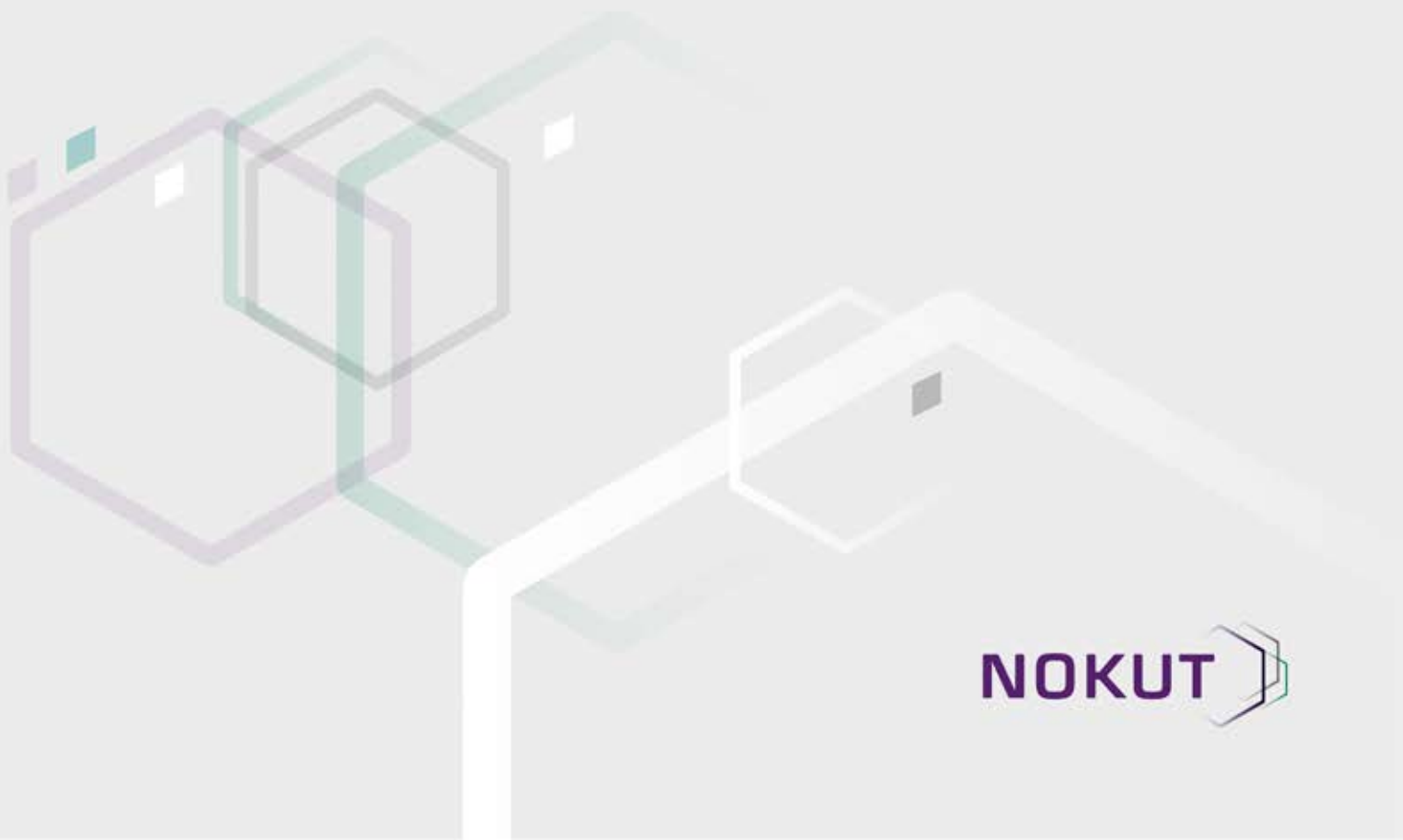


Studiebarometeret: Rapport 2–2015

Studiebarometeret 2014 – analyse av validitet og reliabilitet

Stephan Hamberg



NOKUT 

Sammendrag

Denne analysen tar sikte på å belyse hvorvidt resultatene fra Studiebarometeret i 2014 er representative for studentenes oppfatninger av utdanningskvalitet. Den sentrale delen av analysen er en frafallsanalyse, der vi undersøker om enkelte respondentgrupper er over- eller underrepresentert i forhold til mottagerpopulasjonen, og om det påvirker resultatene. Vi er med andre ord interesserte i å se om de studentene som svarer på undersøkelsen enten var de mest positive eller de mest negative, dvs. om det er systematiske skjevheter i datagrunnlaget. Forskning på studentevalueringer indikerer at det finnes en rekke forhold som kan lede til systematiske skjevheter, blant annet: kjønn, alder, studieerfaring og studiekultur. Vi ser på alle disse i denne analysen. Resultatene viser at resultatene fra undersøkelsen ikke har sterke systematiske skjevheter. Kvinner er mer tilbøyelige til å svare på undersøkelsen enn menn, men de svarer i liten grad forskjellig fra menn. Det er også kun små forskjeller når det gjelder alder. Yngre studenter er mer tilbøyelige til å besvare undersøkelsen, men forskjellen er ikke statistisk signifikant. Når det gjelder tilfredshet er de yngste og de eldste studentene de mest positive, men forskjellene er generelt meget små. Når det kommer til andre bakgrunnsvariabler slik som nasjonalitet, opptakskompetanse, og om studentene har svart på undersøkelsen tidlig eller sent er forskjellene også ubetydelige.

I tillegg til reliabilitetsanalysen har vi også gjort en validitetsanalyse, altså om undersøkelsen faktisk måler det den tar sikte på å måle. For å analysere validiteten av undersøkelsen har vi sett på tre faktorer. Først, hvor mange studenter har valgt å ikke svare på enkelt spørsmål. To, hvor mange studenter har svart «vet ikke» på enkelt spørsmål. Dersom mange studenter har valgt å ikke svare eller svart «vet ikke» på enkelt spørsmål kan det være en indikasjon på at spørsmålet og/eller svar alternativene er upresist formulert eller vanskelig å forstå. Til slutt har vi sett på studentenes kommentarer til spørreskjemaet. Analysen viser at de aller fleste spørsmålene har blitt forstått og besvart. Noen få spørsmål har en relativt høy andel «ikke svart» eller «vet ikke» svar, men analysen viser at det er naturlige årsaker til dette, og at det ikke er fordi spørsmålene eller svar alternativene er dårlig formulert. Det er samtidig viktig å påpeke at disse analysene er gjort på et makronivå og ikke på studieprogramnivå. Det er med andre ord ikke mulig å trekke konklusjoner om representativitet og validitet på studieprogramnivå basert på denne analysen.

Innhold

Sammendrag	1
Frafallsanalyse — hvem besvarte undersøkelsen og var de som svarte mer eller mindre tilfreds? ..	3
Kjønn.....	3
Tilfredshet	4
Alder	7
Andre bakgrunnsvariabler	9
Frafall underveis i skjemautfyllingen.....	9
Var de mest kritiske eller de mest positive studentene mest tilbøyelige til å besvare undersøkelsen?	9
Forskjell i tilfredshet på tidlig og sent svarende studenter	9
Sammenhengen mellom studiepoengproduksjon, svartilbøyelighet og alt-i-alt tilfredshet	10
Sammenheng mellom studentenes totalvurdering og svarprosent på programnivå	11
Hvordan fungerte spørreskjemaet? (Validitetsanalyse)	11
Analyse av frafall i spørreskjemaet	12
Analyse av «vet ikke», «ikke relevant» og «brukes ikke»	13
Hva sier studentene om spørreskjemaet?	17
Konklusjon	17

