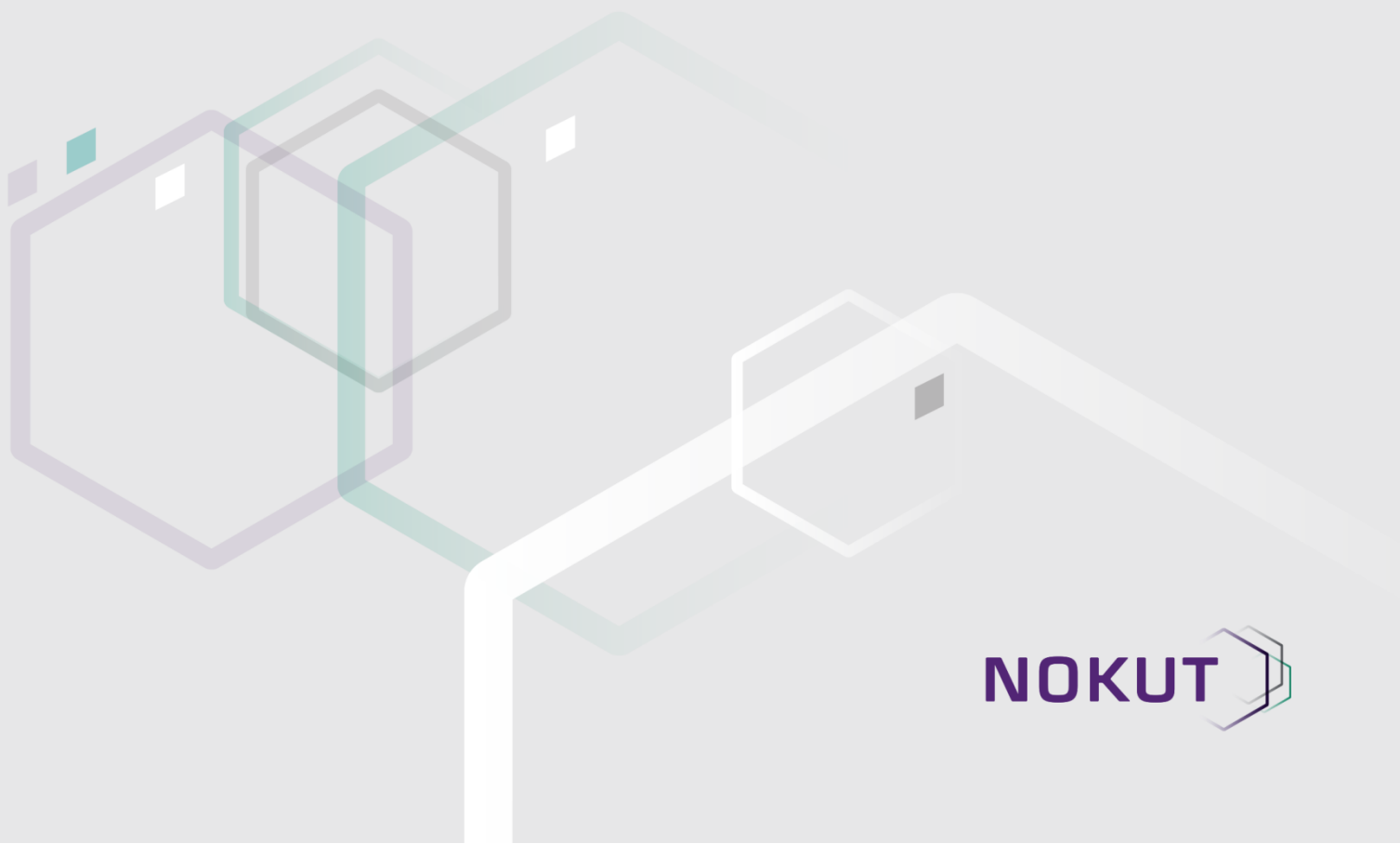


NOKUTs oppsummeringer

# Nasjonal deleksamen i anatomi, fysiologi og biokjemi for sykepleierutdanningen – høsten 2017

Mars 2018



NOKUT 

NOKUTs arbeid skal bidra til at samfunnet har tillit til kvaliteten i norsk høyere utdanning og fagskoleutdanning, samt godkjent høyere utenlandsk utdanning. Med rapportserien "NOKUTs oppsummeringer" vil vi bidra til økt kunnskap om temaer knyttet til fagskole, høyere utdanning og godkjenning av utenlandsk utdanning i Norge. Data til rapportene får NOKUT gjennom arbeidet med akkreditering og godkjenning.

Vi håper at resultatene våre kan være nyttige i arbeidet med godkjenning av utenlandsk utdanning og for lærestedene i arbeidet med å kvalitetssikre og videreutvikle utdanningstilbudene.

<b>Tittel:</b>	Nasjonal deleksamen i anatomi, fysiologi og biokjemi i sykepleierutdanningen - høsten 2017
<b>Forfattere:</b>	Lars Fredrik Pedersen Karl Johan Skeidsvoll Kjersti Tokstad
<b>Dato:</b>	22. mars 2018

## Innhold

<b>1</b>	<b>Innledning</b> .....	<b>1</b>
1.1	Organisering av nasjonal deleksamen .....	1
1.1.1	Eksamensdagen .....	2
1.1.2	Sensur .....	2
1.2	Utvalget .....	2
<b>2</b>	<b>Nasjonale resultater</b> .....	<b>4</b>
<b>3</b>	<b>Analyser</b> .....	<b>5</b>
3.1	Om dataene .....	5
3.2	Karakterpoeng fra videregående skole .....	6
3.3	Institusjonsbidrag.....	8
<b>4</b>	<b>Sensurreliabilitet</b> .....	<b>11</b>
<b>5</b>	<b>Klagesensur</b> .....	<b>14</b>
<b>6</b>	<b>Eksamensoppgaver og sensorveiledning</b> .....	<b>14</b>
<b>7</b>	<b>Konklusjon</b> .....	<b>15</b>
	Vedlegg 1: Karakterfordelingen per institusjon og studiested, 2016 og 2017. ....	17



## 1 Innledning

NOKUT fikk høsten 2014 i oppdrag av Kunnskapsdepartementet å gjennomføre en mulighetsstudie og et pilotprosjekt med nasjonale deleksamener i høyere utdanning. Tre rammeplanstyrte profesjonsutdanninger ble valgt ut til å delta i prosjektet, som pågikk fra 2015 til 2017. Disse er grunnskolelærerutdanningene (GLU 1–7 og GLU 5–10, sistnevnte for studenter som velger matematikk), bachelorgradsstudiet i sykepleie og bachelorgradsstudiet i regnskap og revisjon. En oppsummering av pilotprosjektet kan leses i rapporten *Nasjonal deleksamen – et pilotprosjekt og en mulighetsstudie* fra mars 2017.<sup>1</sup>

Denne rapporten handler om nasjonal deleksamen i anatomi, fysiologi og biokjemi i sykepleierutdanningen som ble gjennomført 18. desember 2017, som er den tredje og siste eksamenen i pilotprosjektet for sykepleierutdanningen. Den første nasjonale deleksamenen i anatomi, fysiologi og biokjemi ble avholdt 17. desember 2015, og i den første rapporten beskrev vi formålet med prosjektet og de viktigste innvendingene mot prosjektet, samt en del av utfordringene. Vi kommer derfor ikke videre inn på disse temaene i denne rapporten, men viser til den første rapporten.<sup>2</sup>

Bachelorgradsstudiet i sykepleie ble tilbudt ved 28 utdanningsinstitusjoner før fusjonene i sektoren trådte i kraft 1. januar 2016. Den første nasjonale deleksamenen som ble avholdt i 2015, omfattet kun 12 av de 28 institusjonene, og var en slags pilot i piloten. Forskjellene mellom institusjonenes undervisningsopplegg og vurderingsformer i anatomi, fysiologi og biokjemi var for store til at det var mulig å gjennomføre en fullskala eksamen allerede i 2015. Etter fusjonene i 2016 og 2017 har antallet institusjoner som tilbyr bachelorgradsstudiet i sykepleie blitt redusert til 14, og i 2016 og 2017 har den nasjonale deleksamenen i anatomi, fysiologi og biokjemi omfattet alle sykepleierstudentene og blitt gjennomført som deres ordinære eksamen i faget.

