



# «Blind» klagesensur ved Universitetet i Oslo

*Pålitelig og rettferdig eller et steg i feil retning?*

## «Blind» appeals of grades at the University of Oslo

*Reliable and fair or a step in the wrong direction?*

Vidar Gynnild

*Professor, Norges teknisk-naturvitenskapelige universitet (NTNU)*

[Vidar.Gynnild@ntnu.no](mailto:Vidar.Gynnild@ntnu.no)

### Sammendrag

Denne artikkelen beskriver og analyserer virkningene av «blind» klagesensur ved Universitetet i Oslo. Før 01.08.14 hadde de nye sensorene ved sensurklage tilgang til resultatene fra ordinær sensur, mens denne muligheten ble tatt bort da «blind» klagesensur ble tatt i bruk fra og med høsten 2014. Undersøkelsen er gjennomført med utgangspunkt i registerdata for to ulike perioder, før og etter at «blind» klagesensur ble innført. Undersøkelsen viser at klage-tilbøyeligheten har økt over tid, uavhengig av sensurordning. Samtidig har «blind» sensur ført til større avvik og større spredning sammenliknet med ordinær sensur. Nyordningen har samtidig ført til en jevnere fordeling mellom endringer i positiv og negativ retning. Norsk Studentorganisasjon var en sentral pådriver for innføring av «blind» klagesensur på nasjonal basis. De ønsket med dette en større grad av likebehandling, og dermed en mer rettferdig vurderingsordning. Denne undersøkelsen viser at nyordningen neppe har vært til studentenes gunst.

### Nøkkelord

Vurdering, klage, sensur, taus kunnskap, skjønn

### Abstract

This study explores the effects of “blind” grade appeals processes at the University of Oslo. Prior to August 1 2014, examiners in charge of appeals processes were granted full access to all information related to ordinary grading; however, this was changed when “blind” appeals procedures were mandated by the Ministry in 2014. The study draws on register data from two different periods, before and after the introduction of “blind” appeals. The analysis confirms that the frequency of appeals has increased over time, unaffected by procedures. However, “blind” appeals resulted in increased discrepancies compared with grades awarded during the ordinary grading processes. At the same time, the new appeals procedures have resulted in a more even distribution of grade changes both in positive and negative directions. The Norwegian Student Association was a driving force to adopt “blind” appeals procedures across all higher education institutions. This study confirms that this decision has not been in the best interest of students in terms of generally achieving improved grades.

### Keywords

Assessment, grade appeals, implicit knowledge

## Innledning

Svært mye forskning dokumenterer betydningen av vurdering og eksamen i høyere utdanning (Brown & Knight, 1994; Rust, 2002; Sadler, 1987, 2010; Snyder, 1971). Vurdering står som regel sentralt i alle faser av studiene, selv om formål og metode varierer. Slutteksamen har vært den mest utbredte vurderingsordningen, men etter årtusensskiftet har ulike typer formativ vurdering, for eksempel midtsemesterprøve og mappevurdering fått større utbredelse. I de senere årene har vi også sett et større mangfold av vurderingsformer i kombinasjon for å ivareta både summative og formative funksjoner, stadig oftere med bruk av digital teknologi.

Denne artikkelen undersøker først regelverket for sensur og klage, spesielt bakgrunnen for at en felles, nasjonal ordning med såkalt «blind» klagesensur ble innført fra 1. august 2014. Før dette benyttet institusjonene ulike ordninger med eller uten «blind» klagesensur. Universitetet i Oslo var blant de institusjonene som ikke benyttet «blind» klagesensur, noe som innebar at klagesensorene både fikk kjennskap til opprinnelig karakter og kandidatens begrunnelse. Norsk studentorganisasjon (NSO) var imidlertid en sterk pådriver for en ordning der klagesensorene ikke skulle få informasjon fra ordinær sensur, og heller ingen begrunnelse for klagen fra klager. NSO mente derfor at «blind» klagesensur ville sikre nøytral vurdering (Mortensen, 2016).

Behandling av *formelle feil* ved eksamen og eksamensavvikling er regulert ved en egen bestemmelse i loven der det gis rett til ny sensur hvis det er påvist feil som kan ha hatt betydning ved vurdering av kandidatens prestasjon (se UH-loven § 5-2). Sensuren kan i tillegg påklages selv om ikke formelle feil eller mangler er påvist.

Nyordningen, som trådte i kraft 01.08.14, innebar at de nye sensorene ikke lenger fikk tilgang til opprinnelig karakter, ei heller kandidatens begrunnelse for klagen. Klagesensur skulle behandles på samme måte som ved første gangs sensur, og dermed i teorien bidra til en mer rettfærdig vurdering. Utfordringene ved «blind» klagesensur har imidlertid vist seg å være betydelige, for eksempel at rettssikkerheten for kandidatene blir svekket ved dårligere opplysning om saken. Kandidatene får ikke anledning til å presentere sine innvendinger mot den påklagde bedømmelsen, og klagesensorene har ikke mulighet til å gjøre seg kjent med og ettergå premisset for den opprinnelige karakteren ved ordinær sensur.

## Data, problemstillinger og metode

Artikkelforfatteren ønsket å gjennomføre en undersøkelse rettet mot effekten av «blind» klagesensur. Det var i den anledning fristende å gjennomføre en sammenliknende studie med utgangspunkt i sensurdata før og etter at «blind» sensur ble innført, og Universitetet i Oslo (UiO) ble valgt på bakgrunn av størrelse og faglig bredde. Forfatteren ble møtt med velvilje fra Seksjon for studieadministrative datasystemer, som selv hadde interesse av undersøkelsen. Artikkelen bygger på data fra og med våren 2006 til og med høsten 2017, og er gjennomført som en sammenliknende studie mellom to perioder – før og etter 1. august 2014. Før dette ble all klagesensur ved UiO gjennomført som «åpen» sensur der opprinnelig karakter og annen informasjon ble gjort tilgjengelig for klagesensorene. Den første perioden er noe lengre enn den andre, men datagrunnlaget er i begge tilfeller meget robust. Analyser av registerdata blir presentert i form av deskriptiv statistikk. Når ikke annet er nevnt, er data fra alle fakultetene brukt, mens data fra Det odontologiske fakultet først ble tilgjengelig fra våren 2014.