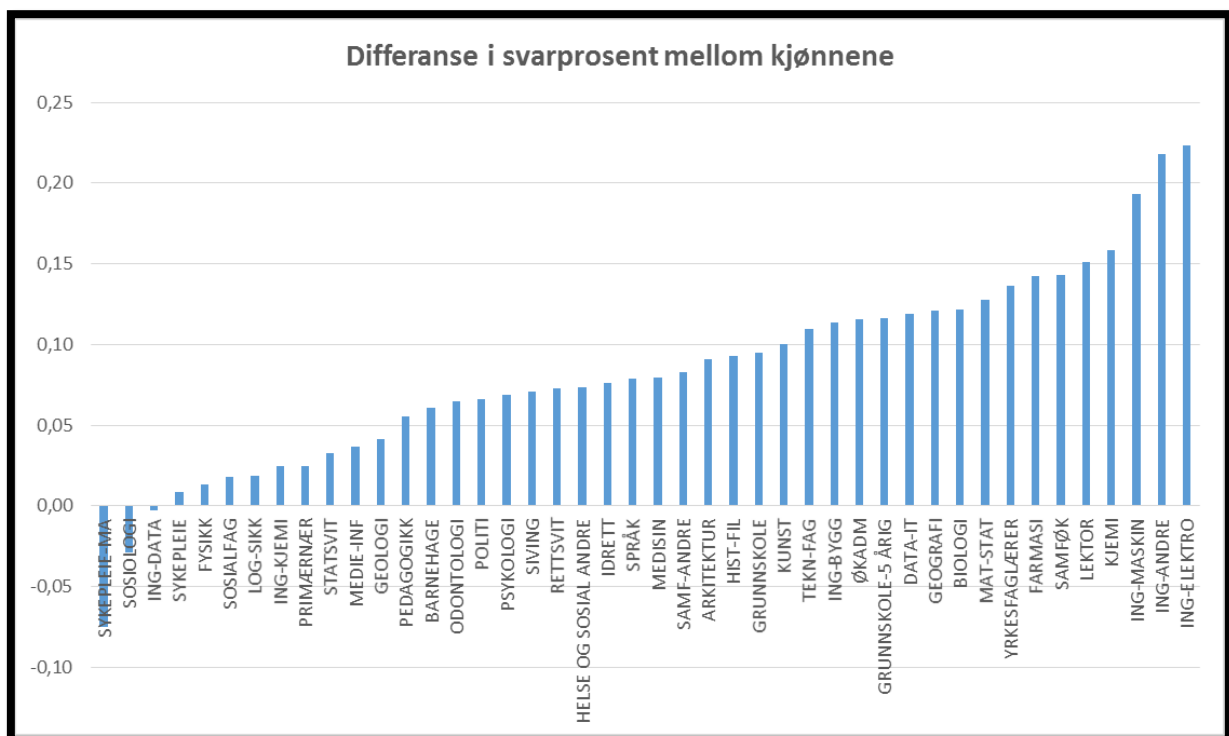
Frafallsanalyse — hvem besvarte undersøkelsen og var de som svarte mer eller mindre tilfreds?

Den nasjonale studentundersøkelsen ble gjennomført for andre gang høsten 2014. Undersøkelsen ble sendt til 58 834 andre og femte års studenter på bachelor- og mastergradsprogrammer i Norge. 24666 (42 %) av disse svarte på hele eller deler av undersøkelsen. I denne delen analyserer vi om enkelte grupper i større grad enn andre besvarte spørreskjemaet og om disse gruppene var mer eller mindre tilfreds med studieprogrammene sine. Vi har analysert frafallet basert på følgende indikatorer: kjønn, alder, statsborgerskap, utdanningslandbakgrunn, opptakskompetanse, studiepoengproduksjon, om studentene er innpassede eller ikke og om studentene er utreisende utvekslingsstudenter eller ikke.

Kjønn

I populasjonen utgjorde kvinner 58,3 % (34 276), og menn 41,7 % (24,486). Blant de som svarte på undersøkelsen utgjorde kvinner 62,4 % (15 367), og menn 37,6 % (9 255)¹. Det er altså en klar tendens til at kvinner i større grad enn menn besvarte spørreskjemaet. Denne forskjellen forekom også i forrige år, men da i noe større grad. Kvinnenes svarprosent var i år 45 % (opp fra 36 % i fjor), og mennenes svarprosent var 38 % (opp fra 27 prosent i fjor). Den totale svarprosenten var i år 42 %, mot 32 % i 2013. På forskjellige utdanningstyper ser vi også at kvinner er mer tilbøyelige til å svare på undersøkelsen enn menn. I figur 1 vises kjønnsdifferansen i svarandel for alle utdanningstypene.

Figur 1: Differanse i svarprosent mellom kvinner og menn for alle utdanningstyper



Figuren leses slik: Søylen viser differanse i svarprosent mellom kjønnene. Dersom menn har høyere svarprosent enn kvinner befinner søylen seg under streken som representerer 0. Der søylene er over streken har kvinner høyere svarprosent enn menn.

¹ Vi mangler data på kjønn for 72 mottagere og 44 respondenter.

Som vi ser fra figur 1 har kvinner høyere svarprosent enn menn i alle utdanningstyper bortsett fra sykepleie, sosiologi og ingeniør-data. Forskjellene i svarandel varierer fra 1 % til 22 %. De tre utdanningstypene med størst differanse er alle store mannsdominerte utdanningstyper. Ingeniør-maskin, Ingeniør-andre og Ingeniør-elektro har henholdsvis 75 %, 66% og 87 % mannlige respondenter. I disse tre utdanningstypene var svarprosenten blant kvinner omkring 20 % høyere enn menn.

At svarprosenten er høyere blant kvinner og menn er ikke nødvendigvis ett problem, men det kommer an på om kvinner også er systematisk mer eller mindre tilfreds enn menn. Vi undersøker dette nærmere i det neste avsnittet.

Tilfredshet

På *alt-i-alt tilfreds* spørsmålet svarte kvinner i snitt 4,06, og menn svarte i snitt 4,05 på en skala fra 1–5 (forskjellen er ikke signifikant, t-test, 0,05 signifikansnivå).² På andre spørsmål og indekser er det en del variasjon i forhold til om menn eller kvinner er mer positive (tabell 1). Kvinner er mer positive når det gjelder indeksene på engasjement, relevans, eksamen og læringsutbytte, mens menn er mer positive når det gjelder indeksene miljø og medvirkning (alle forskjeller er signifikante, t-test, 0,05 signifikansnivå). På spørsmål om faglig tilbakemelding var kvinner mer positive (3,28 mot 3,26), men forskjellen er ikke statistisk signifikant. På spørsmål om faglig oppfølging var menn mer positive (2,89 mot 2,84) og her var forskjellen statistisk signifikant, men forskjellen er så liten at den er substansielt ubetydelig.

Tabell 1: Forskjell i tilfredshet mellom kjønnene

Spørsmål	Menn		Kvinner		Differanse menn og kvinner
	Gjennomsnitt	Standard-avvik	Gjennomsnitt	Standard-Avvik	
alt-i-alt	4,05	0,99	4,06	0,99	-0,01
Tilbakemelding	3,26	1,10	3,28	1,13	-0,02
Oppfølging	2,89	1,20	2,84	1,21	0,05*
Indekser					
Undervisning	3,34	0,81	3,33	0,81	0,01
Miljø	3,80	0,68	3,75	0,68	0,05*
Medvirkning	3,27	0,92	3,22	0,91	0,05*
Engasjement	4,00	0,73	4,12	0,70	-0,12*
Relevans	4,16	0,76	4,31	0,72	-0,14*
Eksamen	3,73	0,74	3,89	0,72	-0,16*
Læringsutbytte	3,64	0,64	3,69	0,61	-0,05*

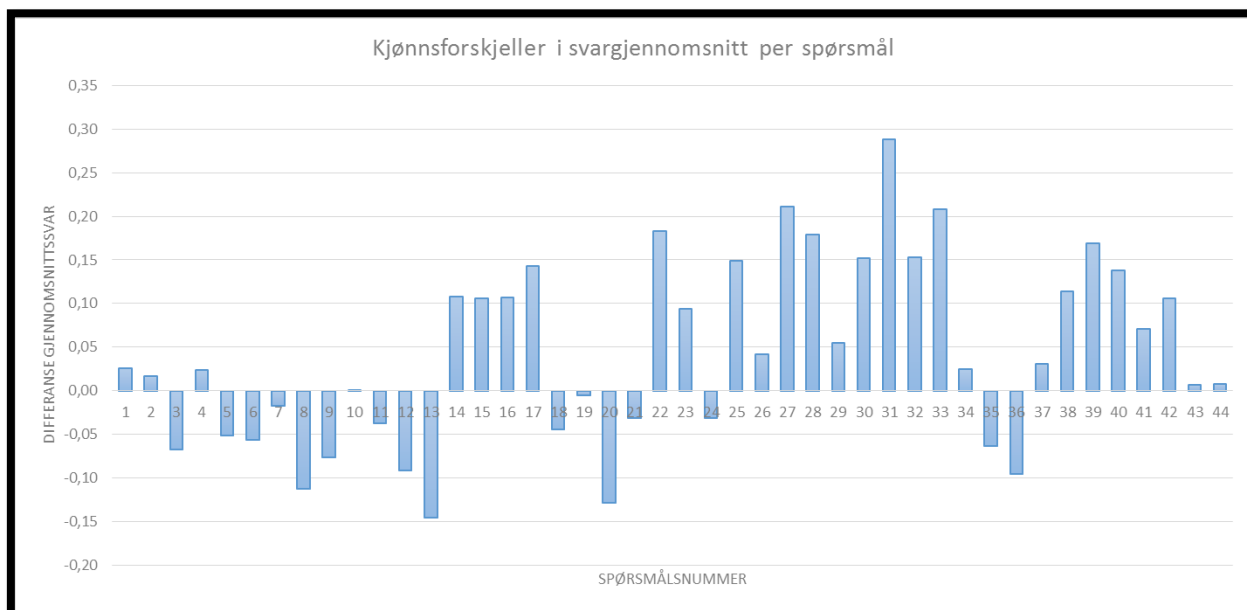
* Signifikant, t-test, 0,01 signifikansnivå. Kvinner er mer positive der differansen er negativ.

