fra effektstørrelsen, er forskjellen i forekomst av psykiske symptomplager mellom de to eksperimentgruppene relativt begrenset, noe som tyder på at det først og fremst er andre forhold som

påvirker hvordan ungdom svarer.³ Likevel vil vi argumentere for at effekten er av betydning, siden det er 3,3 prosentpoeng flere som rapporterer om mange symptomplager når plaget-alternativene står til venstre enn når det står til høyre. Analysene viser for øvrig at rekkefølgeeffekten er uavhengig av ungdommenes kjønn, alder og hvor lang tid de brukte på spørreskjemaet.

Rekkefølgeeffekten varierer derimot tydelig etter ungdommenes skolekarakterer, som har vært brukt som en indikator på kognitive evner. I så måte stemmer våre funn godt overens med funnene knyttet til alder og kognitive evner i Fuchs (2005). Effekten er størst blant de med dårlige karakterer, mens det ikke er noen effekt i det hele tatt blant de med aller best karakterer. Dette indikerer at «sensitiviteten» for svarrekkefølgen er avhengig av hvor flinke ungdommene er på skolen. Om vi legger til grunn at unge med svakere kognitive evner opprinnelig er mer plaget av depressive symptomer enn unge med sterke kognitive evner, vil det å stille alternativet «veldig mye plaget» til venstre dermed kunne bidra til å overdrive denne forskjellen.

Hvorfor finner man rekkefølgeeffekter? Bishop og Smith (2001) presenterer enkelte teorier som har vært fremsatt om hvorfor det finnes primacy-effekter i spørreundersøkelser. En av forklaringene er at de oppstår når man stiller for vanskelige spørsmål eller stiller spørsmål som respondentene ikke har noen klar mening om (f.eks. Payne, 1951; Rugg & Cantril, 1944). Dette, som i den engelskspråklige litteraturen har blitt omtalt som «opinion crystallization» («meningsfiksering»), er etter vår vurdering trolig ikke den viktigste forklaringen på de mønstrene vi fant. Temaet psykisk helse er noe de fleste unge i dag kjenner til, og følelsene som kartlegges er noe som de aller fleste i større eller mindre grad har erfaringer med. I en nylig gjennomført intervjuundersøkelse der ungdom ble bedt om å kommentere på det samme instrumentet som er brukt i denne undersøkelsen, kom det fram at spørsmålene var nokså enkle å forstå og at de opplevdes som relevante i ungdoms liv (Eriksen, Sletten, Bakken & von Soest, 2017).

En annen forklaring har vært omtalt som «the cognitive elaboration model» («den kognitive fordypningsmodellen»). Opprinnelig dreier den seg om hvordan det å bli eksponert for ulike ting i livet først kan kreve stor kognitiv kraft for å håndtere, mens det deretter blir lettere om man eksponeres flere ganger. I spørreskjemametodikken tilsier modellen at når respondentene eksponeres for et spørreskjema visuelt, vil de automatisk bruke mer tid på å lese de første alternativene enn de siste. Så lenge et av de første alternativene er noe respondentene, basert på tidligere livserfaringer, kan relatere seg til, vil dette kunne resultere i en primacy-effekt (Sudman, Bradburn & Schwarz, 1996). Vi oppfatter denne forklaringen som mer relevant, blant annet fordi den underbygges av studien vi innledningsvis viste til, der det ble vist at man ved skjermbaserte spørreundersøkelser har mest fokus på de første svaralternativene man ser (Galesic et al., 2008). Man kan anta at de første svaralternativene man ser, i større grad plasseres i bevisstheten, og at det dermed blir enklere for noen å sette kryss der (Schwarz et al., 1992).

3. Mens effektstørrelsen (Cohens d) av spørreskjemavariant på det kontinuerlige HSCL-målet er 0,14, tilsier forskjellen mellom gutter og jenter en d på 0,60, mens forskjellen mellom klassetrinnene på ungdomsskolen gir en d på 0,20. Ser man kun på elevene med lavest karaktersnitt, er effekten av spørreskjemavariant noe høyere ($d = 0,20$).

En variant av denne forklaringen kan være at respondenter som faktisk blir påvirket av rekkefølgen på svaralternativene er blant de som er mest tilbøyelige til å ta kognitive snarveier, det vil si de som rett og slett velger det enkleste og dermed sparer «kognitiv energi» ved å velge de alternativene man ser først. Dersom man oppfatter at noen av de første alternativene kan stemme sånn noenlunde, er det lettere å gå for disse enn å bruke tid på å vurdere de siste alternativene.