I september 2017 bestemte Kunnskapsdepartementet at nasjonale deleksamener skal videreføres, nå som en permanent ordning, i første omgang innenfor de samme fagområdene og profesjonsutdanningene som har deltatt i pilotprosjektet. Det innebærer at eksamen i anatomi, fysiologi og biokjemi skal fortsette som en nasjonal eksamen inntil videre.

### 1.1 Organisering av nasjonal deleksamen

Det ble satt ned en arbeidsgruppe som utviklet den nasjonale deleksamenen i anatomi, fysiologi og biokjemi i 2015. Med unntak av et par utskiftninger underveis, er det i all hovedsak den samme gruppen som har utviklet alle de tre eksamenene i pilotprosjektet. I tillegg til eksamensoppgaver og sensorveiledninger har arbeidsgruppen også utviklet en felles læringsutbyttebeskrivelse og en detaljert fagbeskrivelse. Læringsutbyttebeskrivelsen og fagbeskrivelsen var oppe til diskusjon i fagmiljøene og i Nasjonalt profesjonsråd for utdanning og forskning innen sykepleie før de ble sendt ut til institusjonene i juni 2016.

<sup>1</sup> [https://www.nokut.no/globalassets/nokut/rapporter/nasjonal-deleksamen7/nasjonal\\_deleksamen\\_et\\_pilotprosjekt\\_og\\_en\\_mulighetsstudie\\_2017.pdf](https://www.nokut.no/globalassets/nokut/rapporter/nasjonal-deleksamen7/nasjonal_deleksamen_et_pilotprosjekt_og_en_mulighetsstudie_2017.pdf)

<sup>2</sup> [https://www.nokut.no/globalassets/nokut/rapporter/nasjonal-deleksamen7/sykepleier/nasjonal\\_deleksamen\\_i\\_anatomi\\_fysiologi\\_og\\_biokjemi\\_for\\_sykepleierutdanningene\\_2016-2.pdf](https://www.nokut.no/globalassets/nokut/rapporter/nasjonal-deleksamen7/sykepleier/nasjonal_deleksamen_i_anatomi_fysiologi_og_biokjemi_for_sykepleierutdanningene_2016-2.pdf)

NOKUT har hatt et overordnet ansvar overfor sektoren og departementet, og et administrativt og logistikkmessig ansvar i forbindelse med eksamensgjennomføringen, som bestilling og fordeling av eksamenspapir, kandidatnumre, distribusjon av besvarelser til sensorkorpset samt formidling av eksamensresultatene til institusjonene.

### **1.1.1 Eksamensdagen**

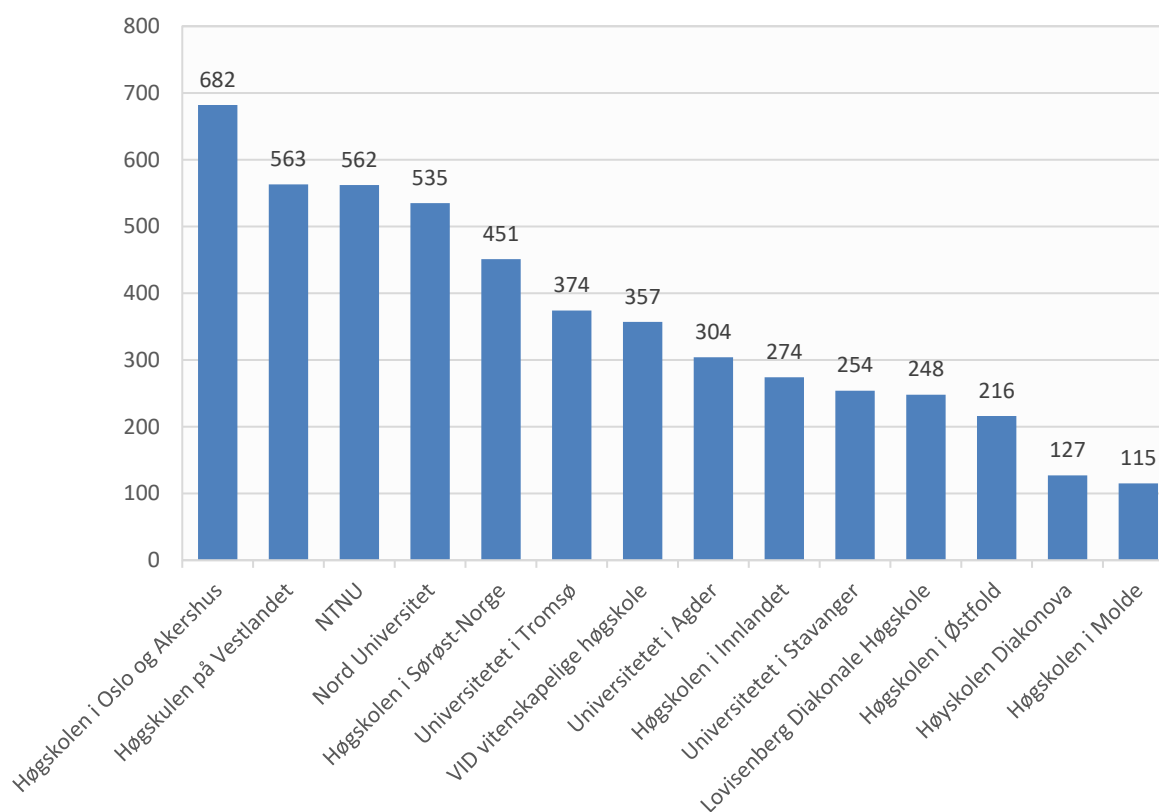
Den tredje nasjonale deleksamenen i anatomi, fysiologi og biokjemi ble avholdt 18. desember 2017 som en firetimers skriftlig eksamen uten bruk av hjelpemidler. Det var 14 institusjoner som deltok, fordelt på 34 studiesteder, og kandidatene var i all hovedsak i sitt første semester. Totalt gjennomførte 5 062 av 5 650 oppmeldte kandidater, som tilsvarer en gjennomføringsgrad på 89,6 prosent.

### **1.1.2 Sensur**

Det er avgjørende å sikre likebehandling og anonymitet. Det ble det oppnevnt et nasjonalt sensorkorps på 80 sensorer, og hver besvarelse ble vurdert av to sensorer. For å få til en karakterkalibrering og en institusjonell fordeling av besvarelsene, delte NOKUT inn sensorene i 159 sensorpar, hvor hver sensor sensurerte sammen med fire andre sensorer. Hvert sensorpar samsensurerte i gjennomsnitt 32 besvarelser, og hver sensor sensurerte i snitt 127 besvarelser, som alle var jevnt og tilfeldig fordelt fra alle læresteder.

## **1.2 Utvalget**

I 2017 var det 5 062 kandidater som avla eksamen, mot 4 838 kandidater i 2016. Figuren under viser antall kandidater per institusjon i 2017, sortert etter antall. Det var flest kandidater fra Høgskolen i Oslo og Akershus.