Som vi ser fra tabell 1 er det ikke noe mønster i om kvinner eller menn er mer positive i sine svar når vi er på indeksene, samt alt-i-alt, faglig tilbakemelding og faglig oppfølging spørsmålene. Hvis vi ser på alle spørsmålene som angår tilfredshet ser vi heller ikke noe klart mønster. Figur 2 viser kjønnsforskjeller i svargjennomsnittene. Forskjellene er gjennomgående små, men kvinner har svart

² I forrige år svarte kvinner i snitt 4,15, og menn 4,08 (differansen var statistisk signifikant i fjor i motsetning til i år).

mer positivt på 27 av spørsmålene og mennene har svart mer positivt på 16 av spørsmålene. På ett av spørsmålene har kvinner og menn svart likt.

Figur 1: Kjønnsforskjeller i gjennomsnittssvar på spørsmål som omhandler tilfredshet



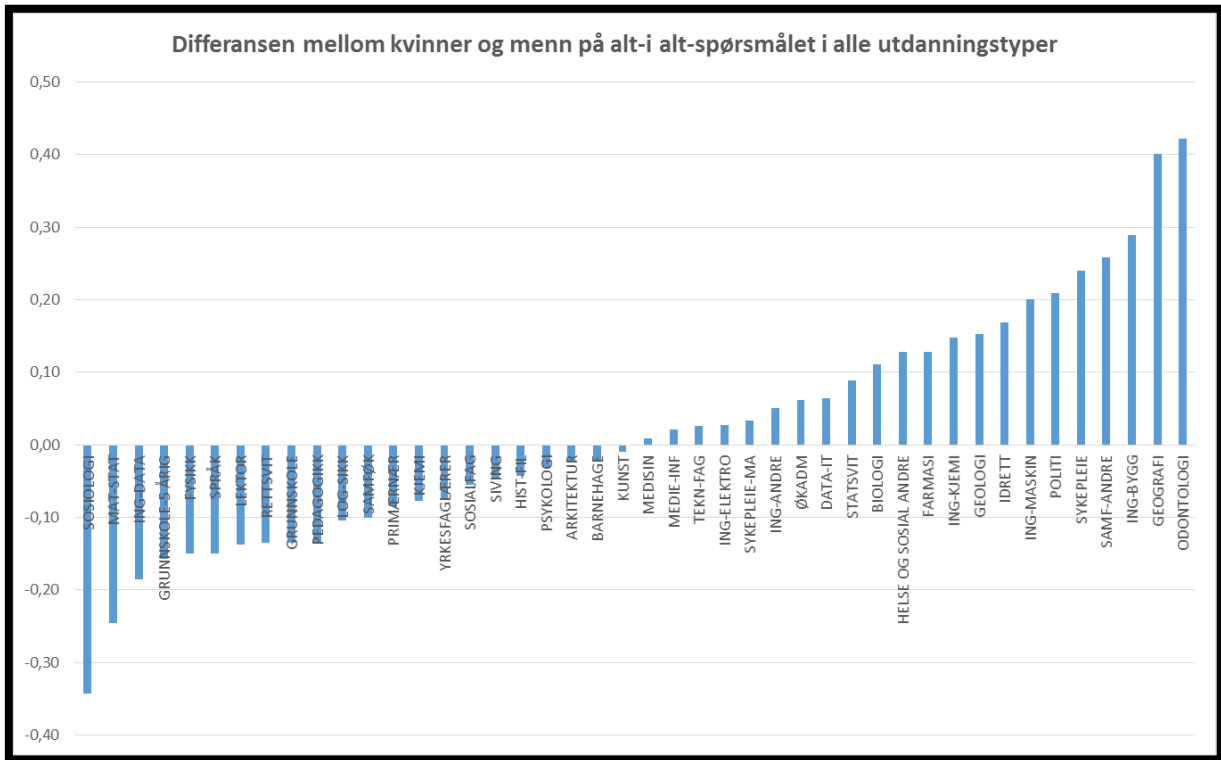
Figuren leses slik: Spørsmålene er angitt i nummerert rekkefølge langs x-aksen. I tilfeller der kvinner har svart mer positivt enn menn, er verdiene/søylene positive (dvs. over streken). I tilfeller der menn har svart mest positivt, befinner søylene seg under streken.

Spørsmålet med størst forskjell er spørsmål 31: «I hvilken grad mener du at eksamens- og innleveringsoppgavene hittil i studieprogrammet ditt: oppfordret til refleksjon og kritisk tenkning». Her svarer kvinner i gjennomsnitt 4,02 og menn 3,74. De andre spørsmålene hvor kvinner svarer klart mer positivt enn menn er «Om eksamen og innleveringsoppgaver [...] oppfordret til å bruke kunnskap fra flere emner for å løse oppgaver», «I hvilken grad gir studieprogrammet [...] kunnskap som er viktig i arbeidslivet» og «hvor tilfreds er du med teoriopplæringens relevans for praksisutøvelse». Spørsmål der menn er mest positive i forhold til kvinner er spørsmål 13 «hvor tilfreds er du med hvordan kritikk og synspunkter fra studentene blir fulgt opp», og spørsmål 20 «hvor tilfreds er du med kommunikasjonen mellom praksissted og høyskole/universitet». Det er heller ikke

Når vi ser på indeksene og på enkelt spørsmål er det altså slik at kvinner svarer mer positivt enn menn på ett flertall av indeksene og enkelt spørsmål. Men som vi så i tabell 1 er ikke alle forskjellene statistisk signifikante. Vi har ikke undersøkt om forskjellene i svargjennomsnitt på enkelt spørsmål er statistisk signifikante.

I forrige runde av Studiebarometeret fant vi at kvinner både var mer tilbøyelige til å svare på undersøkelsen og at de jevnt over var mer positive enn mennene som svarte. Analysen over viser noen av de samme resultatene, men det er mindre forskjeller i den grad kvinner er mer positive. Et interessant spørsmål er om kvinner er mer positive av studiekvaliteten også innenfor hver utdanningstype. I figur 3 viser vi differansen i gjennomsnittssvarene på alt-i-alt spørsmålet for alle utdanningstypene.

Figur 3: Forskjell mellom kvinner og menn i alt-i-alt tilfredshet per utdanningstype



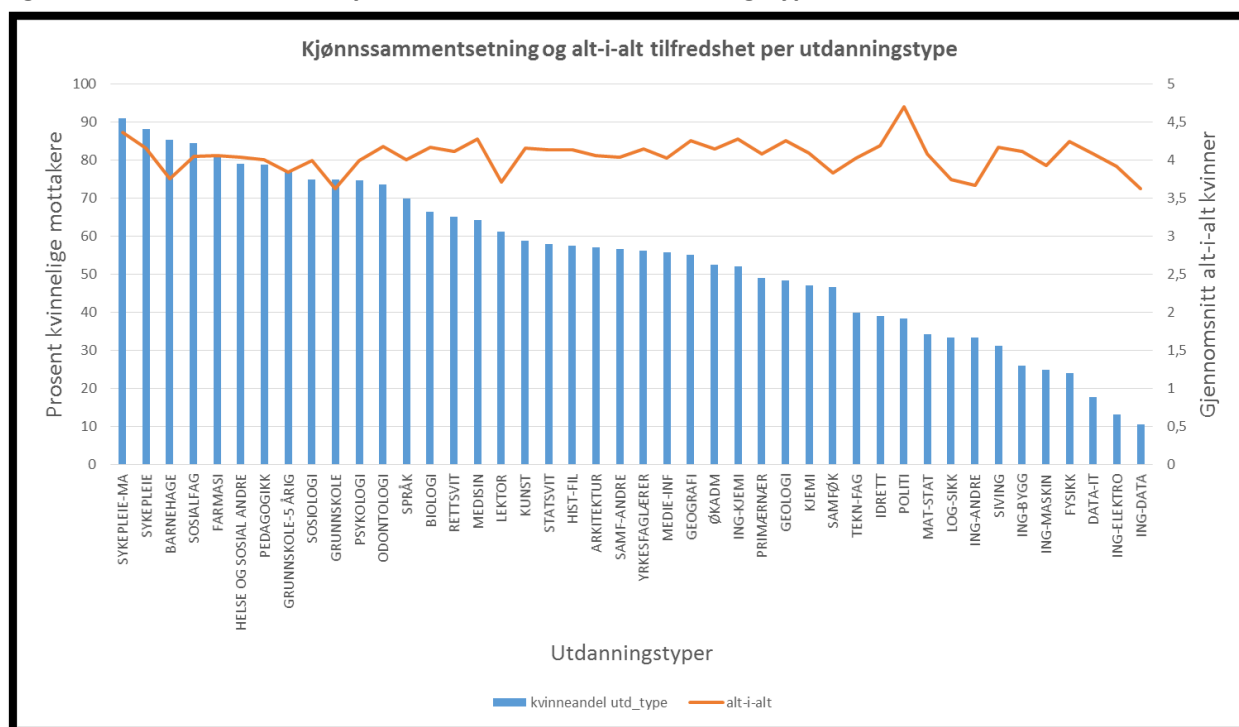
Figuren leses slik: Differansen mellom kjønnene i grad av totalt fornøydhets per utdanningstype fremkommer av søylene. Der kvinner er mer fornøyd enn menn er søylene positive (over streken), der hvor menn er mer fornøyd er søylene negative (under streken).