En slik forklaring kan trolig bidra til å kaste lys over hvorfor rekkefølgeeffekten er størst blant de med svakest karakterer. De med de svakeste skoleprestasjonene vil som regel være de som ville måtte bruke lengst tid på å lese både spørsmål og svaralternativer, og på den måten vil de også bruke mer «kognitiv energi» dersom de skal få med seg alle detaljer om spørsmålet. Elever med høyt karakternivå derimot, vil vanligvis være blant de som har de beste evnene til å konsentrere seg, samtidig som de i gjennomsnitt har bedre leseferdigheter og evner til å fortolke tekst.

Et argument som kan tale mot at kognitive evner er avgjørende for å forstå rekkefølgeeffektene, er at alder ikke samspiller med svarrekkefølgen. Siden mange modnes kognitivt i tenåringsfasen, skulle man forvente at svarrekkefølgen slår mest ut på svarene til de yngste. Analysene viser derimot at rekkefølgeeffektene er like utslagsgivende, uansett om ungdommene er i begynnelsen eller i slutten av tenårene. En mulig forklaring kan være at samtlige aldersgrupper i undersøkelsen har nådd et kognitivt nivå som gjør at spørsmålene vi fokuserer på her er helt greie å forholde seg til. Som nevnt er det å forstå meningen som ligger i både spørsmål og svaralternativer, for de aller fleste ikke spesielt vanskelig. Når vi likevel ser en samspillseffekt mellom svarrekkefølge og karakterer, kan det tenkes at dette målet bedre fanger opp en mindre gruppe av ungdom – uansett alder – som har vansker med å konsentrere seg, og med å vurdere og ta til seg informasjon, og at karakterer derfor gir utslag i analysene.

En annen mulig forklaring på karaktereffekten kan være at elever med gode karakterer er mer motivert for å svare, og å svare grundig, på spørreundersøkelsen. Tidligere forskning har vist at karakternivå er forbundet med elevenes motivasjon, for eksempel indre motivasjon, som innebærer en glede over å gjøre oppgaver i seg selv (Gottfried, 1985; Steinmayr & Spinath, 2009). Indre motivasjon har også blitt satt i sammenheng med hvordan respondenter fyller ut spørreskjema. I flere studier (Krosnick, 1991; Krosnick & Presser, 2010) kobles indre motivasjon til at respondentene i spørreundersøkelser tenker igjennom og svarer nøye, mens lav indre motivasjon i større grad vil lede til at respondentene ikke reflekterer så grundig og typisk velger første passende alternativ, og som dermed kan medføre primacy-effekter. Dette betyr i så fall at ikke bare kognitive faktorer alene er avgjørende, men også at graden av indre motivasjon for å svare på spørsmålene kan bidra til å forklare hvorfor spørreskjemaeffektene varierer med elevenes karakterer.

IMPLIKASJONER

At vi finner rekkefølgeeffekter når vi måler psykiske helseplager blant ungdom i et elektronisk spørreskjema, viser at konteksten for hvordan målinger foregår har betydning for resultatene. Analysene tyder på at rapporteringen av psykiske helseplager i Ung i Oslo-

undersøkelsene er noe høyere enn det man ville funnet dersom man fra starten av hadde valgt HSCL-varianten med svaralternativet «ikke plaget» til venstre. Det er likevel svært vanskelig å avgjøre hva som gir det beste og det mest riktige bildet. At det er en viss primacy-effekt – i den forstand at en viss prosent av de som besvarer spørreskjemaet har en tendens til å krysse av på svaralternativer til venstre – vil gjelde uansett hvordan respondentene får presentert alternativene. Dette betyr trolig at når man bruker den opprinnelige varianten av HSCL med alternativet «ikke plaget» til venstre, er det en fare for å underestimere omfanget av plager. Om det 'riktige nivået' ligger et sted midt imellom eller nærmere en av variantene, er trolig umulig å avgjøre helt presist. Vi vil likevel anbefale at man i undersøkelser der målet er å beskrive omfanget av psykiske plager blant ungdom, bruker varianten som har alternativet «ikke plaget» til venstre. Ikke bare fordi dette sikrer et bedre sammenlikningsgrunnlag på tvers av undersøkelser, men også fordi det etter vår mening er bedre å gi et noe mer konservativt anslag på omfanget av problemer enn å overdrive det. Særlig kan dette være viktig i en tid der det har vært veldig mye fokus på psykiske utfordringer ungdom står overfor.

HSCL brukes også i klinisk arbeid. Vår studie er en påminnelse om at rekkefølgeeffekter også kan oppstå i de tilfeller der instrumentet brukes for å diagnostisere enkeltindivider. I slike tilfeller vil rekkefølgeeffekter imidlertid være av en betydelig mer kritisk karakter for individet enn i studier av psykisk helse i den generelle befolkningen.

Samtidig er resultatene fra våre analyser også relevante for andre spørreundersøkelser, ikke bare de som anvender HSCL-spørsmål. Denne typen primacy-effekter kan selvsagt også gjøre seg gjeldende når ungdom blir stilt spørsmål om øvrige helseaspekter eller holdninger med lignende svaralternativer.