Universitetet i Oslo består av følgende åtte fakulteter: Det historisk-filosofiske fakultet (HF); Det juridiske fakultet (JF); Det matematisk-naturvitenskapelige fakultet (MN); Det medisinske fakultet (MF); Det odontologiske fakultet (OD); Det samfunnsvitenskapelige fakultet (SV); Det teologiske fakultet (TF) og Det utdanningsvitenskapelige fakultet (UV). Forkortelsene blir etter dette benyttet, og forskningsspørsmålene er som følger:

- Hva var bakgrunnen for innføring av «blind» klagesensur på nasjonal basis?
- På hvilken måte har «blind» klagesensur påvirket karaktersettingen sammenliknet med tiden før denne ordningen ble innført høsten 2014?
- Gis det argumenter for å hevde at «blind» sensur har bidratt til en mer pålitelig og rettferdig vurdering ved Universitetet i Oslo?

Vurdering er et redskap for læring, men kan også være avgjørende for karrieremuligheter og yrkesvalg. Ettersom vurdering ikke er en eksakt vitenskap, inngår bruken av skjønn som et vesentlig element. Pålitelighet, gyldighet og rettferdighet er sentrale dimensjoner, og ideelt sett skal samme besvarelse vurderes ut fra felles kriterier og standarder av ulike bedømmere, og med samme resultat. Denne artikkelen undersøker virkningene av ulike regelverk ved behandling av klage på sensurvedtak med utgangspunkt i registerdata fra én institusjon.

## Kort historikk

Fra 1973 til 1989 ble klage på sensurvedtak regulert av en forskrift med hjemmel i forvaltningsloven (Kunnskapsdepartementet, 1988). I forarbeidene til universitetsloven av 1989 ble det imidlertid foreslått at klagebestemmelsene skulle tas inn i universitetsloven. I forhold til forskrift av 1973 var endringene denne gangen av redaksjonell art. Kandidaten kan be om en redegjørelse for karakterfastsettingen, vanligvis innen to uker, muntlig eller skriftlig (Universitetsloven, 1990). Loven gir også mulighet til skriftlig klage over karakterfastsettingen, og dersom klagen ansees som berettiget, skal karakteren endres. Om så ikke skjer, kan studenten be om skriftlig begrunnelse og opprettholde klagen selv om karakteren blir forbedret. Opprettholdes klagen, skal det gjennomføres ny sensur med nye sensorer (Universitetsloven, 1990, s. 13). Karakteren kan endres begge veier, og vedtaket kan ikke påklages.

I universitetslovens § 52 av 1995 er klagesaksbehandlingen ytterligere forenklet i forhold til forvaltningslovens prinsipper for klagebehandling. Kandidaten kan etter dette be om en begrunnelse, inkludert tilgang til de generelle prinsippene som er lagt til grunn for bedømmelsen. Begrunnelsen kan gis skriftlig eller muntlig etter sensors valg. Dersom kandidaten klager på karaktervedtaket uten krav om begrunnelse, gis det krav om ny sensur (Universitetsloven, 1997).

I brev til kunnskapsdepartementet datert 17.06.13 bemerker Norsk studentorganisasjon (NSO) at universitetsloven ikke konkretiserer hvilke opplysninger de nye sensorene skal ha (Oftedal, 2013). NSO viste i den anledning til et brev fra departementet av 10. juni 1999 der det ble lagt vekt på at den nye sensuren burde være mest mulig lik den første: «De nye sensorene skal derfor ikke ha tilgang til andre opplysninger enn de første sensorene» (Kjellemo, 1999). Departementets brev førte til en rekke henvendelser fra institusjoner som mente denne tolkningen var uheldig, med henvisning til forvaltningslovens § 33 femte ledd med krav om klageinstansens opplysning av saken. Det ble også pekt på at departementets forslag var i strid med gjeldende praksis ved flere institusjoner.

I et nytt brev av 31. mai 2001 begrunnet departementet standpunktet med henvisning til universitets- og høyskolelovens § 52 nr. 4 der det heter at «Ny sensurering skal da foretas» (Aalde, 2001). Dette lå til grunn for departementets syn om at «vurderingen skal være så lik vanlig førstegangssensur som mulig» (Kjellemo, 1999), noe som ble utdypet på følgende vis:

Dokumentasjon om førstegangssensorens karakterfastsettelse og begrunnelse er derfor unødvendig for at klagesensorene skal kunne utføre sin oppgave. En ny sensurering uten informasjon om første sensur vil også sikre en helt ubundet vurdering av besvarelsen, og forhindre at det oppstår et inntrykk blant studentene om at klagevurderingene er motivert av lojalitet sensorene i mellom (Aalde, 2001).

Departementet la likevel vekt på at den pålagte løsningen var i strid med praksis flere steder. I tillegg ble det vist til at universitetslovens § 59 slo fast at ny karakter bare skulle fastsettes dersom det var et påtakelig misforhold mellom prestasjon og karakter (Universitetsloven, 1990). Bestemmelsen forutsatte at klagesensorene fikk informasjon om karakter også ved første gangs sensur. Etter dette fant departementet at begge løsninger måtte godtas og lot institusjonene velge ordning etter eget ønske (Aalde, 2001). I sitt brev til departementet i 2013 la NOS vekt på prinsippet om likebehandling, og mente det burde være én felles ordning for klagesensur (Ofteidal, 2013). Departementet innså etter dette at endringer i loven måtte til, og følgende forslag ble foreslått under § 5-3 fjerde ledd:

En kandidat kan klage skriftlig over karakteren for sine egne prestasjoner innen tre uker etter at eksamensresultat er kunngjort. Ny sensurering skal da foretas. Ved ny sensur skal sensorene kun få kunnskap om de generelle retningslinjer som er gitt for bedømmelsen ... (*Høringsnotat – forslag til endringer i universitets- og høyskoleloven*, 2013).