Det er relativt store forandringer her i forhold til forrige år. I fjor var kvinner mer positive enn menn 31 av 44 utdanningstyper, mens i år er kvinner og menn mest fornøyd i 22 utdanningstyper hver. Det er heller ikke slik at kvinner er mer positive i store eller små programmer. Totalt er det ca. 11000 respondenter i de utdanningstypene hvor menn er mest positive, og ca. 13500 respondenter i de programmene kvinnene er mest positive. Størst avvik er det blant studenter som studerer odontologi og geografiske fag. Kvinner som studerer odontologi er 0,42 poeng (på en skala fra 1–5) mer fornøyd enn menn, mens kvinner som studerer geografiske fag er 0,40 poeng mer fornøyd. Blant de som studerer sosiologi er mennene 0,34 poeng mer fornøyd enn kvinnene.³ Det er viktig å påpeke at det på noen utdanningstyper er svært få menn eller kvinner, dette betyr at et ekstremt svar fra en kvinne eller mann kan påvirke gjennomsnittet kraftig.

Et interessant spørsmål er om disse resultatene har en sammenheng med kjønnssammensetningen innenfor utdanningstypene. Er det for eksempel slik at kvinner er mer fornøyd innenfor typiske mannsdominerte utdanningstyper? Vi ser på dette i figur 4.

³ Vi har ikke testet om forskjellen mellom menn og kvinner i hver utdanningstype er statistisk signifikant.

Figur 4: Er kvinner mer fornøyde i mannsdominerte utdanningstyper?



Figuren leses slik: De blå søylene representerer andelen av kvinner som mottok spørreundersøkelsen. Prosentandelen er beskrevet på den høyre aksene. Den røde linjen representerer kvinnenes gjennomsnitt score på alt-i-alt tilfreds spørsmålet og tallet står på den høyre aksene. Hvis man ser på utdanningstypen «politi» kan man se at kvinnene representerte ca. 39% av mottakerne og gjennomsnittet på alt-i-alt spørsmålet var ca. 4.7.

Figur 4 viser at det ikke er noen sammenheng mellom kjønns sammensetningen innenfor utdanningstypene og kvinners alt-i-alt tilfredshet. Det er med andre ord ikke slik at kvinner som studerer i mannsdominerte utdanningstyper er mer fornøyde enn kvinner som studerer i kvinnedominerte utdanningstyper.

Oppsummert er det klart at kvinner var mer tilbøyelige til å besvare spørreskjemaet, men de er ikke mere positive enn menn. Der vi i fjor så en positiv slagside i materialet siden kvinner både var mer tilbøyelige til å svare og svarte mer positivt enn menn, har vi i år ikke samme tendens. I år som i fjor svarer flere kvinner på skjemaet, men det er liten eller ingen forskjell på hvor positive menn og kvinner er.

Alder

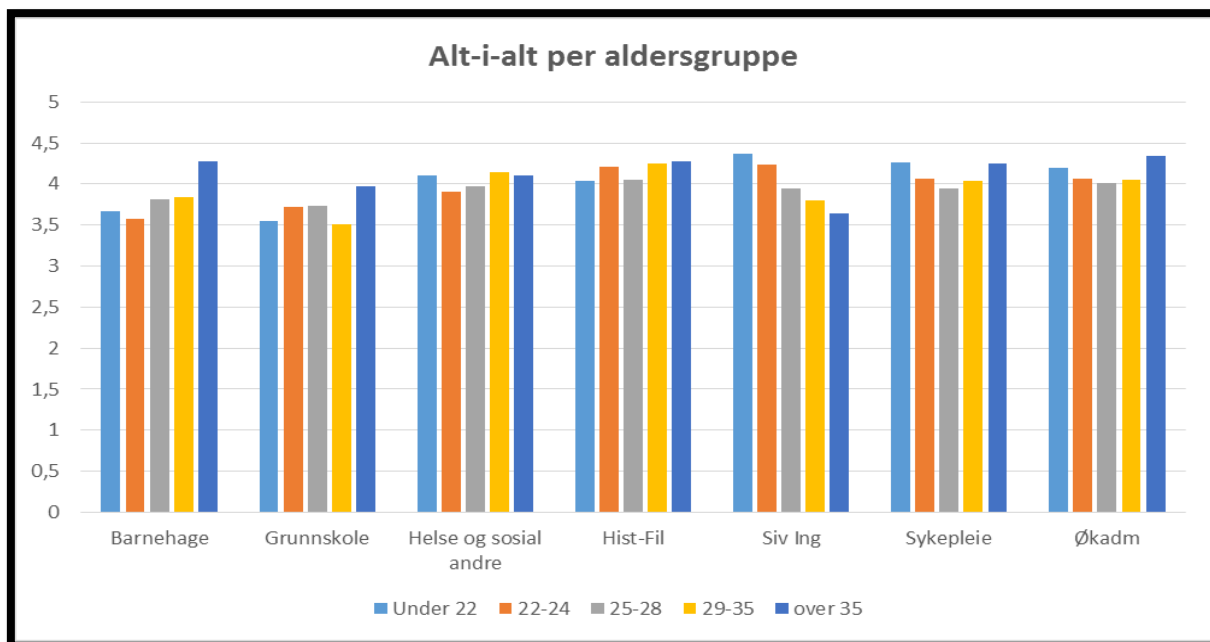
I populasjonen var gjennomsnittsalderen 25,6 år og medianalderen 24 år. Av de som svarte på undersøkelsen var gjennomsnittsalderen 25,5 og medianalderen 23 år. Den gruppevisse fordelingen viser at flertallet av de som svarte er under 24 år og altså rundt median alderen. Samtidig var nesten 20 % av de som svarte over 29 år. Tabell 2 viser antall og andel mottagere og svarende, samt svarprosent per aldersgruppe og alt-i-alt scoren per aldersgruppe.

Tabell 2: Alder, svarinnngang og alt-i-alt tilfredshet

	ANTALL MOTTAGERE	ANDEL MOTTAGERE	ANTALL SVARENDE	ANDEL AV ALLE RESP	SVARPROSENT	ALT-I- ALT	STANDARD- AVVIK(ALT)
UNDER 22	16047	0,27	6834	0,28	0,43	4,10	0,95
22-24	19518	0,33	8284	0,34	0,42	4,04	0,99
25-28	12147	0,21	5032	0,20	0,41	3,98	1,01
29-35	5811	0,10	2397	0,10	0,41	4,00	1,03
OVER 35	5302	0,09	2119	0,09	0,40	4,21	0,96

Gjennomsnittsmålene og aldersfordelingen viser at svarprosenten var tilnærmet den samme for alle aldersgruppene. Dette er en endring fra forrige år da de eldste studentene var mer tilbøyelige til å svare. I år som i fjor er de eldste og yngste studentene de mest fornøyde. Forskjellene er relativt små, men gjennomsnittssvaret til de eldste studentene er høyere enn alle andre grupper og den forskjellen er statistisk signifikant på 0,01 nivå. Forskjellen mellom de yngste studentene og de tre neste gruppene er også statistisk signifikant. Det er ingen statistisk forskjell mellom de tre midterste gruppene.⁴

En analyse av alder og tilfredshet på utdanningstype er vanskelig å gjennomføre. For det første er det få utdanningstyper med tilstrekkelig antall svarende i de forskjellige aldersgruppene, og for det andre er det også mange utdanningstyper med studieprogram på ulike gradsnivå. En analyse av noen av de store utdanningstypene (figur 4) viser at resultatene er sprikende. For noen utdanningstyper er de eldste studentene mest fornøyd, mens for andre utdanningstyper er det andre aldersgrupper som er mest fornøyd.

Figur 4: Forskjell i alder og alt-i-alt tilfredshet

Oppsummert kan vi si at det er liten variasjon i forhold til alderen på de som har besvart undersøkelsen. Det er viser seg at de eldste og yngste studentene er noe mer tilfreds (vi har kun undersøkt alt-i-alt spørsmålet), men at forskjellen, selv om den er statistisk signifikant, er substansielt

⁴ Vi har testet forskjellen på alt-i-alt scoren mellom gruppene med anova/Bonferroni tester.

liten. Vi har med andre ord liten grunn til å frykte at studentenes aldersfordeling påvirker resultatene i noen særlig grad. Disse resultatene stemmer godt overens med fjorårets resultater.