Selv om vi har avdekket noen tydelige rekkefølgeeffekter, har vår studie noen begrensninger. Ideelt sett skulle vi hatt et mer presist mål på kognitive ferdigheter enn karakterer, siden karakterer ikke skiller mellom motivasjonsfaktorer og hvor flinke elever er til å lese og prosessere tekst. For å komme nærmere hvilke mekanismer som skaper karaktereffektene, vil det for etterfølgende studier være interessant å inkludere mål som kan indikere hvor motivert ungdom er til å svare på spørreskjemaer. Det kan også være interessant å studere om karaktereffektene skyldes forhold ved elevene som ikke belyses i denne studien, som for eksempel sosioøkonomisk bakgrunn eller innvandrerbakgrunn.

KONKLUSJON

Vår studie viser viktigheten av å være bevisst metodiske konsekvenser av både hvordan man konstruerer surveyundersøkelser og hvordan man vurderer funn og resultater fra undersøkelsene. Undersøkelsen viser videre nytten av split-ballot eksperimenter for å avdekke rekkefølgeeffekter i spørreskjemaundersøkelser, og at målinger av psykiske helseplager synes å være sensitive overfor svarrekkefølge. Det finnes sjelden et fasitsvar for hva som vil gi det mest korrekte estimatet av fenomener man måler, med mindre fenomenet kan knyttes opp til mer objektive mål som faktiske handlinger. I så måte vil en studie hvor man sammenligner rekkefølgeeffekter for holdningsmål og handlingsmål være nyttig, og særlig hvis man i tillegg har mål på faktiske handlinger som ikke baserer seg på selvrappor-

tering. Vår anbefaling er uansett at man bør benytte spørreskjemaoppsett som gir konservative estimater framfor at estimatene blåses opp. Det at rekkefølgeeffekten varierer etter kognitive evner, kan potensielt ha store konsekvenser for tolkningen av resultater fra forskning. Hvis estimatene til enkelte grupper avviker fra estimatene for andre grupper, og dette egentlig er et metodisk artefakt som varierer med kognitive evner, kan det medføre at situasjonen til den gitte gruppen feilaktig tolkes som verre (eller bedre) enn den egentlig er, sammenlignet med andre. En måte å motvirke slike effekter på, kan være å bruke indekser der spørsmålene er formulert med ulikt fortegn eller ulike svarrekkefølger. Det viktigste er uansett at man, både som forsker og som leser, er klar over at subtile forskjeller i hvordan data samles inn, kan spille inn på resultatene.

FINANSIERING

Helsedirektoratet gir tilskudd til Ungdatasenteret ved NOVA for å gjennomføre Ungdata. Arbeidet med denne artikkelen er finansiert innenfor rammen av dette tilskuddet.

REFERANSER

- Andersen, P. L. & Bakken, A. (2015). *Ung i Oslo 2015*. NOVA Rapport 8/17. Oslo: Høgskolen i Oslo og Akershus.
- Bakken, A. (2007). Ungdommens klasseskiller. I Å. Strandbu & T. Øia (red.), *Ung i Norge: Skole, fritid og ungdomskultur* (s. 42–61). Oslo: Cappelen Akademisk Forlag.
- Bakken, A. (2017). *Ungdata 2017. Nasjonale resultater*. NOVA Rapport 10/17. Oslo: Høgskolen i Oslo og Akershus.
- Bishop, G. & Smith, A. (2001). Response-order effects and the early gallup split-ballots. *Public Opinion Quarterly*, 65(4), 479–505. doi: 10.1086/323575
- Bor, W., Dean, A. J., Najman, J. & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian and New Zealand Journal of Psychiatry*, 48(7). doi: 10.1177/0004867414533834
- Chan, J. C. (1991). Response-order effects in likert-type scales. *Educational and Psychological Measurement*, 51(3), 531–540. doi: 10.1177/0013164491513002
- Chang, L. & Krosnick, J. A. (2010). Comparing oral interviewing with self-administered computerized questionnaires. An experiment. *Public Opinion Quarterly*, 74(1), 154–167. doi: 10.1093/poq/nfp090
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. utg.). Hillsdale, NJ: Lawrence Erlbaum.
- Collishaw, S. (2015). Annual Research Review: Secular trends in child and adolescent mental health. *Journal of Child Psychology and Psychiatry*, 56(3), 370–393. doi: 10.1111/jcpp.12372
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H. & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Sciences*, 19(1), 1–15. doi: 10.1002/bs.3830190102
- Eriksen, I. M., Sletten, M. A., Bakken, A. & von Soest, T. (2017). *Stress og press blant ungdom*.