Figur 1: Antall kandidater per institusjon i 2017.

Tabell 1 viser antall kandidater som avla eksamen i 2017 og 2016, fordelt på institusjon og studiested.

Tabell 1: Antall kandidater i 2017 og 2016 per institusjon og studiested.

Institusjon og studiested	Antall 2017	Antall 2016
Høgskolen i Oslo og Akershus - Kjeller	181	171
Høgskolen i Oslo og Akershus - Pilestredet	465	478
Høgskolen i Oslo og Akershus - Sandvika	36	29
Høgskolen i Sørøst-Norge - Drammen	159	180
Høgskolen i Sørøst-Norge - Porsgrunn	157	181
Høgskolen i Sørøst-Norge - Vestfold	135	133
Høgskulen på Vestlandet - Bergen	258	184
Høgskulen på Vestlandet - Førde	112	72
Høgskulen på Vestlandet - Haugesund	106	104
Høgskulen på Vestlandet - Stord	87	90
Høgskolen i Innlandet*	274	256
Høgskolen i Molde	115	163
Høgskolen i Østfold	216	172
Høyskolen Diakonova	127	127
Lovisenberg Diakonale Høgskole	248	225

Nord Universitet - Bodø	158	166
Nord Universitet - Levanger og Namsos**	236	211
Nord Universitet - Mo i Rana	101	78
Nord Universitet - Vesterålen	40	20
NTNU - Gjøvik	222	231
NTNU - Trondheim	204	197
NTNU - Ålesund	136	120
Universitetet i Agder - Grimstad	129	132
Universitetet i Agder - Kristiansand	175	150
Universitetet i Stavanger	254	272
Universitetet i Tromsø - Hammerfest	66	100
Universitetet i Tromsø - Harstad	97	65
Universitetet i Tromsø - Narvik	47	50
Universitetet i Tromsø - Tromsø	164	132
VID vitenskapelige høgskole - Bergen	177	170
VID vitenskapelige høgskole - Oslo	180	179
<b>Totalt</b>	<b>5 062</b>	<b>4 838</b>

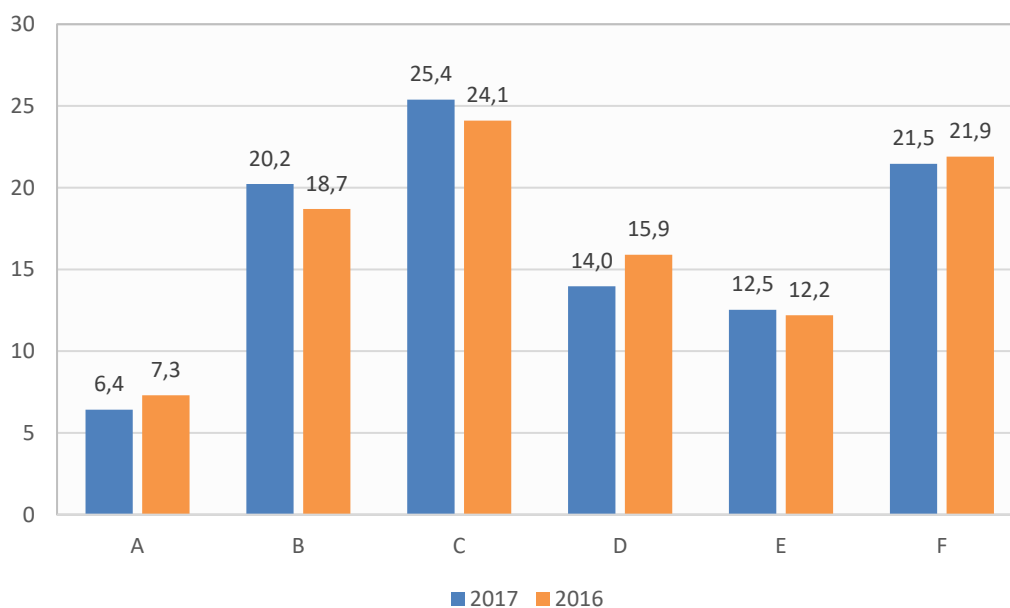
\*Tallet for 2016 er Høgskolen i Hedmark. 2017 er Elverum og Kongsvinger.

\*\*I 2016 var ikke tallene for Levanger og Namsos. 2017-tallet er slått sammen.

## 2 Nasjonale resultater

Her presenterer vi det nasjonale resultatet fra nasjonal deleksamen i anatomi, fysiologi og biokjemi for 2017. Figur 2 viser karakterfordelingen i 2017 sammenlignet med 2016. Det er kun små endringer i karakterfordelingen mellom de to årene. Andelen stryk er omtrent uendret, det samme gjelder andelen som fikk E. I 2017 var det en noe lavere andel som fikk A og D, mens det var en noe høyere andel som fikk B og C. Se vedlegget for en oversikt over karakterfordelingen per institusjon og studiested.





Figur 2: Karakterfordelingen for hele landet i 2016 og 2017.

Videre i rapporten presenterer vi mer detaljerte analyser av resultatene. Først undersøker vi sammenhengen mellom institusjonstilhørighet og resultat på nasjonal deleksamen, kontrollert for tilgjengelige bakgrunnsvariabler. Deretter presenterer vi analyser av sensurreliabiliteten.

### 3 Analyser

I denne delen av rapporten undersøker vi sammenhengen mellom resultatet på nasjonal deleksamen og tilgjengelige bakgrunnsvariabler, samt institusjonsbidrag. Vi gjør også analyser av sensorreliabiliteten. Vi starter med en kort gjennomgang av datagrunnlaget.

#### 3.1 Om dataene

I sensurarbeidet brukte sensorene et vurderingsskjema hvor de skulle rapportere poeng på hver deloppgave for hver kandidat, og kandidatens poengsum. Poengsum varierer mellom 0 og 100. Hver student fikk besvarelsen sin sensurert av to sensorer. Vurderingsskjemaene viser både antall poeng kandidatene fikk på hver deloppgave fra hver sensor, samt felles poeng (samsensur) per deloppgave som sensorene ble enige om. Disse poengene ble summert til en poengsum. Vurderingsskjemaene viser også kandidatens karakterer på skalaen A til F. Karakter ble satt på bakgrunn av poengsum. I de videre analysene av resultatene tar vi utgangspunkt i antall poeng per deloppgave, ettersom disse variablene inneholder mer detaljert informasjon enn karakteren.