Andre bakgrunnsvariabler

Vi også informasjon om studentene er norske eller utenlandske statsborgere, om de er utreisende utvekslingsstudenter, hvilket land de har sin utdanningsbakgrunn fra, deres opptakskompetanse og om de er innpasset eller ikke. De aller fleste (langt over 90 %) er norske, har sin utdanningsbakgrunn fra Norge, har generell studiekompetanse og er hverken utreisende eller innpasset. Våre analyser (ikke vist) indikerer at forskjellen mellom disse gruppene er svært små, både i svarprosent og alt-i-alt tilfredshet. Disse gruppene er så små og frafallet fra undersøkelsen så liten at de i liten grad påvirker svardataene når vi ser på resultatene på utdanningstype- og nivå. Det er allikevel klart at noen av disse variablene kan påvirke resultatene ved enkelte studieprogram i betydelig grad.

Frafall underveis i skjemautfyllingen

Samlet sett er frafallet underveis i utfyllingen av spørreskjemaet nokså beskjedent. Totalt sett var det 8,2 % av alle som startet å svare på undersøkelsen som ikke fullførte hele undersøkelsen. Tallene er noe forskjellige når vi ser sammenligner menn og kvinner. 7,9 % av kvinnene frafalt, mens 8,7 % av mennene frafalt. Dette stemmer godt overens med den noe høyere svarandelen for kvinner (se ovenfor).

Tilbøyeligheten til å fullføre undersøkelsen varierer lite i forhold til alder. Av de under 29 år fullfører 91,9 % hele undersøkelsen, mens av de over 29 år fullfører 91,4 % hele undersøkelsen. Disse variasjonene er så små at de ikke påvirker resultatene i noen spesiell grad.

Var de mest kritiske eller de mest positive studentene mest tilbøyelige til å besvare undersøkelsen?

I en studentundersøkelse slik som Studiebarometeret er et sentralt spørsmål om de mest positive eller negative studentene er mest tilbøyelige til å svare. Dersom en slik tendens finnes i datagrunnlaget får vi en selvselektert utvalgsskjevhet. For å besvare dette spørsmålet skulle man helst ha gjort intervjuer med de studentene som har valgt å ikke besvare undersøkelsen, men en slik ekstraundersøkelse er meget ressurskrevende og har ikke blitt gjort. Vi belyser denne problemstillingen på tre måter: Først ser vi på om det er forskjell i svarene til de som har svart på undersøkelsen med en gang de fikk den, og de som ikke svarte før de hadde fått mange purringer. Deretter ser vi på forskjellen mellom studenter som har hatt høy studiepoengproduksjon og de som har hatt lav studiepoengproduksjon. Til sist ser vi på sammenhengen mellom studentenes totalvurdering og svarprosenten på programnivå.

Forskjell i tilfredshet på tidlig og sent svarende studenter

Dersom man antar at de mest engasjerte studentene også er de mest tilfredse studentene, er det sannsynlig at de som besvarte undersøkelsen først og uten flere oppfordringer er mer positive enn de studentene som svarte på undersøkelsen sent og etter mange oppfordringer. For *alt-i-alt tilfredshet* spørsmålet er gjennomsnittet for hele svarpopulasjonen 4,05. Tilsvarende verdi for de studentene som svarte på undersøkelsen i løpet av den første dagen de hadde mulighet (14. oktober) er også 4,05 (4131 studenter svarte på undersøkelsen den første dagen). Verdien for de studentene som ikke svarte før november, dvs. to uker etter invitasjonen og etter flere purringer, er 3,97 (1050

studenter svarte på undersøkelsen i november). Denne forskjellen er relativt liten, men statistisk signifikant (t-test, 0,05 signifikansnivå).

Spredningen innenfor gruppene er liten, men de er noe forskjellige mellom gruppene. Standardavviket for de som svarte tidlig er 1,03, mens for de som svarte sent er 0,98. Det betyr at det var litt mindre variasjon i svarene til de som svarte sist enn de som svarte først.

Dersom vi ser på de studentene som falt fra underveis er gjennomsnittsvaret på alt-i-alt spørsmålet 3,99. Dette spørsmålet kommer sent i undersøkelsen og kun 412 av de 2026 som falt fra svarte på dette spørsmålet. Standardavviket for disse studentene er 1,04, altså nokså likt som de andre studentene.

Dersom vi antar at studentene som falt fra og de studentene som kun svarte etter flere purringer er mindre engasjerte, så er det altså slik at de minst engasjerte studentene også er de minst tilfredse studentene.

Oppsummert kan vi si at det er en liten forskjell i tilfredshet mellom de studentene som er engasjerte og svarte raskt på undersøkelsen og de studentene som svarte sent eller ikke fullførte undersøkelsen. Selv om vi ikke kan si noe om de som ikke svarte på undersøkelsen i det hele tatt (58 % av populasjonen), er det mulig at disse studentene er lite engasjerte og også mindre tilfreds enn de studentene som svarte på undersøkelsen. Dersom en slik antagelse er korrekt vil det si at vi har en noe «positiv» slagside i dataene.

Sammenhengen mellom studiepoengproduksjon, svartilbøyelighet og alt-i-alt tilfredshet

En av bakgrunnsvariablene som er registrert for nesten samtlige mottagere og respondenter er gjennomsnittlig studiepoengproduksjon, definert som gjennomsnittet av oppnådde studiepoeng de siste tre semestrene. Det er sannsynlig at studenter med høy studiepoengproduksjon jevnt over er mer engasjerte enn de med lav studiepoengproduksjon, og at dette kan påvirke dere syn på studiekvalitet. Normert produksjon for en fulltidsstudent er 30 studiepoeng per semester. Gjennomsnittlig studiepoengproduksjon blant alle mottager er 24,8, mens gjennomsnittlig studiepoengproduksjon blant respondentene er 26,2. Fordelingen, i de ulike studiepoenggruppene, angis i tabell 3.

Tabell 3: Fordeling av mottageres og respondenters studieprogresjon sammenholdt med svarprosent

STUDIEPOENG	ANDEL AV ALLE MOTTAGERE	ANDEL AV ALLE RESPONDENTER	SVARPROSENT	ANTALL SVARENDE
0 POENG	3 %	2 %	25 %	514
1-16	14 %	10 %	31 %	2460
17-24	17 %	15 %	38 %	3676
25-30	59 %	64 %	46 %	15845
OVER 30	8 %	8 %	46 %	2041

Tabell 3 viser at det er en klar tendens til at jo høyere studiepoengproduksjon studentene har, jo større tilbøyelighet har de til å svare på undersøkelsen. Det er allikevel interessant å se at hele 25 % av studentene med ingen produserte studiepoeng har svart på undersøkelsen.

Den bivariate korrelasjonen mellom studiepoengproduksjon og «alt-i-alt tilfredshet» er 0,02. Korrelasjonen er signifikant, men er likevel substansielt ubetydelig. Studenter med høy studiepoengproduksjon vurderer altså studieprogrammet (alt-i-alt) noen mer positivt enn de med lav studiepoengproduksjon.

For å tydeliggjøre dette ser vi nærmere på tre grupperes gjennomsnitt svar på alt-i-alt spørsmålet.

- Studenter som produserer mindre enn 25 studiepoeng (eksklusive de som produserer 0): 3,99
- Studenter som produserer mellom 25 og 30 studiepoeng: 4,06
- Studenter som produserer over 30 studiepoeng: 4,11

Sammenhengen mellom gruppene og alt-i-alt tilfredshet er omtrent identisk med forrige år, men tallene er noe lavere (i fjor var gjennomsnittet per gruppe 4,04; 4,10; 4,15). Forskjellen mellom gruppe 1 og 2 er signifikant. Det samme gjelder forskjellen mellom gruppe 1 og 3. Derimot er forskjellen mellom gruppe 2 og 3 ikke signifikant. Ettersom den klart største gruppen av studenter produserer mellom 25 og 30 studiepoeng har vi liten grunn til å tro at studiepoengproduksjon påvirker resultatene i nevneverdig retning.

Sammenheng mellom studentenes totalvurdering og svarprosent på programnivå

Dersom det er de mest positive studentene er de som er mest tilbøyelige til å svare på undersøkelsen, vil de studieprogrammene med høyest svarandel være de som har høyest gjennomsnitt score *alt-i-alt tilfredshet* spørsmålet. Vi ser nærmere på dette her. I 2014 var det totalt 878 studieprogrammer som tilfredsstilte terskelverdiene for resultatvisning på studiebarometeret.no.⁵ Den bivariate korrelasjonen mellom svarprosent og alt-i-alt tilfredshet er 0,07 blant disse programmene. Korrelasjonen er signifikant, men substansielt ubetydelig, og vi kan dermed ikke si at det er noen sterk sammenheng mellom svarandel og alt-i-alt tilfredshet.