- Erfaringer, årsaker og utbredelse av psykiske helseplager*. NOVA Rapport 6/17. Oslo: Høgskolen i Oslo og Akershus.
- Friedman, H. H., Herskovitz, P. J. & Pollack, S. (1994). *The biasing effects of scale-checking styles on response to a likert scale*. Paper presentert på The Proceedings of the American Statistical Association Annual Conference: Survey Research Methods.
- Fuchs, M. (2005). Children and adolescents as respondents. Experiments on question order, response order, scale effects and the effect of numeric values associated with response options. *Journal of Official Statistics*, 21(4), 701–725.
- Galesic, M., Tourangeau, R., Couper, M. P. & Conrad, F. G. (2008). Eye-tracking data: New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, 72(5), 892–913. doi: 10.1093/poq/nfn059
- Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology*, 77(6), 631–645.
- Grøgaard, J. B. & Arnesen, C. Å. (2016). Kjønnforskjeller i skoleprestasjoner: Ulik modning? *Tidsskrift for ungdomsforskning*, 16(2), 42–68.
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43(1), 59–74. doi: 10.1007/s11135-007-9077-3
- Holbrook, A. L., Krosnick, J. A., Moore, D. & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally – The impact of question and respondent attributes. *Public Opinion Quarterly*, 71(3), 325–348. doi: 10.1093/poq/nfm024
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. doi: 10.1002/acp.2350050305
- Krosnick, J. A. & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51(2), 201–219. doi: 10.1086/269029
- Krosnick, J. A. & Presser, S. (2010). Questionnaire design. I J. D. Wright & P. V. Marsden (red.), *Handbook of Survey Research* (2. utg.). West Yorkshire, England: Emerald Group.
- Kuncel, N. R., Credé, M. & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75(1), 63–82. doi: 10.3102/00346543075001063
- Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly*, 72(5), 914–934. doi: 10.1093/poq/nfn050
- Mathews, C. O. (1927). The effect of position of printed response words upon children's answers to questions in two-response types of tests. *Journal of Educational Psychology*, 18(7), 445–457.
- Mathews, C. O. (1929). The effect of the order of printed response words on an interest questionnaire. *Journal of Educational Psychology*, 20(2), 128–134.
- McClendon, M. J. (1986). Response-order effects for dichotomous questions. *Social Science Quarterly*, 67(1), 205–211.
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26(1), 67–82. doi: 10.1093/esr/jcp006
- O'Halloran, A., Hu, S. S., Malarcher, A., McMillen, R., Valentine, N., Moore, M. A., ... Gerzoff, R. B. (2014). Response order effects in the Youth Tobacco Survey: Results of a split-ballot experiment. *Survey Practice*, 7(3).
- Payne, S. L. (1951). *The Art of Asking Questions*. Princeton, NJ: Princeton University Press.

- Rugg, D. & Cantril, H. (1944). The wording of questions in public opinion polls. I H. Cantril (red.), *Gauging Public Opinion* (s. 23–50). Princeton, NJ: Princeton University Press.
- Sanjeev, M. A. & Balyan, P. (2014). Response order effects in online surveys: An empirical investigation. *International Journal of Online Marketing*, 4(2), 28–44. doi: 10.4018/ijom.2014040103
- Schuman, H. & Presser, S. (1996). *Questions & Answers in Attitude Surveys*. Thousand Oaks, CA: Sage Publications.
- Schwarz, N., Hippler, H.-J. & Noelle-Neumann, E. (1992). A cognitive model of response–order effects in survey measurement. I N. Schwarz & S. Sudman (red.), *Context Effects in Social and Psychological Research* (s. 187–201). New York, NY: Springer New York.
- Siqveland, J., Moum, T. & Leiknes, K. A. (2016). *Måleegenskaper ved den norske versjonen av Symptom Checklist 90 Revidert (SCL-90-R)*. Oslo: Folkehelseinstituttet.
- Steinmayr, R. & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, 19(1), 80–90. doi: 10.1016/j.lindif.2008.05.004
- Sudman, S. N., Bradburn, M. & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass.
- Tourangeau, R., Couper, M. P. & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly*, 68(3), 368–393. doi: 10.1093/poq/nfh035
- von Soest, T. & Wichstrøm, L. (2014). Secular trends in depressive symptoms among Norwegian adolescents from 1992 to 2010. *Journal of Abnormal Child Psychology*, 42(3), 403–415. doi: 10.1007/s10802-013-9785-1
- Øia, T. (2007). *Ung i Oslo. Levekår og sosiale forskjeller*. NOVA Rapport 6/07. Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.
- Øia, T. (2012). *Ung i Oslo 2012. Nøkkeltall*. NOVA Notat 7/12. Oslo: Norsk institutt for forskning om oppvekst, velferd og aldring.