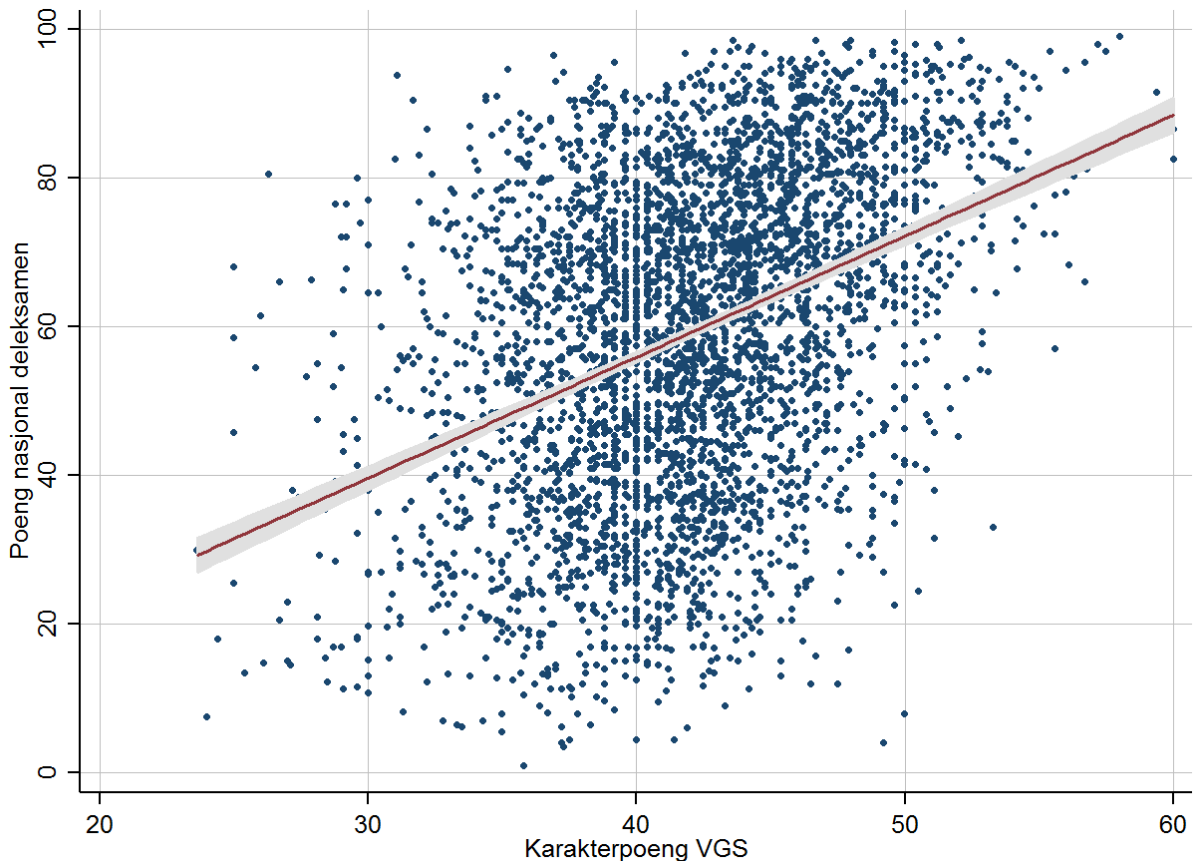
Innholdet i alle vurderingsskjemaene er satt sammen til et datasett hvor det også er koblet på bakgrunnsinformasjon om kandidatene, som kandidatenes kjønn, alder og karakterpoeng fra videregående skole.<sup>3</sup> Datasettet inneholder ikke komplett informasjon om alle de 5 062 kandidatene. For eksempel mangler karakterpoeng fra videregående opplæring for flere kandidater. I tillegg er det noen kandidater som er kodet til «missing» på flere av variablene fra vurderingsskjemaet. Dette gjelder poeng på deloppgaver og poengsum. Dette skyldes enten mangelfull utfylling av vurderingsskjemaet, eller tekniske feil i utfyllingen som gjorde det vanskelig å lese av innholdet automatisk ved genereringen av datasettet. Det er viktig å understreke at disse feilene ikke har hatt innvirkning på selve sensureringen og fastsettelsen av endelig karakter.

### **3.2 Karakterpoeng fra videregående skole**

Korrelasjonen mellom karakterpoeng fra videregående opplæring og poengsum på nasjonal deleksamen er positiv og signifikant forskjellig fra null ( $r=0,38$ ,  $p<0,001$ ). En regresjonsanalyse (ordinary least square estimator) med poengsum på nasjonal deleksamen som avhengig variabel og karakterpoeng fra videregående skole som uavhengig variabel, gir en statistisk signifikant sammenheng med en koeffisient på 1,6. Det betyr at en økning på ett karakterpoeng fra videregående skole gir en økning på 1,6 poeng på nasjonal deleksamen, som et gjennomsnitt for alle kandidatene. Karakterpoeng fra videregående opplæring forklarer 15 prosent av variasjonen i poengsum på nasjonal deleksamen. Karakterpoeng fra videregående opplæring er en viktig forklaringsfaktor når det gjelder resultatene på nasjonal deleksamen, men andre faktorer står allikevel for en langt større del av variasjonen. Sammenhengen er vist i figur 3. Hvert punkt er en kandidat som tok nasjonal deleksamen, og figuren viser regresjonslinjen med 95 prosent konfidensintervall.

---

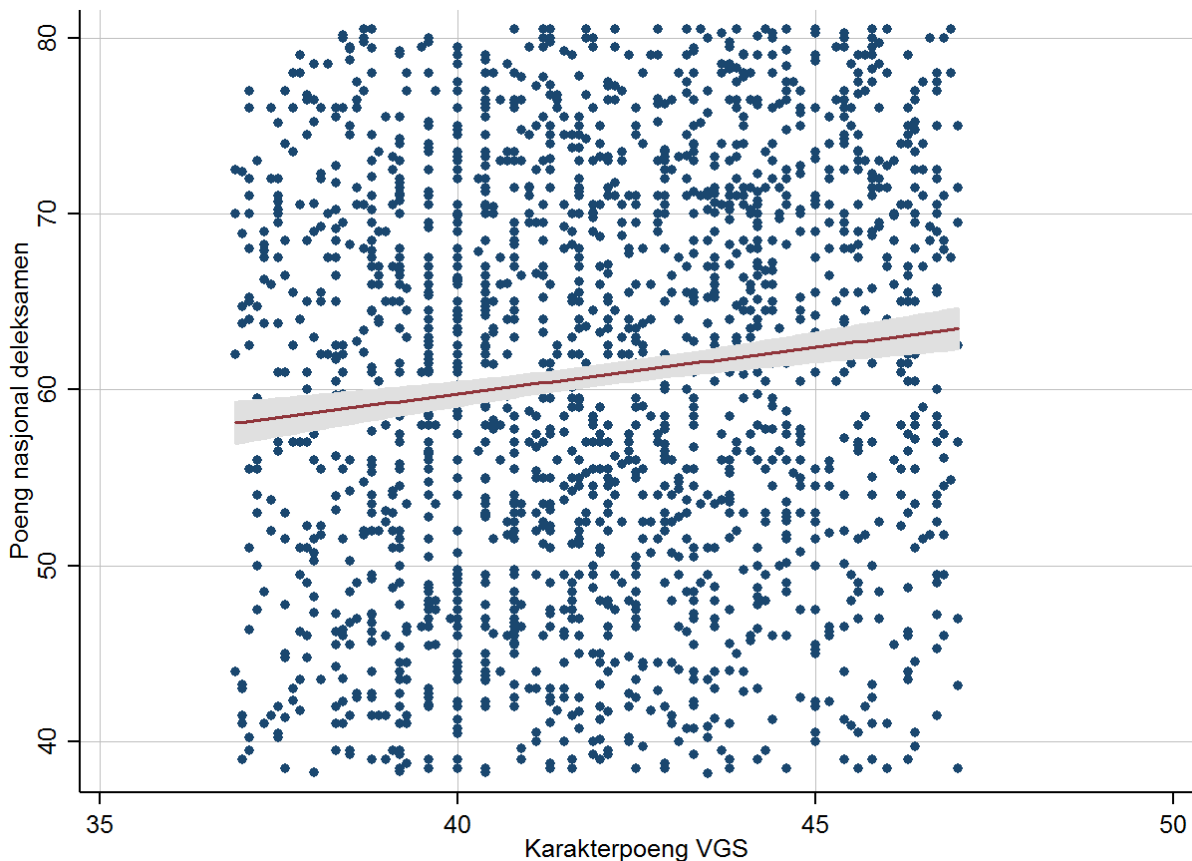
<sup>3</sup> Karakterpoeng fra VGS varierer mellom 20 og 60 poeng. Kandidatene i datasettet har i gjennomsnitt 41,9 karakterpoeng (std. avvik 5,1). Gjennomsnittsalderen til kandidatene var 23,5 år (std. avvik 6,1), og 87 prosent var kvinner.