Blant de programmene som tilfredsstilte terskelverdien var gjennomsnittet på alt-i-alt tilfredshet 4,03. De 50 programmene med høyest svarprosent scorer i gjennomsnitt 4,2, mens de 50 programmene med lavest svarprosent (av de som tilfredsstiller terskelverdiene) har en gjennomsnittlig score på 3,93. Forskjellen i gjennomsnittsscore er relativ liten og heller ikke statistisk signifikant.

Oppsummert finner vi ingen sterk sammenheng i alt-i-alt tilfredshet og svarprosenten per studie program. Vi finner med andre ord ingen grunn til å tro at de mest tilfredse studentene har vært mer tilbøyelige til å svare på undersøkelsen.

Hvordan fungerte spørreskjemaet? (Validitetsanalyse)

I denne delen av analysen ser vi nærmere på om resultatene fra undersøkelsen er valide, altså om undersøkelsen faktisk måler det den tar sikte på å måle. Vi belyser dette ved å analysere frafall gjennom spørreskjemaet (andel blanke og «vet ikke» svar), noe som kan indikere om et spørsmål er upresist formulert eller vanskelig å forstå. Vi har også undersøkt hva studentene som benyttet seg av det åpne spørsmålet har skrevet om spørreskjemaet.

Undersøkelsen som danner grunnlaget for Studiebarometeret ble gjennomført for første gang i 2013. En analyse av reliabilitet og validitet ble gjennomført våren 2014, og i den analysen beskrives utviklingen av spørreskjemaet (Bakken og Lied 2014). Analysen viste at noen spørsmål var vanskelige å svare på og at andre spørsmål hadde forbedringspotensial. Dette gjaldt i hovedsak spørsmål knyttet til tidsbruk, eksamen og vurderingsformer og spørsmål knyttet til valg av studieprogram. På bakgrunn av denne analysen, egne erfaringer og innspill fra sektoren gjorde vi noen forandringer av disse

⁵ Fra 2014 er NOKUTs terskelverdier for å publisere data på portalen at studieprogrammet har flere enn 9 svarende og mer enn 19,5 % svar, eller at studieprogrammet har mellom 6 og 9 svarende og mer enn 49,5 % svar. Disse grenseverdiene eksisterer for at tallene som presenteres skal være robuste og for å sikre anonymisering av de studentene som har svart på undersøkelsen.

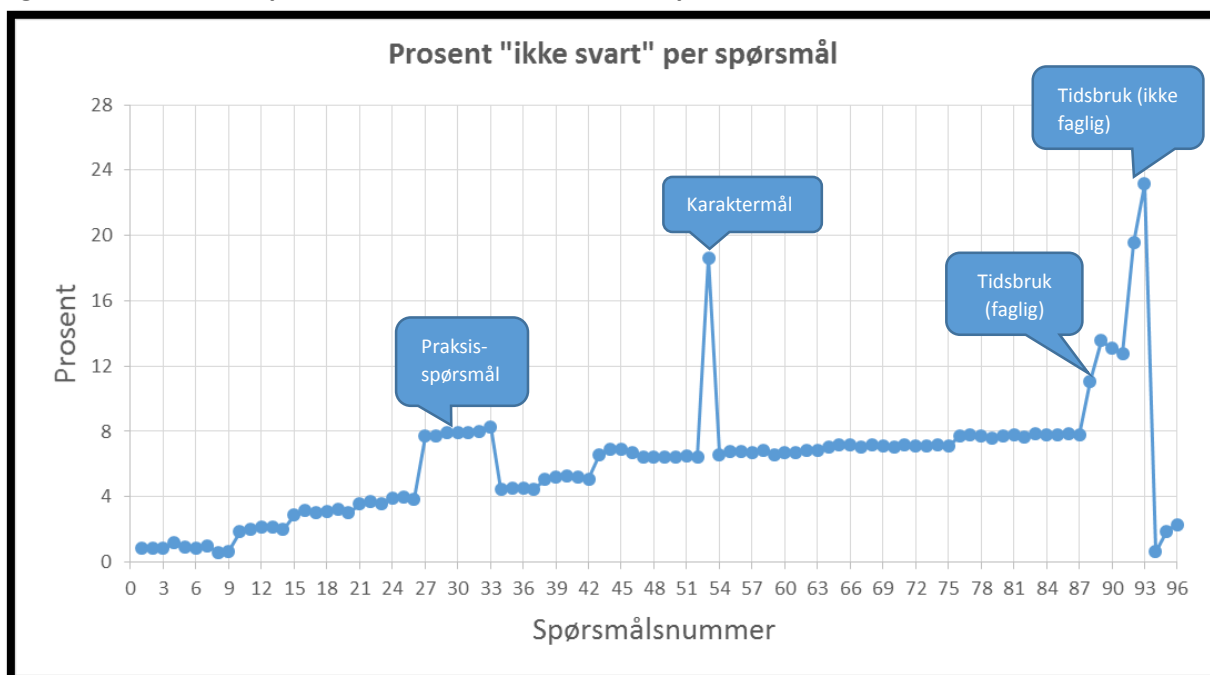
spørsmålene i 2014.⁶ I tillegg la vi i 2014 til ett spørsmålsbatteri om praksis. Som vi ser i analysen nedenfor har disse forandringene hatt varierende suksess, men totalt sett viser analysen at det spørreskjemaet som ble brukt i 2014 inneholder gode spørsmål som de aller fleste studentene forstår og besvarer.

Analyse av frafall i spørreskjemaet

Som nevnt over var det totalt 8.2 % av de som startet å svare på undersøkelsen som ikke fullførte hele undersøkelsen. I denne analysen ser vi på alle spørsmålene for å få en bedre forståelse av om studentene forstår spørsmålene og i hvilken grad vi kan si at de er valide.

Respondentene kan på mange av spørsmålene svare «vet ikke», «brukes ikke» og «ikke relevant», i tillegg kan studenter la være å svare på spørsmålene. I analysen under har vi skilt ut hvor mange studenter som ikke har svart på spørsmålene og hvor mange studenter som har svart «vet ikke». Dersom vi ser på kategorien «ikke svart» øker den gradvis ettersom vi kommer lenger ut i spørreskjemaet (se figur 5). Dette er det vi kan kalle «naturlig frafall». Selv om undersøkelsen kun tar 10–11 minutter i gjennomsnitt å fullføre, så faller flere studenter fra jo lenger ut i undersøkelsen vi kommer. Det er større variasjon i den andelen som har svart «vet ikke», «brukes ikke» eller «ikke relevant», men som analysen under viser, har også dette en naturlig forklaring og påvirker derfor ikke undersøkelsens validitet i nevneverdig grad. I analysen under ser vi først på andelen av studenter som har ikke har svart på enkelt spørsmål, deretter ser vi på andelen som har svart «vet ikke».

Figur 5: Prosent av respondenter som ikke har besvart spørsmål



Figur 5 viser at det meste av frafallet øker gradvis jo lenger ut i spørreskjemaet man kommer. Det første spørsmålsbatteriet som handler om «valg av studiested» har en «ikke besvart» andel på ca. 1 %. Den første unaturlige stigningen i prosenten av «ikke besvart» er spørsmålene knyttet til praksis

⁶ For nærmere beskrivelse av endringer som ble gjort se Bakken, Pål: [Studiebarometeret 2014: gjennomføring og svarinnngang](#)

(spørsmål 27–33). Disse spørsmålene var nye i 2014 og det er mulig de studentene som ikke svarte følte at de ikke hadde nok kunnskap til å svare. Samtidig kunne studentene svare «vet ikke» på disse spørsmålene så det er noe underlig at «ikke besvart» andelen er såpass mye høyere enn for de foregående og etterfølgende spørsmålene. Spørsmål 53 lyder «hvilket mål har du til dine karakterer», den store «ikke besvart» andelen på dette spørsmålet indikerer at mange studenter er ukomfortable med å oppgi karaktermål. Samtidig tyder kommentarer i det åpne kommentarfeltet på at flere studieprogrammer ikke bruker karakterer eller kun bruker «bestått / ikke bestått» og dette kan også ha bidratt til den relativt store andelen av studenter som ikke har svart på spørsmålet. Det er allikevel slik at nesten alle studentene som ikke besvarer dette spørsmålet fortsetter undersøkelsen. Den store økningen i «ikke besvart» mot slutten av undersøkelsen dreier seg om studentenes tidsbruk.

Spørsmål 88-90 gjelder tidsbruk på faglige aktiviteter, mens spørsmål 91–93 gjelder ikke-faglig tidsbruk. De tre spørsmålene om faglig tidsbruk er «Studier totalt: Alle timer brukt på faglig aktivitet», «Antall timer brukt på organiserte læringsaktiviteter» og «antall timer brukt på egenstudier». Disse tre spørsmålene er det henholdsvis 11,08, 13,57 og 13,1 % som ikke har svart på. De tre neste spørsmålene dreier seg om tidsbruk på «betalt arbeid», «ubetalt arbeid tilknyttet studier» og «annet ubetalt arbeid». Disse spørsmålene, spesielt de to siste, er det mange studenter som ikke har svart på (henholdsvis 12,8, 19,5 og 23,15 %). I undersøkelsens kommentarfelt har en del studenter påpekt at det er vanskelig å anslå hvor mange timer de bruker i gjennomsnitt per uke på forskjellige aktiviteter. Dette er et av de spørsmålsbatteriene vi endret fra 2013. Vi endret spørsmålet og svarmulighetene på spørsmålet om faglig tidsbruk, samt at spørsmålene om ikke faglig tidsbruk ble delt inn i tre spørsmål. I forhold til forrige år, er det omtrent like mange som ikke svarte i 2014 som ikke svarte eller svarte «vet ikke» i 2013.