I den endelige lovteksten ble § 5-3 fjerde ledd presisert på følgende vis:

En kandidat kan klage skriftlig over karakteren for sine egne prestasjoner innen tre uker etter at eksamensresultat er kunngjort. Ny sensurering skal da foretas. Ved ny sensurering skal sensorene ikke få opprinnelig karakter, sensors begrunnelse for denne eller studentens begrunnelse for klagen ... (*Universitets- og høyskoleloven*, 2005)

Notatet med forslag til lovendringer ble sendt til 74 høringsinstanser. Blant disse var universiteter og høyskoler, arbeidstakerorganisasjoner, departementer, studentombud og en rekke andre. Blant utdanningsinstitusjonene var svarfordelingen om lag lik for og imot forslag om «blind» sensur, mens blant universitetene var seks av ni imot ordningen. Alle studentorganisasjonene var derimot samstemmig for å innføre «blind» klagesensur. Hele 42 av 74 høringsinstanser hadde ingen synspunkter, men de som uttalte seg mente det burde være én felles praksis. Få høringsinstanser begrunnet sitt syn teoretisk eller prinsipielt.

Det juridiske fakultet ved Universitetet i Bergen uttalte følgende: «En blind sensur er ingen klageordning, men bare en frittstående «second opinion», og gir verken økt kvalitet ved bedømmelsen eller noen meningsfull tilbakemelding til studenten, førstehånds sensor eller institusjonen» (Hørings svar fra Det juridiske fakultet, Universitetet i Bergen, 16.09.13).

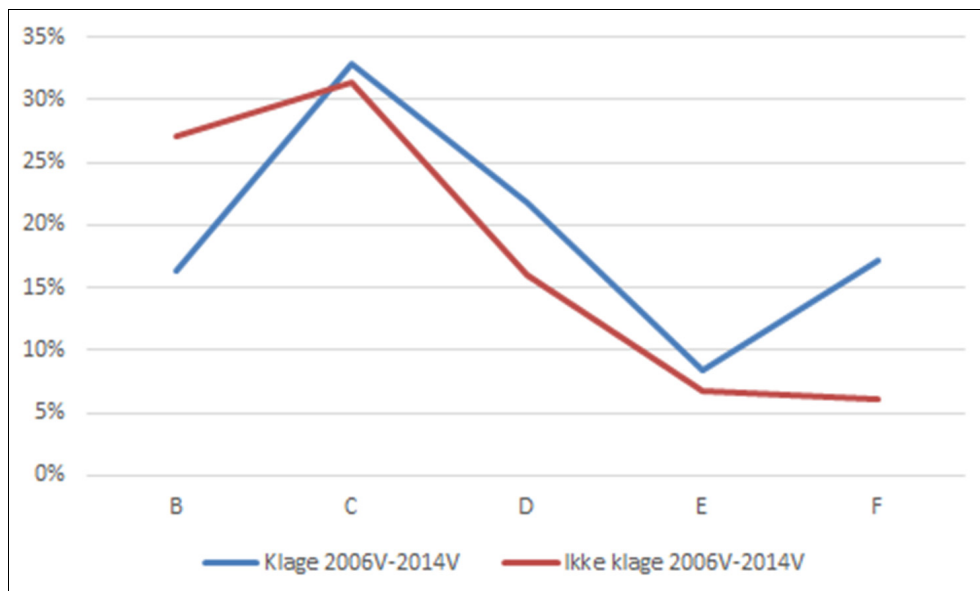
Institusjoner som argumenterte mot «blind» sensur utviser etter forfatterens vurdering større evalueringsfaglig innsikt sammenliknet med dem som støtter endringsforslaget. Universitetet i Oslo støttet ikke forslaget med den begrunnelse at saken bør opplyses så godt som mulig, og at studentene bør komme til orde i prosessen: «Dersom ny sensurering foretas kun på grunnlag av eksamensoppgaven og den forelagte sensorveiledningen, vil dette sidestilles med opprinnelig sensur og gi det samme potensialet for uriktig eller vilkårlig

resultat» (Høringsnotat, UiO, 2013). Studentparlamentet ved UiO var av en annen oppfatning, der spørsmålet om rettferdighet sto sentralt i argumentasjonen. Loven ble likevel endret med virkning fra 1. august 2014. Etter dette fikk ikke lenger klagesensor vite klagers opprinnelige karakter, og ingen begrunnelser for klagen ble akseptert.

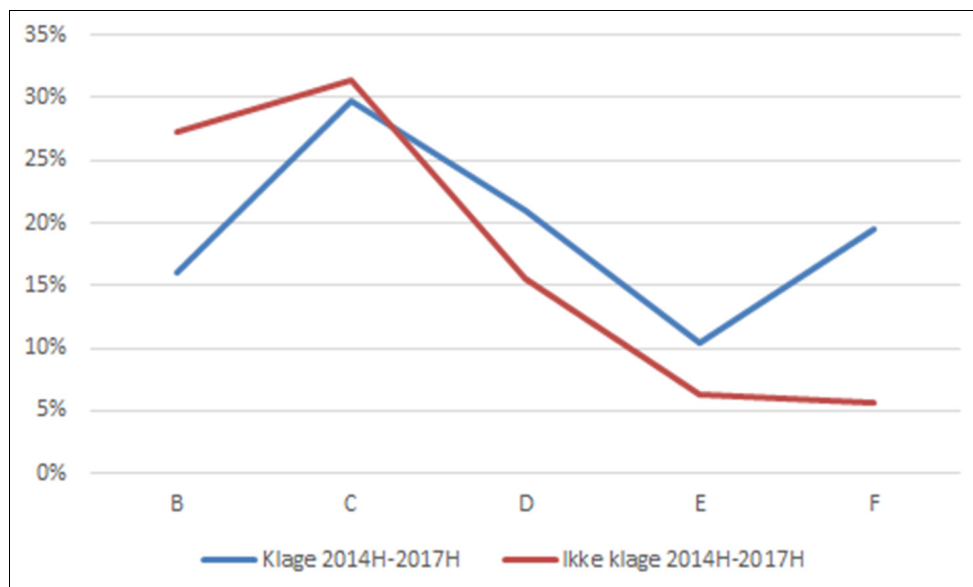
I retningslinjene for begrunnelse og klage ved Universitetet i Oslo blir imidlertid studentene oppfordret til å be om begrunnelse før de eventuelt bestemmer seg for å klage, men dette er ingen betingelse. Ved ønske om begrunnelse får kandidaten informasjon om generelle retningslinjer for sensur samt sensors vurdering. Sensurvedtak ved alle skriftlige eksamener kan påklages, mens muntlig og praktiske prøver er unntatt fra denne regelen (<https://www.uio.no/studier/eksamen/begrunnelse-klage/>).