Figur 3: Sammenhengen mellom karakterpoeng fra videregående skole og poengsum på nasjonal deleksamen i 2017. Regresjonslinje med 95 prosent konfidensintervall.  $N=4086$ .

Hvis man ser på fordelingen i figur 3, kan en visuell tolkning tyde på at sammenhengen mellom karakterpoeng fra videregående og poengsum på nasjonal deleksamen ikke er særlig sterk for de kandidatene som ligger rundt gjennomsnittet på begge variablene. Det ser ut til at mye av stigningen til regresjonslinjen påvirkes av de kandidatene som ligger et stykke utenfor gjennomsnittet. Dette gjelder særlig de som har mange karakterpoeng fra videregående. Det kan derfor være interessant å undersøke denne sammenhengen for bare de kandidatene som ligger rundt gjennomsnittet på begge variablene.

I Figur 4 har vi gjentatt analysen i Figur 3, men fjernet alle kandidater som ligger utenfor gjennomsnittet pluss/minus ett standardavvik, både på karakterpoeng og på poengsum på nasjonal deleksamen. Visuelt sett ser det ikke ut til å være noen sammenheng. Regresjonsmodellen viser at en økning på ett karakterpoeng fra videregående skole nå bare gir en økning på 0,53 poeng på nasjonal deleksamen. Selv om sammenhengen også her er statistisk signifikant, forklarer modellen bare 1,3 prosent av variasjonen i poengsum på nasjonal deleksamen.



Figur 4: Sammenhengen mellom karakterpoeng fra videregående og poengsum på nasjonal deleksamen i 2017. Lineær regresjonslinje med 95 prosent konfidensintervall. Kun kandidater innenfor gjennomsnittet +/- ett standardavvik på begge variablene. N=1865.

### 3.3 Institusjonsbidrag

I det følgende analyserer vi sammenhengen mellom kandidatenes studiested og poengsum på nasjonal deleksamen. De ulike institusjonene og studiestedene kan ha en ulik sammensetning av studenter. Vi kontrollerer derfor for kandidatenes alder, kjønn og karakterpoeng fra videregående opplæring. Det er viktig å understreke at analysen kun er egnet til å gi en indikasjon på studiestedenes bidrag til resultatet. Det er antageligvis andre faktorer som også spiller inn, og som vi ikke har mulighet til å kontrollere for i modellen.

Modell 1 i tabell 2 viser sammenhengen mellom de tre bakgrunnsvariablene og poengsum på nasjonal deleksamen. Alder og karakterpoeng har positiv sammenheng med poengsum på nasjonal deleksamen. Kvinner oppnådde i gjennomsnitt noe lavere poengsum på deleksamen enn menn.

I modell 2 er kandidatenes institusjon og studiested lagt til som en uavhengig variabel ved hjelp av dummykoding. Dermed kan vi undersøke sammenhengen mellom studiested og resultatet på nasjonal deleksamen, kontrollert for bakgrunnsvariablene alder, kjønn og karakterpoeng fra videregående. Ved dummykoding må alltid kategoriene (som her er institusjon-studiested) sammenlignes med en referansekategori (base). Vi valgte studiestedet med flest kandidater, HiOA Pilestredet, som

referansekategori. Referanse kategorien ligger også nærmest det nasjonale gjennomsnittet på den avhengige variabelen. I tabellen er studiestedene sortert etter verdien på koeffisienten.

Resultatene viser at kandidater ved syv studiesteder har et signifikant høyere resultat på nasjonal deleksamen sammenlignet med referanse kategorien, kontrollert for bakgrunnsvariablene. Den sterkeste positive sammenhengen mellom studiested og resultat finner vi for NTNU Trondheim. Åtte studiesteder har et signifikant lavere resultat. Den sterkeste negative sammenhengen mellom studiested og resultat finner vi for Høgskolen i Molde. For resten av studiestedene er det ingen statistisk signifikant forskjell fra referanse kategorien.

Som nevnt bør disse resultatene forstås som en *indikasjon* på studiestedenes bidrag til resultatet på nasjonal deleksamen. Vi bør derfor utvise forsiktighet ved tolkninger av data og bruk av resultater. Vi tenker allikevel at resultatene er interessante. Institusjonene kan sammenligne seg med hverandre, og institusjoner kan sammenligne sine ulike studiesteder. Vi håper det gir en mulighet for fagmiljøene til å utveksle erfaringer og lære av hverandre.

Tabell 2: Regresjonsmodell (OLS) med poengsum (0-100) på nasjonal deleksamen som avhengig variabel.

	Modell 1		Modell 2	
	b	std.feil	b	std.feil
Konstant	-26,92***	3,37	-16,51***	3,86
Alder	0,62***	0,07	0,59***	0,07
Kvinne	-3,07***	0,95	-2,12*	0,94
Karakterpoeng VGS	1,79***	0,06	1,54***	0,07
Institusjon				
NTNU - Trondheim			8,13***	1,75
Høgskolen i Innlandet - Elverum			6,26***	1,71
HVL - Førde			6,07**	2,12
UiA - Kristiansand			5,31**	1,83
VID - Bergen			4,15*	1,83
Lovisenberg Diakonale Høgskole			4,10*	1,65
Høgskolen i Østfold			3,94*	1,73
UiA - Grimstad			3,40	2,07
NTNU - Ålesund			2,52	2,05
VID - Oslo			2,11	1,82
HVL - Bergen			1,73	1,67
HVL - Haugesund			1,47	2,47
Nord - Levanger			1,39	2,06
Høgskolen Diakonova			0,88	2,18
HiOA - Pilestredet			(ref.)	-
Høgskolen i Innlandet - Kongsvinger			-1,33	2,99
HiOA- Sandvika			-1,84	3,65
Nord - Bodø			-2,35	1,93
HSN - Porsgrunn			-2,67	1,99
HSN - Vestfold			-2,98	2,07
UiT - Narvik			-3,10	3,68
UiT - Hammerfest			-3,20	2,73
NTNU - Gjøvik			-3,44	1,75
HSN - Drammen			-3,92*	1,98
UiT - Harstad			-4,01	2,54
Nord - Vesterålen			-4,35	3,52
UiT - Tromsø			-4,42*	1,90
HiOA - Kjeller			-5,63**	1,84
Nord - Namsos			-5,75*	2,28
Nord - Mo i Rana (Helgeland)			-6,66**	2,33
HVL - Stord			-7,33**	2,38
UiS			-7,62***	1,66
Høgskolen i Molde			-8,87***	2,32

Modell 1: Justert  $R^2=0,17$ . Modell 2: Justert  $R^2=0,21$ . Antall obs. (begge modeller): 4086.

\* $p<0,05$ , \*\* $p<0,01$ , \*\*\* $p<0,001$ .