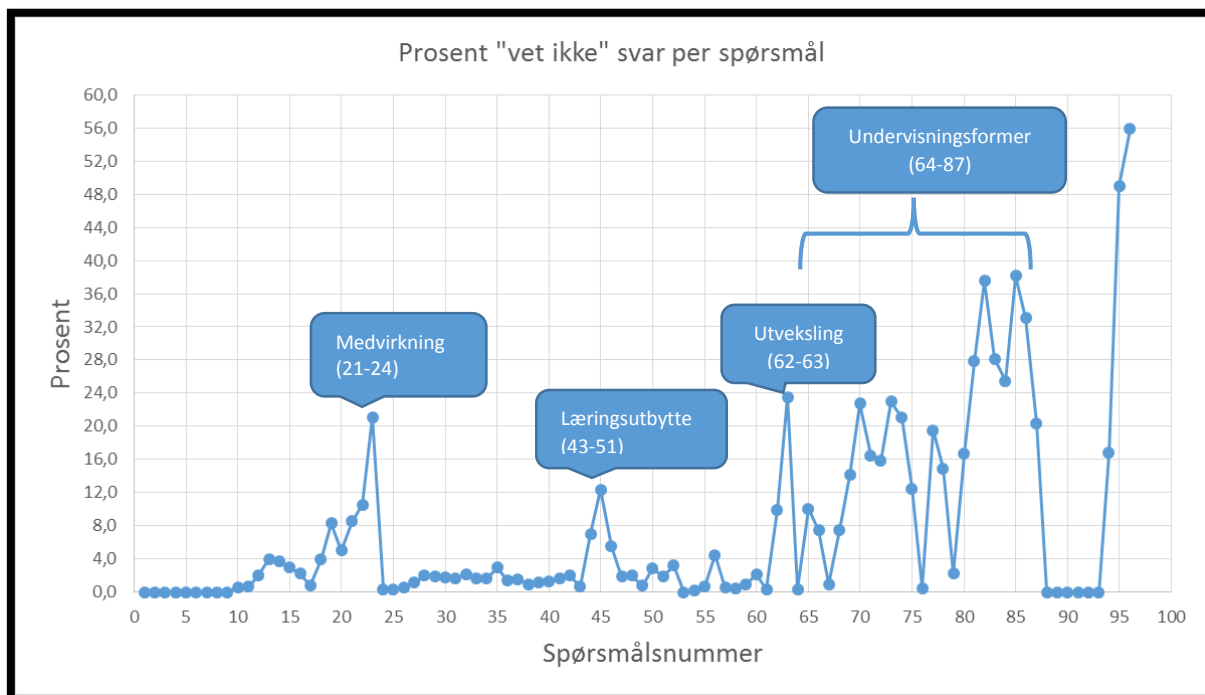
De tre siste spørsmålene dreier seg om relevansen av tiden man bruker på «ikke faglig» arbeid. De siste tre spørsmålene ble kun stilt til de studentene som hadde anmerket at de hadde brukt minimum en time på «ikke faglige aktiviteter». De studentene som fikk disse spørsmålene (ca. 15000) har i stor grad svart på dem, men som vi ser under har mange svart «vet ikke».

At andelen «ikke besvart» øker naturlig (bortsett fra de unntakene som har blitt gjort rede for) viser i utgangspunktet at spørsmålene treffer godt og blir forstått av respondentene.

Analyse av «vet ikke», «ikke relevant» og «brukes ikke»

I tillegg til andelen av «ikke besvart» er det viktig å se på andelen av respondentene som har svart «vet ikke» på spørsmålene. Høye andeler av «vet ikke» svar kan være et annet signal på at spørsmålene ikke har blitt forstått eller at svaralternativene ikke er gode nok. Figur 6 viser at andelen «vet ikke», «brukes ikke» og «ikke relevant» varierer mye.

Figur 6: Andel «vet ikke» svar per spørsmål⁷



For å få en bedre forståelse av hvorfor studentene har svart «vet ikke», må vi se på de spesifikke spørsmålene og svaralternativene.

Den første bolken av spørsmål med høy «vet ikke» svarandel er spørsmål som dreier seg om studentenes medvirkning (spørsmål 21–24). Studentene blir her spurt om «Hvor tilfreds de er med:»

1. Studentenes mulighet for å påvirke innhold og opplegg i studieprogrammet
2. Hvordan kritikk og synspunkter fra studentene blir fulgt opp
3. Det lokale studentdemokratiet (f.eks. tillitsvalgte og studentorganisasjon)

Tabell 4: Medvirkning - svarandel «vet ikke»

Spørsmål	Medvirkning		
	Evne til å påvirke studieprogram	Kritikk fulgt opp	Studentdemokrati
Antall «Vet ikke»	2101	2602	5205
Andel «vet ikke»	8,5	10,5	21,1

Selv om svarandelen «vet ikke» på disse spørsmålene er høye er det relativt liten grunn til bekymring. En sannsynlig årsak er at studentene ikke kjenner godt nok til situasjonen ved sitt studieprogram.

Den neste bolken med høy svarandel «vet ikke» gjelder spørsmål knyttet til «eget læringsutbytte» (spørsmål 43–51). Her ble studentene spurt om «hvor tilfreds er du med eget læringsutbytte hittil i studiet, når det gjelder:»

1. Teoretisk kunnskap
2. Kunnskap om vitenskapelig arbeidsmetode og forskning
3. Egen erfaring med forsknings- og utviklingsarbeid

⁷ På noen spørsmål kunne studentene svare «vet ikke» på andre spørsmål «brukes ikke» eller «ikke relevant» Figuren viser prosenten på alle tre varianter.

4. Yrkes- og fagspesifikke ferdigheter
5. Evne til refleksjon og kritisk tenking
6. Samarbeidsevne
7. Evne til å jobbe selvstendig
8. Muntlig kommunikasjonsevne
9. Skriftlig kommunikasjonsevne

Tabell 5: Eget læringsutbytte – svarandel «vet ikke»

Eget læringsutbytte										
Spørsmål	Teoretisk kunnskap	Kunnskap vitenskap. arb. metode	Erfaring FoU	Yrkes ferdigheter	Refleksjon og kritisk tenkning	Samarbeids- evne	Arbeide selvstendig	Muntlig kommunikasjon	Skriftlig kommunikasjon	Tenke nytt
Antall «Vet ikke»	170	1734	3044	1359	461	493	189	721	465	782
Andel «vet ikke / ikke relevant»	0,7	7,0	12,3	5,5	1,9	2,0	0,8	2,9	1,9	3,2

Som vi ser i tabell 5 svarer forholdsvis mange studenter «vet ikke» på spørsmål 2 og 3. Igjen ser vi ingen problemer med dette spørsmålet. Årsaken til den høye «vet ikke» andelen er sannsynligvis at en del studenter er usikre på om de har fått slik kunnskap eller om de har hatt erfaring med FoU. Det faktum at 12,3 % av studentene ikke vet om de har hatt egen erfaring med FoU sier muligens noe om at studieprogrammene ikke er dyktige nok til å formidle hva FoU er, og i hvilken grad studentene bruker dette i læringen sin.

De neste spørsmålene med høy svarandel «vet ikke» gjelder spørsmål om utveksling (spørsmål 62 og 63). Her svarer henholdsvis 9,9 % og 23,5 % av studentene at de ikke vet om studieprogrammet «informerer godt om utvekslingsmuligheter» og «legger godt til rette for utveksling til utlandet». Heller ikke her tror vi det er spørsmålet som er årsaken til den høye «vet ikke» andelen, men heller mangel på informasjon fra studieprogrammets side, samt at mange studenter ikke har undersøkt om mulighetene til utveksling.

Spørsmål 64–87 dreier seg om undervisnings- og arbeidsformer. I det første batteriet spør vi: «I hvilken grad brukes disse undervisnings- og arbeidsformene i studieprogrammet ditt:»

1. Forelesning
2. Seminar
3. Gruppe uten lærer
4. Skriftlig arbeid til innlevering
5. Prosjektarbeid
6. Feltarbeid / egen datainnhenting
7. Laboratorieøving
8. Annen praktisk øving
9. Case
10. Simulering/rollespill
11. Praksisperioder
12. Digitale arbeidsformer

I det påfølgende batteriet spør vi: «I hvilken grad bidrar disse undervisnings- og arbeidsformene til læringen din?» på begge disse batteriene er det til dels stor variasjon i andel «vet ikke». Vi viser dette i tabell 6 og 7.