## Resultater

Vi undersøker først karakterfordelingen blant klager før og etter innføring av «blind» sensur. Grafene i Figur 1 og Figur 2 inkluderer data fra alle fakultetene i de aktuelle tidsperiodene. Her blir det visualisert at kandidater som klagde fikk noe dårligere karakter ved ordi- nær sensur sammenliknet med gruppen som ikke valgte å klage på karakterene.

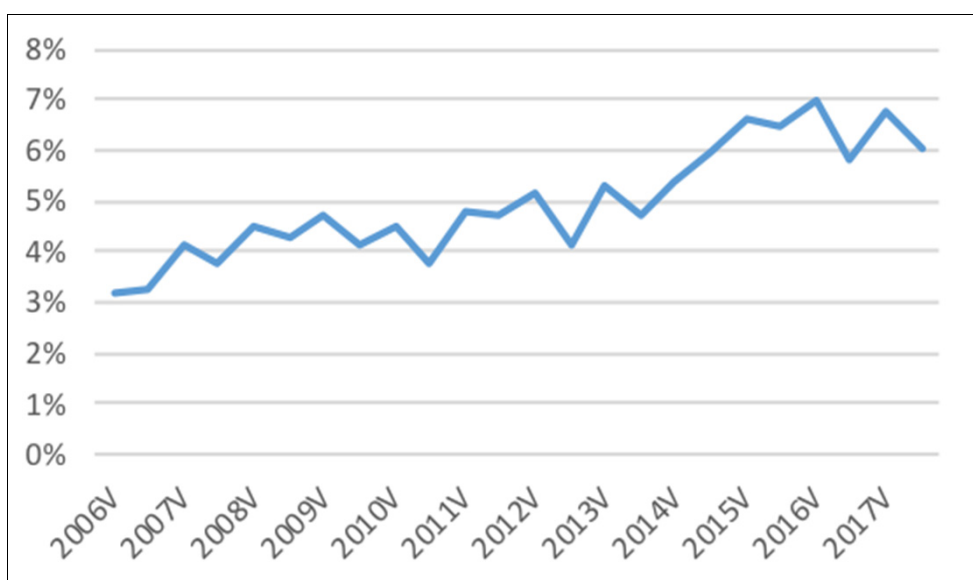


**Figur 1.** Karakterfordeling blant studenter som klager (n=25073) sammenliknet med studenter som ikke klager (n=486930) i perioden før blind sensur ble innført høsten 2014



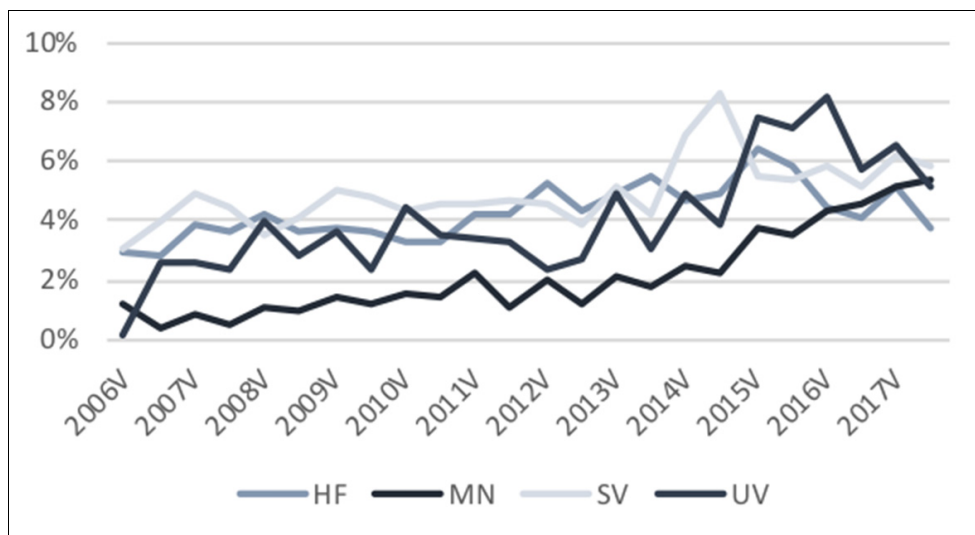
**Figur 2.** Karakterfordeling blant studenter som klager (n=17098) sammenliknet med studenter som ikke klager (n=220830) i perioden etter blind sensur ble innført høsten 2014

Klageomfanget for hele perioden viser en markant økning fra 3 % til 6 % fra våren 2006 til og med høsten 2017 (Figur 3). Dersom en ser perioden under ett, innebærer dette betydelig økt ressursbruk i forbindelse med klagesaker. Statistikken bygger på data fra alle fakultetene ved UiO med unntak av odontologi, som først innrapporterte data fra og med våren 2014.