## 4 Sensurreliabilitet

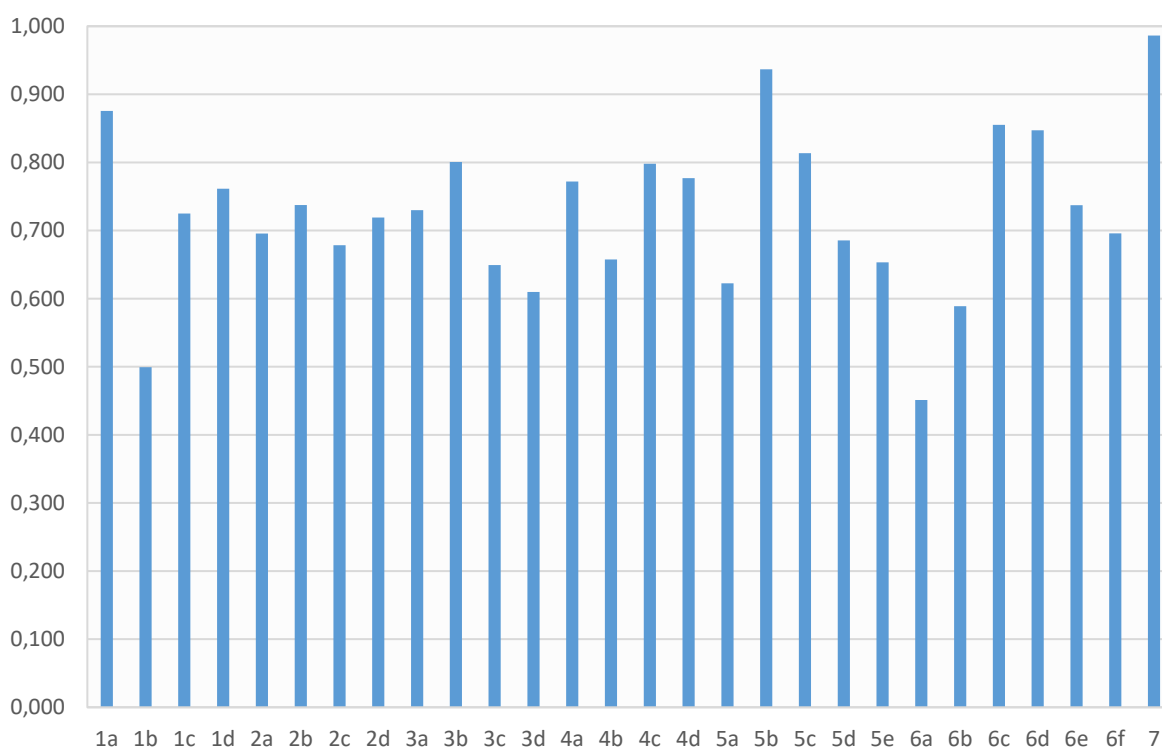
Vi brukte Cohens Kappa, også kalt Kappa-koeffisient, for å måle sensurreliabiliteten. Kappa tar utgangspunkt i graden av enigheten mellom sensorene, men tar hensyn til den underliggende sannsynligheten for at sensorene gir en lik vurdering ved ren tilfeldighet. Kappa er derfor bedre egnet til å måle graden av enighet mellom to sensorer, sammenlignet med å kun beregne prosentvis enighet direkte. Kappa kan beregnes med og uten vektning. Uvektet Kappa er mest aktuelt for nominale variabler, altså når verdiene ikke kan rangeres. Vi har beregnet vektet Kappa (med lineær vektning). Beregningen tar da hensyn til avstanden mellom sensorenes vurderinger. For eksempel vil tilfeller hvor sensor X gir 5 poeng og sensor Y gir 2 poeng, føre til en lavere Kappa enn dersom sensor X gir 5 poeng og sensor Y gir 4 poeng. Kappa varierer fra 0 (ingen enighet) til 1 (perfekt enighet). Følgende skala<sup>4</sup> er ofte brukt:

- 0,0–0,20: svak enighet
- 0,21–0,40: rimelig enighet
- 0,41–0,60: moderat enighet
- 0,61–0,90: betydelig enighet
- 0,91–1,0: nesten perfekt enighet

Figuren under viser Kappa for hver deloppgave og for poengsum på deleksamen. Beregningen er gjort på alle kandidatene i datasettet. Det innebærer at det ikke er de samme to sensorene som har vurdert alle kandidatene, men for hver kandidat er det de samme to sensorene som har vurdert hver oppgave. Denne fremgangsmåten avviker fra en forutsetning for Kappa, nemlig at dataene skal inneholde vurderinger fra de samme to sensorene for alle observasjoner. Resultatet bør derfor tolkes med noe forsiktighet, men er allikevel egnet til å gi en indikasjon på graden av enighet mellom sensorene for de ulike oppgavene. For alle oppgavene er det minst en moderat enighet mellom sensorene, og for 25 av 28 oppgaver er enigheten enten betydelig eller nesten perfekt.

---

<sup>4</sup> Landis, J. & Koch, G. (1977). «The Measurement of Observer Agreement for Categorical Data» *Biometrics* 33: 159–174.

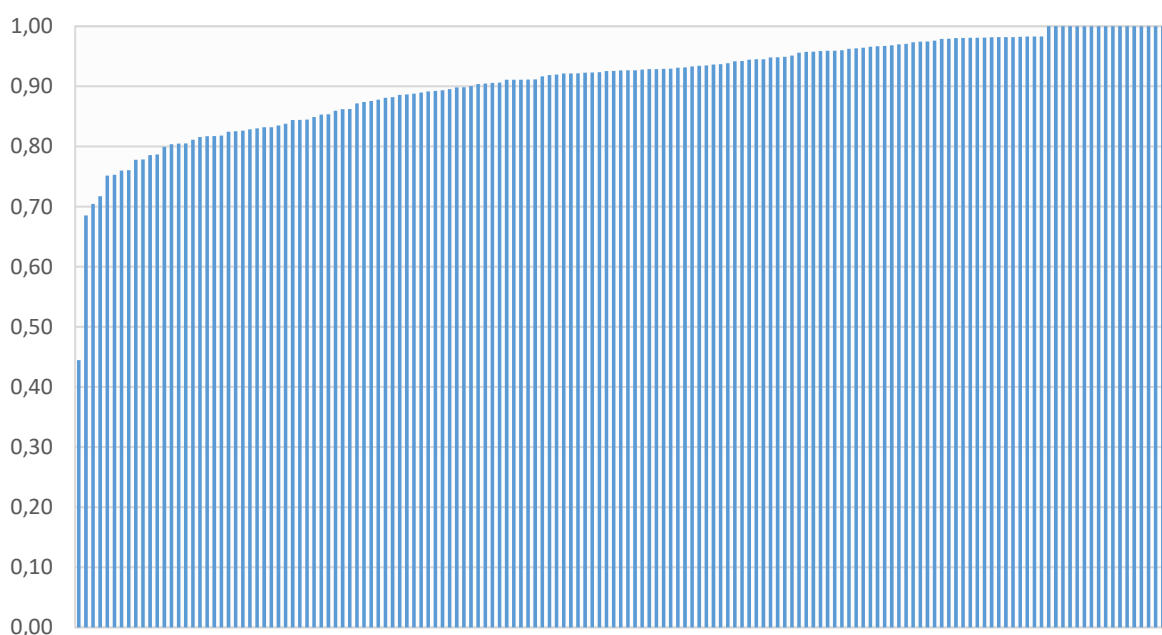


Figur 5: Vektet (lineær) Kappa for hver deloppgave og for poengsum på deleksamen.

Vi har også brukt Kappa for å måle graden av enighet mellom sensorene i hvert sensorpar. Sensorveiledningen inneholdt en karakterskala hvor det ble angitt hvilken karakter på skalaen A til F som skulle gis ut i fra totalt antall poeng som kandidaten fikk, men det var åpent for at sensorene kunne gjøre skjønsmessige vurderinger. Vi kodet poengsum (0-100) fra begge sensorene om til en beregnet karakter (F-A) ut i fra denne karakterskalaen, og brukte disse omkodede variablene for å måle graden av enighet. Resultatet er vist i figuren under, hvor hver søyle representerer Kappa-verdien for ett sensorpar, basert på alle kandidatene som sensorparet sensurerte sammen. Det var 159 sensorpar som sensurerte besvarelsene i nasjonal deleksamen. Figuren viser resultatet for 154 sensorpar. Det skyldes ulike tekniske feil knyttet til utfyllingen av vurderingsskjemaene og avlesning av data<sup>5</sup>. Analysen viser at det var nesten perfekt enighet for flertallet av sensorparene. For resten av sensorparene var det betydelig enighet, med unntak av ett sensorpar som oppnådde moderat enighet.