Tabell 6: Bruk av undervisnings- og arbeidsformer – svarandel «vet ikke»

Bruk av undervisnings og arbeidsformer i studieprogrammet												
Spørsmål	Forelesning	Seminar	Gruppe uten lærer	Skriftlig arb. til innlevering	Prosjektarbeid	Feltarbeid	Laboratorieøving	Annen praktisk øving	Case	Simulering	Praksis	Digitale arbeidsformer
Antall «Vet ikke / ikke relevant»	86	2477	1838	234	1855	3488	5605	4071	3902	5675	5187	3076
Andel «vet ikke/ ikke relevant»	0,3	10,0	7,5	0,9	7,5	14,1	22,7	16,5	15,8	23,0	21,0	12,5

Tabell 7: Nytte av undervisnings- og arbeidsformer – svarandel «vet ikke»

Nytte av undervisnings og arbeidsformer i studieprogrammet												
Spørsmål	Forelesning	Seminar	Gruppe uten lærer	Skriftlig arb. til innlevering	Prosjektarbeid	Feltarbeid	Laboratorieøving	Annen praktisk øving	Case	Simulering	Praksis	Digitale arbeidsformer
Antall «Vet ikke / brukes ikke»	121	4812	3672	556	4115	6862	9263	6937	6281	9405	8158	5006
Andel «vet ikke / brukes ikke»	0,5	19,5	14,9	2,3	16,7	27,8	37,6	28,1	25,5	38,1	33,1	20,3

Når det gjelder den store andelen av «vet ikke» / «ikke relevant» på de to spørsmålsbatteriene angående undervisning og arbeidsformer er størstedelen av disse på undervisningsformer vi vet brukes relativt lite. Hvis vi ser på «forelesning» ser vi at kun rundt 0,3 % av studentene ikke har svart på dette spørsmålet, men denne prosenten øker med undervisningsformer vi vet er lite brukt. Heller ikke her er grunnen til den store «vet ikke» andelen dårlige spørsmål eller svaralternativer, men slett at studentene ikke har brukt mange av de alternative arbeidsformene.

Tilslutt i denne analysen ser vi på kort på tidsbruk spørsmålene og spørsmålene relatert til «arbeidets relevans». Studentene fikk totalt seks spørsmål om tidsbruk. De første tre spørsmålene dreide seg om tidsbruk på faglige aktiviteter og de tre siste spørsmålene om tidsbruk på ikke-faglige aktiviteter. Studentene ble bedt om å skrive inn antall «timer per uke, i gjennomsnitt på dette studiet (ikke medregnet ferier), du bruker på:» Studentene kunne her skrive et tall, eller la være å svare. På disse spørsmålene var det ikke mulig å svare «vet ikke». Som vi så i analysen over (figur 5), var det en god del studenter som ikke svarte på disse spørsmålene, samt at en del studenter svarte meget høye verdier (over 70 timer) og mange som svarte «0».

Spørsmålene om tidsbruk på «ikke-faglige aktiviteter» var knyttet direkte til spørsmål om «arbeidets relevans». Dersom en student hadde oppgitt at hun hadde brukt 1 time eller mer på enten «betalt jobb», «ubetalt arbeid i tilknytning til studier» eller «annet ubetalt arbeid», fikk hun også spørsmålet: «I hvilken grad er arbeidet ditt relevant for utdanningen?» Uansett om en student hadde svart 1 time på betalt arbeid, og 0 timer på ubetalt arbeid fikk studenten 3 spørsmål om arbeidets relevans. Dette har ført til en meget høy «vet ikke / ikke aktuelt» svarandel på de tre siste spørsmålene (se figur 6, spørsmål 94-96). Det er altså ikke spørsmålene angående arbeidets relevans det er problemer med, men koblingen mellom de to siste spørsmålsbatteriene.

Oppsummert kan vi si at de fleste spørsmålene har blitt forstått og besvart av respondentene. Noen spørsmål har mange «vet ikke» svar, men det er liten grunn til å tro at dette er pga. spørsmålene

eller svaralternativene. Kommentarer fra respondentene kan tyde på at mange studenter finner det vanskelig å estimere tidsbruken sin, men vi har allikevel ingen grunn til å tro at tallene fra de som har svart ikke er reliable eller valide. Til slutt kan det virke som om en del studenter er ukomfortable med å oppgi sine karaktermål.

Hva sier studentene om spørreskjemaet?

Det siste elementet i spørreskjemaet er et kommentarfelt hvor vi ber studentene kommentere spesielt på spørreskjemaet. Ca. 8000 studenter skrev noe i dette feltet. Majoriteten av kommentarene var positive til spørreskjemaet, og over 4000 studenter skrev at skjemaet var «bra», «veldig bra», «fungerte bra» eller lignende kommentarer. Av de studentene som var mer negative, dreide kommentarene seg nesten kun om lengden på spørreskjemaet. Mange studenter kommenterte også på hvor vanskelig det var å beregne antall timer brukt på faglige og ikke-faglige aktiviteter. Andre konstruktive kommentarer som gikk igjen var ofte relatert til hjelpetekster, eller mangel på slike. Flere studenter etterspurte hjelpetekster på de forskjellige undervisnings- og arbeidsformene. Som forventet var det også mange studenter som ønsket seg flere kommentarfelter hvor de kan begrunne sine svar, og beskrive sin egen situasjon mer i detalj. Det var også en del deltidsstudenter som følte at mange av spørsmålene var lite relevante for deres studiesituasjon. Alt i alt kan vi si at spørreskjemaet har blitt godt mottatt av studentene.

Konklusjon

Undersøkelsen som danner grunnlaget for studiebarometeret.no, er ikke basert på et tilfeldig utvalg. Det er derfor viktig at vi analyserer om det er systematiske skjevheter i datagrunnlaget som kan påvirke resultatene. I denne analysen har vi sett på noen av de viktigste kildene til skjevhet, slik som kjønn, alder, studieprogresjon og flere andre variabler. Analysen viser at det er liten grunn til å frykte at resultatene er preget av systematiske skjevheter. Selv om kvinner i større grad har svart på undersøkelsen er det ingen tegn til at kvinner er systematisk mer positive eller negative enn menn. Når det gjelder alder er det de yngste og eldste studentene som er mest positive, men forskjellene er små og substansielt ubetydelige. Analysen viser en liten forskjell i tilfredshet på de studentene som har svart på undersøkelsen tidlig, og de som har svart sent og etter flere purringer. Dersom vi antar at de som studentene som svarte tidlig er mer engasjerte betyr det at de mest engasjerte studentene også er de mest tilfredse. Denne forskjellen er liten, men statistisk signifikant. Dette kan også bety at de studentene som ikke har svar på undersøkelsen er lite engasjerte og også noe mindre tilfreds enn de som har svart på undersøkelsen. Dersom det er slik at de mest engasjerte studentene er mest tilfreds og også er mer tilbøyelige til å svare, så bør det også være slik at de programmene med høyere svarandel er gjennomsnittlig mer fornøyd enn programmer med lav svarprosent. I analysen over er ikke dette tilfellet. Forskjellen mellom de 50 programmene med høyest og lavest svarandel er ikke statistisk signifikant. Med andre ord, selv om de studentene som svarer tidlig er noe mer tilfreds enn de som svarer sent finner vi ikke den samme sammenhengen når det gjelder svarandel per program og tilfredshet.

Oppsummert kan vi si at resultatene i Studiebarometeret er relativt representative. Vi finner få sterke bevis for at det finnes systematiske skjevheter i datagrunnlaget, noe som styrker de funnene vi presenterer i portalen. Det er interessant å merke at årets datagrunnlag er noe forskjellig fra forrige år. I fjor fant vi en noe positiv slagside, spesielt siden kvinner systematisk svarte mer positivt i enn menn. Bortsett fra den forskjellen er resultatene stort sett de samme i fjor og i år.

Når det gjelder validiteten i dataene er også disse gode. Spørsmålene er stort sett forstått og besvart. Det er et naturlig frafall i svarprosent jo lenger ut i undersøkelsen vi kommer, men dette er ikke overraskende for en undersøkelse med nesten 100 spørsmål. Det er også en del spørsmål der

respondentene har svart «vet ikke» eller «brukes ikke» eller «ikke relevant», men analysen viser at dette sannsynligvis er fordi studentene har liten eller ingen erfaring med de aktuelle temaene. Det er med andre ord ikke fordi spørsmålene eller svaralternativene er dårlige. Validitetsanalysen viser også at årets undersøkelse er noe bedre enn forrige års undersøkelse. En lignende analyse i 2014, samt andre vurderinger førte til en del endringer i spørreskjemaet og årets analyse viser at de forandringene som ble gjort førte til et bedre spørreskjema i 2014.