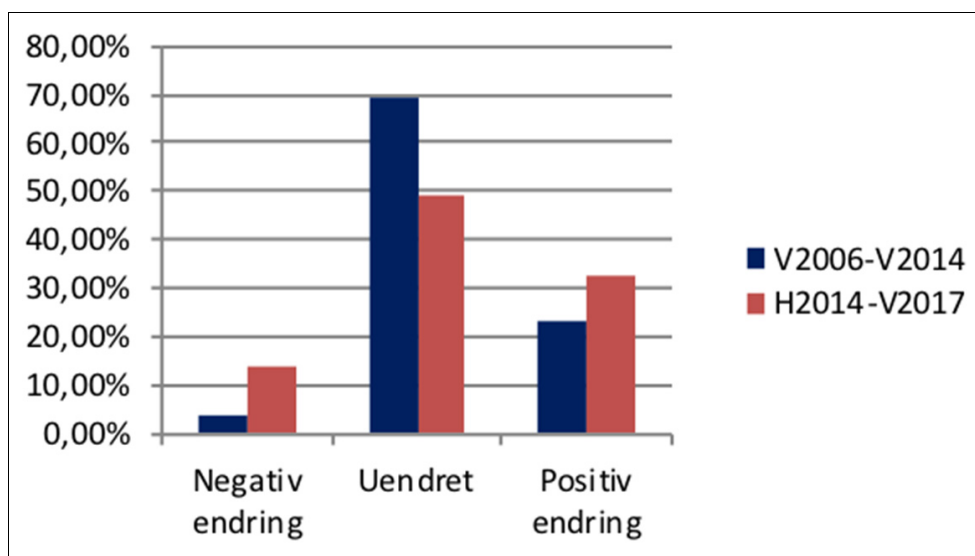
**Figur 3.** Utvikling med hensyn til klageomfang i prosent av alle eksamensbesvarelser fra og med våren 2006 til og med våren 2017. Basert på totalt n=867559 eksamener, hvorav n=43637 ble påklaget

Utviklingen ved HF, MN, SV, UV er framstilt i Figur 4. Det matematisk-naturvitenskapelige fakultet (MN) har laveste klageandelen for nesten hele perioden, men en sterk økning fra 1 % våren 2006 til 5 % høsten 2017. Alle hadde en økning fra 2006 til 2017. Det juridiske fakultet skiller seg ut med et klageomfang fra 9 % våren 2006 til 12 % høsten 2017.



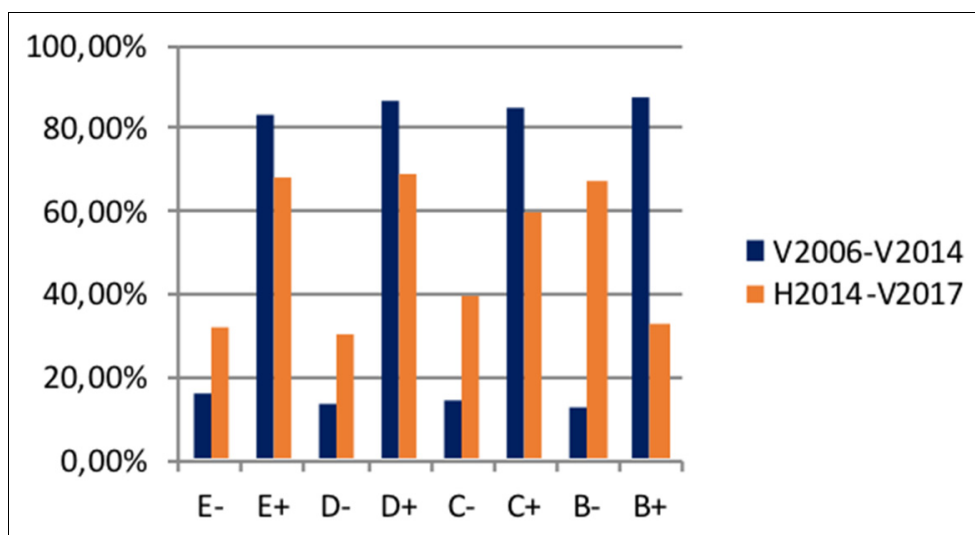
**Figur 4.** Prosentvis endring i klageomfang for fire fakulteter ved Universitetet i Oslo fra våren 2006 til og med høsten 2017. Basert på 699680 eksamener, hvorav 27811 ble påklaget

Figur 5 framstiller retningen (positiv, uendret, negativ) før og etter innføring av «blind» sensur høsten 2014. Langt flere endringer fant sted både i positiv og negative retning etter at «blind» sensur ble innført. Mens nesten 70 % av alle klager i perioden fra 2006 til 2014 endte uendret, er tilsvarende tall i perioden etter «blind» sensur ble innført nede i under 50 %.



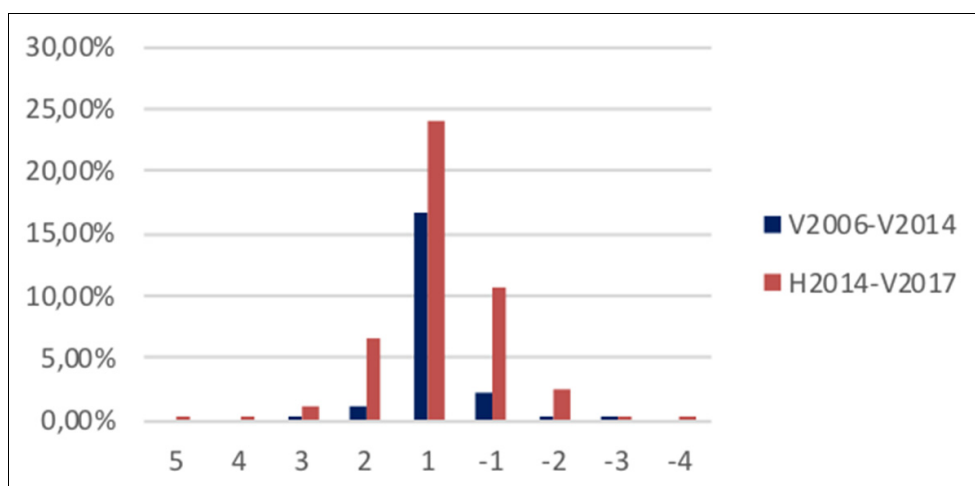
**Figur 5.** Retning på endringer av opprinnelig karakter før og etter «blind» sensur ble innført. Basert på 21399 klager i perioden 2006–2014, og 12536 klager i perioden 2014–2017