<sup>5</sup> Les «om dataene» tidligere i rapporten for mer informasjon.





Figur 6: Vektet (lineær) Kappa for hvert sensorpar når det gjelder total poengsum.

I krysstabellen under vises den beregnede karakteren fra hver sensor i de ulike sensorparene. Tallene er i prosent og tar utgangspunkt i data for alle kandidatene. Vi ser da for eksempel at 19,9 prosent av kandidatene fikk en poengsum tilsvarende en C fra begge sensorene. På samme måte ser vi at 1,8 prosent av kandidatene fikk en poengsum tilsvarende en C fra den ene sensoren, og en poengsum tilsvarende en D fra den andre sensoren. For rundt 82 prosent av kandidatene kom begge sensorene frem til en poengsum som gir lik karakter ifølge karakterskalaen.

Tabell 3: Sammenligning av beregnet karakter fra hver sensor for alle kandidatene. Tallene er prosent. N=4901.

	F	E	D	C	B	A	Totalt
F	<b>20,0</b>	0,9	0,0	0,0	0,0	0,0	20,9
E	1,5	<b>9,9</b>	1,8	0,0	0,0	0,0	13,1
D	0,0	1,5	<b>10,4</b>	1,8	0,0	0,0	13,8
C	0,0	0,1	2,2	<b>19,9</b>	2,8	0,0	24,9
B	0,0	0,0	0,0	2,5	<b>16,7</b>	1,2	20,4
A	0,0	0,0	0,0	0,0	1,5	<b>5,4</b>	6,9
Totalt	21,5	12,4	14,4	24,1	21,0	6,6	<b>100,0</b>

## 5 Klagesensur

Nasjonal deleksamen følger UH-loven i spørsmål om krav om begrunnelse og krav om ny sensur. 681 kandidater ba om begrunnelse for karaktergivning. Begrunnelser ble sendt fra NOKUT i form av sensorenes samsensur på hver deloppgave sammen med sensorveiledningen. Slik kunne kandidatene selv gå gjennom sin besvarelse og se hvordan svarene er vurdert.

360 kandidater klaget på karakteren. I tabell 4 er det angitt antall krav om ny sensur fordelt på opprinnelige karakterer, og hvordan besvarelsene ble vurdert etter ny sensur.

Tabell 4: Oversikt over ny sensur

Opprinnelig karakter	Antall	Antall forbedret karakter	Antall dårligere karakter	Opp 1 karakter	Ned 1 karakter	Opp mer enn 1 karakter
B	13	2	1	2	1	0
C	24	12	1	12	1	0
D	22	6	4	6	4	0
E	12	6	3	6	3	0
F	289	38	0	33	0	5
SUM	360	64	9	59	9	5

Oppsummert var det 73 kandidater (20,3 prosent av de som klaget) som fikk endret karakter, mens 287 kandidater fikk uendret karakter etter klagesensur.

## 6 Eksamensoppgaver og sensorveiledning

Etter at hovedsensuren var avsluttet, sendte vi ut en spørreundersøkelse til sensorene. 62 av 80 sensorer besvarte undersøkelsen. Undersøkelsen besto av fem spørsmål som skulle besvares på en skala fra 1 til 6. Sensorene ble blant annet bedt om å vurdere kandidatens faglige nivå og hvor tydelig eksamensoppgavene var formulert. Videre ble de bedt om å besvare seks åpne spørsmål om eksamensoppgavene, sensorveiledningen og organiseringen av sensuren, med mulighet for ytterligere kommentarer. Eksamenssettets vanskegrad ble vurdert til 3,2 (6 betyr vanskelig), og på spørsmål om eksamenssettet var dekkende for emnet var samlet vurdering 4,7 (6 betyr dekkende).<sup>6</sup>

Fra sensorenes svar kommer det frem at et flertallet mente at eksamenssettet var tydelig formulert, men at enkelte oppgaver kunne vært mer presise. Noen oppgaver ble vurdert til å ha for lav vanskegrad. Hvis vi eksempelvis ser på figur 5 ovenfor, er det størst uenighet blant sensorene på oppgavene 1b og 6a, til dels også 6b. Disse oppgavene tester allmennkunnskap, og det blir et spørsmål om man skal kreve vitenskapelige begreper eller også godkjenne svar som bruker dagligdags språk uten tekniske fagbegreper. En slik situasjon åpner for at sensorene følger noe forskjellige kriterier for

<sup>6</sup> Eksamensoppgaver og sensorveiledninger kan finnes på <https://www.nokut.no/norsk-utdanning/hoyere-utdanning/nasjonal-deleksamen/>

poenggivning. Likevel, en enkel Rasch-analyse viser at poenggivningen fra samsensuren på disse oppgavene enkeltvis svarer godt til kandidatenes totale skår på besvarelsen.

Sensorene omtalte generelt sensorveiledningen som tydelig, og de vurderte den til å være et godt arbeidsredskap. Spørreundersøkelsen viste at det er noe uenighet om hva de faglige kriteriene burde ha vært, men vi har ikke kunnskap om i hvilken grad dette påvirket sensurarbeidet. En del mente at detaljstyringen var for høy, mens andre igjen savnet større presisjon.

Kvalitative vurderinger av eksamensoppgaver/-besvarelser og sensorveiledningen,<sup>7</sup> viser at det gjennomgående er uklart hvilke krav som stilles til delvis skår, som for eksempel 1 poeng på en oppgave som gir opptil 2 poeng. Dette problemet blir større når oppgaven kan gi flere poeng. Spørsmålet er om man kan måle suksessivt større faglig kompetanse per poengnivå. Man bør vurdere en begrenset poengtilordning hvor kun heltall er mulig, og en sensorveiledning som tydelig skiller mellom «ingen kreditt», «delvis kreditt» og «full kreditt», hvor uttellingen med større presisjon gjenspeiler reelle kompetansenivåer.

## 7 Konklusjon

Vi har i denne rapporten sammenstilt nasjonale resultater fra 2016 og 2017. Resultatene fra pilottestingen i 2015 er ikke tatt med siden denne gjennomføringen ikke var i fullskala-format. De nasjonale resultatene i 2017 viser liten endring fra året før. Andelen kandidater med F (21,5 %) og E (12,5 %) er tilnærmet samme nivå som året før. Fordelingen på de øvrige fire karakterene viser en liten forskyvning mot B og C. Andelen A (6,4 %) har sunket med 0,9 prosentpoeng.

På samme måte som i fjorårets rapport, viser analysen at karakterpoeng fra videregående til en viss grad kan predikere hvor godt kandidatene skårer på nasjonal deleksamen. Målt på hele kullet, forklarer det likevel ikke mer enn 15 % av variasjonen i poengsum på eksamen. Når vi fjerner kandidatene med høyest og lavest karakterpoeng (fra videregående) og poengsum (på nasjonal deleksamen), så vi at karakterpoeng i liten grad predikerer resultatet på nasjonal deleksamen.

Når det gjelder institusjonsbidraget, fant vi forskjeller mellom institusjonene i resultatet på nasjonal deleksamen, selv når vi kontrollerte for forskjeller i kandidatenes alder, kjønn og karakterpoeng fra videregående. Resultatet bør kun tolkes som en indikasjon på institusjonsbidraget, men det kan være relevant for institusjonene å bruke dette som et utgangspunkt for sammenligning og erfaringsdeling.