Retningen på endringene for de klage som førte til karakterendring er framstilt i Figur 6. Her er bare de klage som faktisk førte til karakterendring tatt med. Symbolet B+ betyr at kandidaten har gått fra karakteren B til A, mens B- betyr at kandidaten har gått fra B til C, D, E eller F. Av Figur 6 framgår det at en langt mindre andel oppnår en bedre karakter etter at «blind» sensur ble innført, mens flere ender opp med en dårligere karakter. Mens cirka 85 % av karakterene som ble endret før «blind» sensur endte i studentenes gunst, ble andelen kraftig redusert etter «blind» sensur. Men selv da endte karakterendringer til klagerens gunst i 60–70 % for alle karaktertrinn med unntak av karakteren B der kun 33 % førte til bedring. Karakteren B framstår derfor som den mest «risikable» å klage på. Figur 6 viser at sannsynligheten for å få en dårligere karakter er stor etter at «blind» sensur ble innført.



**Figur 6.** Detaljerte karakterendringer før og etter «blind» sensur. Symbolene + og – indikerer retning av endring i positive eller negativ retning ut fra opprinnelig karakter. Basert på 3708 klager i perioden 2006–2014 og 4685 klager i perioden 2014–2017

Figur 7 visualiserer størrelsen på endringene før og etter «blind» sensur. «Blind» sensur resulterte i en større andel karakterendringer og større avvik i forhold til opprinnelig karakter.



**Figur 7.** Størrelsen på endringer før og etter «blind» sensur ble innført. Basert på 4381 klager i perioden 2006–2014 og 5736 klager i perioden 2014–2017. Tallene avviker fra Figur 6 ettersom også de som klagde på stryk er tatt med

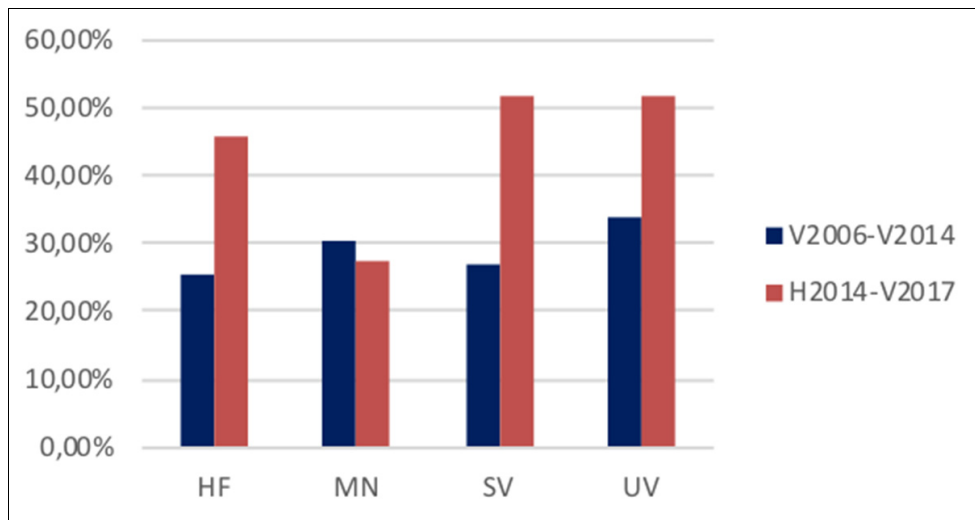
Den prosentvise fordelingen i Figur 7 skjuler datagrunnlaget, som er gjengitt i Tabell 1. Før «blind» sensur ble innført var ingen endringer større enn tre trinn, mens det etterpå var 19 slike endringer, 16 endringer på fire trinn og tre endringer på fem trinn (Tabell 1). I tillegg skal det sies at datagrunnlaget er betydelig mindre etter at «blind» sensur ble innført i 2014.

**Tabell 1.** Størrelse på karakterendringer før og etter «blind» sensur ble innført

|                    | Størrelse på endring |    |     |      |      |     |     |    |   |
|--------------------|----------------------|----|-----|------|------|-----|-----|----|---|
|                    | -4                   | -3 | -2  | -1   | 1    | 2   | 3   | 4  | 5 |
| <b>V2006-V2014</b> | 0                    | 3  | 44  | 489  | 3572 | 253 | 20  | 0  | 0 |
| <b>H2014-V2017</b> | 4                    | 57 | 312 | 1350 | 3028 | 822 | 148 | 12 | 3 |

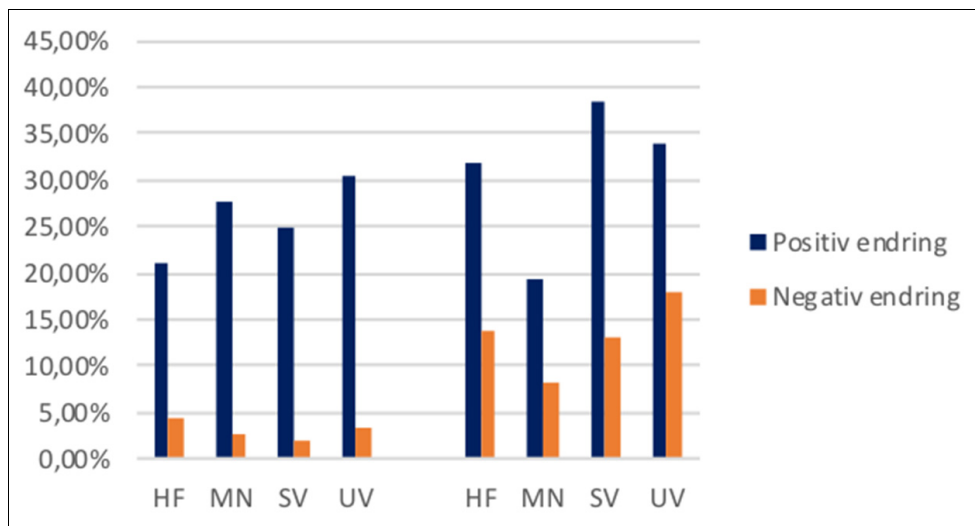


Figur 8 illustrerer fakultetsvise endringer som følge av klage på sensurvedtak før og etter «blind» sensur ble innført. Juridisk fakultet (JF) er utelatt fordi det var mange klager på bestått med ukjent resultat i datasettet. Den relative andelen av karakterendringer økte etter at «blind» sensur ble innført med unntak av Det matematisk-naturvitenskapelige fakultet (MN). Figur 8 illustrerer at «blind» sensur fungerer annerledes for besvarelser med «fasit-svar» enn for besvarelser som har større rom for sensors skjønnsutøvelse (HF, SV og UV).



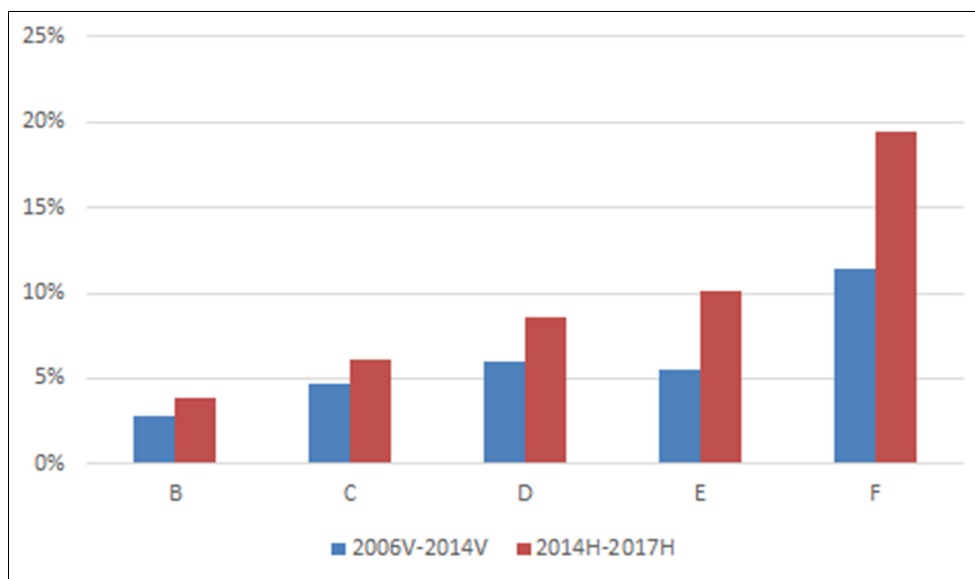
**Figur 8.** Fakultetsvise endringer som følge av klage på sensurvedtak før og etter «blind» sensur ble innført. Basert på 13209 klager i perioden 2006–2014, og 6779 klager i perioden 2014–2017

Figur 9 dokumenterer retning på endring (positiv eller negativ) før og etter at «blind» sensur ble innført. Figuren viser at andelen av både positive og negative endringer har økt med «blind» sensur, med unntak av Det matematisk-naturvitenskapelige fakultet (MN).



**Figur 9.** Retning på karakterendringer før (til venstre) og etter «blind» sensur ble innført (til høyre). Basert på samme antall klager som i Figur 8

Figur 10 viser relativt klageomfang for hvert karaktertrinn før og etter «blind» sensur ble innført. Klageomfanget har økt for alle karaktertrinn etter at «blind» sensur ble innført. Den største økningen skjer ikke uventet på lavere karaktertrinn, spesielt for stryk (F).



**Figur 10.** Prosentvis relativ andel av sensurklager før og etter kravet om «blind» sensur ble innført; n=512003 før «blind» sensur ble innført, mens n=237928 etter at «blind» sensur ble innført

## Analyse

Kandidater som framsetter klage, presterer noe dårligere på opprinnelig sensur sammenliknet med dem som ikke klager både før og etter blind sensur ble innført (Figur 1 og Figur 2). Prosentvis andel av klage på sensurvedtak har samtidig økt for fakultetene enkeltvis og samlet (Figur 3 og Figur 4), slik at alle fakultetene har et prosentvis større klageomfang i 2017 sammenliknet med 2006. Dette har skjedd til tross for omfattende reformer i sektoren. Her kan nevnes kvalifikasjonsrammeverket (2009), bruken av læringsutbyttebeskrivelser, karakterbeskrivelser og sensurveiledninger. Digital informasjonstilgang kan ha bidratt til økt klageutløsel, men felles, nettbaserte skjema ble ikke innført ved UiO før høsten 2016. Den store økningen i antall klager i 2014 (Figur 3) kan derfor mest sannsynlig tilskrives ordningen med «blind» klagesensur. Klageomfanget har økt for alle karaktertrinn fra B, mens andelen av uendrete karakterer også har gått ned, samtidig som endringene noe sjeldnere skjer til kandidatens gunst. Studentene framstår alt i alt som den tapende part karaktermessig sett, dersom en sammenlikner med resultatene av klagesensur før høsten 2014.

I brev til KD datert 17.06.13 argumenterte Norsk Studentorganisasjon (NSO) sterkt for at blind sensur burde bli gjeldende praksis i all høyere utdanning. Dette ble begrunnet ut fra et ønske om likebehandling, og at bestemmelsene i Universitets- og høyskoleloven burde ligge til grunn ved sensurklage. Ifølge NSO ville ny sensur uten tilgang til ytterligere dokumentasjon «sikre en helt nøytral vurdering av besvarelsen og forhindre at det oppstår et inntrykk [...] om at klagevurderingene er motivert av lojalitet sensorene imellom» (NSO, 17.06.13). Ett argument var også at «blind» sensur ville være den mest *rettferdige* vurderingsordningen. All vurdering bygger imidlertid på grader av skjønn, og kan aldri bli «nøytral» i en absolutt betydning. Analysen viser da også at karakterendringer som følge av «blind» klagesensur er minst ved Det matematisk-naturvitenskapelige fakultet, som i vesentlig grad underviser emner med noe mindre grad av skjønn enn øvrige fakulteter. Jo større grad av skjønnsetning i vurderingen, desto større risiko oppstår for «feil» og avvik mellom sensorer ved vurdering av samme besvarelse. Sadler (2009) omtaler dette som integritetsutfordringer ved vurdering, og sorterer feilene i fire kategorier: Tilfeldige og systematiske feil, feil vurderingsprinsipp og feilfokuset vurdering.