Analysen av sensurreliabilitet viser stor grad av enighet mellom sensorene. Ett sensorpar skiller seg ut med moderat enighet. Det er tre deloppgaver i oppgavesettet hvor sensorene oppnår bare moderat enighet.

NOKUT har stått ansvarlig for gjennomføringen av nasjonal deleksamen i anatomi, fysiologi og biokjemi tre ganger. Logistikkmessig har det fungert godt, all den tid det er en manuell papirbasert eksamen der besvarelsene er distribuert til sensorkorpset gjennom en nettportal. En eventuell

---

<sup>7</sup> Dette er momenter fra en analyse utført av Øystein Gutterud, Naturfagssenteret ved UiO

overføring til digitalt format og fulldigitalisert system for sensurbehandling, vil kunne gjøre gjennomføringen smidigere og sikrere.

Nasjonale deksamener for bachelorgradsstudiet i sykepleie har bidratt til tettere samarbeid mellom fagmiljøene, og et skarpere fokus på utvikling av eksamensoppgaver og sensorveiledning. Nasjonale deksamener bærer med seg et potensial for å bidra til kvalitetsarbeid lokalt og nasjonalt.

## Vedlegg 1: Karakterfordelingen per institusjon og studiested, 2016 og 2017.

Institusjon	A		B		C		D		E		F		Gj.snitt		Antall	
	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017
HiOA - Kjeller	6,4	3,3	22,8	19,9	24,6	16,6	19,3	13,8	11,1	17,1	15,8	29,3	2,5	1,9	171	181
HiOA - Pilestredet	5,7	7,3	16,7	18,9	21,1	26,9	16,3	14,6	13,6	8,6	26,6	23,7	2,0	2,3	478	465
HiOA- Sandvika	3,5	5,6	3,5	19,4	27,6	27,8	27,6	5,6	6,9	22,2	31,0	19,4	1,8	2,2	29	36
HSN - Drammen	5,6	5,7	14,4	20,1	18,9	26,4	17,2	12,6	10,0	13,2	33,9	22,0	1,9	2,3	180	159
HSN - Porsgrunn	2,2	3,8	11,1	13,4	19,3	29,9	14,9	15,9	13,3	12,7	39,2	24,2	1,6	2,1	181	157
HSN - Vestfold	4,5	5,9	13,5	21,5	27,8	20,0	15,0	16,3	13,5	8,9	25,5	27,4	2,0	2,2	133	135
HVL - Bergen	15,2	7,0	36,4	25,2	28,3	27,9	9,2	14,7	5,4	11,6	5,4	13,6	3,3	2,6	184	258
HVL - Førde	15,3	4,5	22,2	27,7	30,6	25,0	16,7	14,3	9,7	16,1	5,6	12,5	3,0	2,5	72	112
HVL - Haugesund	3,9	3,8	12,5	17,9	24,0	34,9	14,4	9,4	12,5	16,0	29,8	17,9	1,9	2,3	101	106
HVL - Stord	2,2	2,3	6,7	9,2	25,6	20,7	15,6	17,2	13,3	25,3	33,3	25,3	1,6	1,7	87	87
Høgskolen i Innlandet - Elverum*	5,9	7,3	17,6	20,9	30,5	34,6	18,0	9,1	11,3	10,9	16,8	17,2	2,4	2,5	256	220
Høgskolen i Innlandet - Kongsvinger*	5,9	1,9	17,6	11,1	30,5	29,6	18,0	18,5	11,3	14,8	16,8	24,1	2,4	1,9	256	54
Høgskolen i Molde	5,5	4,3	14,1	9,6	22,1	20,0	21,5	14,8	10,4	12,2	26,4	39,1	2,0	1,6	163	115
Høgskolen i Østfold	3,5	9,7	19,8	24,5	25,0	22,7	18,0	14,4	9,9	9,3	23,9	19,4	2,2	2,5	172	216
Høgskolen Diakonova	4,7	4,7	29,1	25,2	26,8	27,5	15,0	11,0	7,1	11,0	17,3	20,5	2,6	2,4	127	127
Lovisenberg Diakonale Høgskole	13,3	14,5	28,4	33,9	24,0	23,0	16,0	12,1	8,5	6,5	9,8	10,1	2,9	3,1	225	248
Nord - Bodø	3,0	1,9	12,1	17,1	20,5	22,2	10,2	17,1	16,9	18,4	37,4	23,4	1,6	2,0	166	158
Nord - Levanger**	5,2	5,2	16,1	21,3	22,3	29,4	18,0	19,9	19,9	12,5	18,5	11,8	2,1	2,5	211	136
Nord - Mo i Rana (Helgeland)	3,9	0,0	6,4	5,0	15,4	18,8	14,1	20,8	15,4	7,9	44,9	47,5	1,3	1,3	78	101
Nord - Namsos**	5,2	0,0	16,1	8,0	22,3	18,0	18,0	13,0	19,9	24,0	18,5	37,0	2,1	1,4	211	100
Nord - Vesterålen	0,0	2,5	15,0	10,0	30,0	25,0	15,0	12,5	10,0	7,5	30,0	42,5	1,9	1,6	20	40

\* I 2016 splittet vi ikke resultatene i Elverum og Kongsvinger. Tallene for 2016 er totalen for begge utdanningene

\*\* I 2016 splittet vi ikke resultatene i Levanger og Namsos. Tallene for 2016 er totalen for begge utdanningene

Institusjon	A		B		C		D		E		F		Gj.snitt		Antall	
	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017	2016	2017
NTNU - Gjøvik	4,3	4,6	15,2	18,5	25,1	23,9	14,7	12,2	14,7	15,8	26,0	15,8	2,0	2,1	231	222
NTNU - Trondheim	26,4	21,1	29,4	36,3	20,3	20,1	8,6	9,8	7,1	5,4	8,1	7,4	3,4	3,4	197	204
NTNU - Ålesund	7,5	5,2	18,3	14,0	25,8	34,6	19,2	19,1	15,0	12,5	14,2	14,7	2,4	2,4	120	136
UiA - Grimstad	4,6	3,1	19,7	21,7	25,8	32,6	20,5	10,9	17,4	17,1	12,1	14,7	2,4	2,4	132	129
UiA - Kristiansand	12,0	14,9	28,0	27,4	26,7	30,3	18,0	8,6	5,3	8,0	10,0	10,9	2,9	3,0	150	175
UiS	1,8	2,4	11,8	10,6	22,1	25,6	20,6	20,1	17,3	13,0	26,5	28,4	1,8	1,8	272	254
UiT - Hammerfest	4,0	1,5	12,0	7,6	17,0	16,7	18,0	12,1	18,0	16,7	31,0	45,5	1,7	1,3	100	66
UiT - Harstad	1,5	3,1	10,8	13,4	27,7	11,3	7,7	16,5	16,9	10,3	33,9	45,4	1,7	1,5	64	97
UiT - Narvik	4,0	0,0	8,0	17,0	14,0	10,6	22,0	12,8	18,0	27,7	34,0	31,2	1,6	1,5	50	47
UiT - Tromsø	9,1	5,5	21,2	18,3	27,3	26,2	15,9	11,0	9,1	12,8	17,4	26,2	2,5	2,1	132	164
VID - Bergen	14,1	9,0	28,2	26,0	27,1	32,2	8,8	10,7	7,7	13,0	14,1	9,0	2,9	2,8	170	177
VID - Oslo	10,6	5,6	21,2	24,4	30,2	22,8	14,0	17,8	8,9	15,6	15,1	13,9	2,7	2,5	179	180
Nasjonalt	7,3	6,4	18,7	20,2	24,1	25,4	15,9	14,0	12,1	12,5	21,8	21,5	2,3	2,3	4 838	5 062