*Tilfeldige feil* rammer den enkelte eller noen få, og kan skyldes feil kandidaten selv er ansvarlig for, som for eksempel tidsnød, psykisk press eller misforståelser. Slike feil kan selvsagt også skyldes feil fra sensors side, for eksempel på grunn av uoppmerksomhet. *Systematiske feil* påvirker hele eller større deler av gruppen. Eksempler ved muntlig eksamen kan være sympatier eller antipatier ut fra kjønn, etnisitet, alder med videre. Feil kan også oppstå ved at sensor misforstår beveler som løser oppgaven etter en bestemt metode, eller ved sampling av oppgaver. Underveis i sensuren oppdager for eksempel faglærer at enkelte oppgaver framstår som spesielt lette eller vanskelige med det resultat at råscorer kan bli justert opp eller ned for å kompensere for dette. Dette er informasjon som sensorene ved «blind» sensur går glipp av, og som på denne måten står i fare for å vurdere besvarelsene på et annet grunnlag enn hva tilfellet var ved ordinær sensur.

Spørsmålet om *vurderingsprinsipp* refererer til valget mellom *normbasert* og *kriteriebasert* vurdering. Selv om vi her til lands er pålagt å benytte kriteriebasert vurdering, viser både erfaring og forskning at dette er et krevende prinsipp å etterleve (Sadler, 2005). Derfor blir normbasert vurdering, gjerne med henvisning til normalfordeling, noen ganger benyttet som «korrektiv», selv om en slik praksis strir mot de offisielle retningslinjene. *Feilfokuseret vurdering* innebærer at aspekter ved besvarelsen eller kandidatens atferd blir trukket inn som del av vurderingsgrunnlaget, for eksempel innsats, progresjon og motivasjon. Integritetsutfordringene gjelder for all sensur, men sjansene for å oppdage og korrigere feil blir kraftig redusert når begrunnelse(n) for klagen ikke er tilgjengelig.

## Diskusjon

Kravet om to sensorer ved alle ordinære eksamener med unntak av masteroppgaver, derav én ekstern, falt bort i 2003 med innføringen av kvalitetsreformen. Endringen står imidlertid ikke i veien for bruk av to sensorer. Etter årtusenskiftet har likevel større vekt blitt lagt på bruken av karakterbeskrivelser og sensurveiledninger, noe som setter krav til vurderingskompetanse, og til hvordan denne blir etablert og praktisert:

... achievement standards are, in essence, abstract concepts ... [and] cannot be reduced to set procedures that can be applied by non-experts. ... They cannot be written down as detailed verbal descriptions, categories or lists which can then be used by students and assessors alike. (Sadler, 2009, s. 820)

Verbale beskrivelser og forklaringer i seg selv er utilstrekkelige. Ordene trenger en konkret referanse, fordi «...verbal descriptions do not have unique meanings or fixed, context-free interpretations» (Sadler, 2009, s. 820–821). «Codifications are inherently incapable of meeting the requirements because key terms lack the necessary attributes» (Sadler, 2014, s. 273).

... the starting point is an exemplar of a standard, not a verbal description of a standard. In determining the quality of complex student work, recognition – not definition – is the primary act. The purpose of a description is to account for a particular judgment. ... The words form the necessary link between an immediate concrete referent and the abstract concept of quality. (Sadler, 2009, s. 821)

En spesiell utfordring består i at læringsutbyttebeskrivelsene kun angir «terskelverdien eller minstekravet for å bestå studieprogrammet» (Kulturdepartementet, 2017, s. 47). Departementet nøyer seg med å slå fast at det er «viktig å formidle at kravene for beste karakter er kvalitativt annerledes enn kravene til bestått» (Kulturdepartementet, 2017, s. 47). Hvilke standarder som gjelder for de øvrige karakterene, blir dermed overlatt til klagesensorenes skjønn, og resultatene er dokumentert i denne undersøkelsen. Fra og med høsten 2014 til

og med våren 2017 fikk 148 kandidater forbedret karakteren sin med tre karaktertrinn, mens 57 fikk redusert karakteren med tre trinn. Tilsvarende tall for perioden fra våren 2006 til og med våren 2014 var 20 og 3. Slike tall reiser en rekke spørsmål.

Denne undersøkelsen gir ikke svar på hvilke vurderingsprinsipper som har vært benyttet, eller om ulike prinsipper har vært benyttet ved ordinær sensur og klagesensur. Ettersom sammenlikningsgrunnlaget oftest er større ved ordinær sensur, kan en spørre om relativ vurdering da har vært oftere benyttet enn ved klagesensur. En slik praksis er imidlertid ikke i tråd med kravet om kriteriebasert vurdering, men er likevel mye benyttet fordi kriterier og standarder kan være krevende å operasjonalisere. Om klagesensorene benyttet «kriteriebasert» vurdering, er uvisst, og KD har ikke gitt føringer for praksis. Pålegg om sensorveiledninger, generelle og spesielle karakterbeskrivelser er velmente tiltak, men er i seg selv utilstrekkelig som bidrag til konsistent vurdering: «... our current reliance on criteria, rubrics, moderation and standardising grade distributions is unlikely to tackle the proven lack of grading consensus» (Bloxham, den-Outer, Hudson, & Price, 2016, s. 479).

I brev til KD med ønske om «blind» sensur viste NSO til et brev fra departementet der det ble argumentert for at ny sensur skulle gjennomføres mest mulig lik den første sensuren (NSO, 17.06.13). Dette ble tolket som at «de nye sensorene verken skulle ha tilgang til de første sensorenes karakterfastsettelse og begrunnelse eller studentenes begrunnelse for klagen» (NSO, 17.06.13). KD åpnet deretter for valgfri sensurordning, noe NSO tolket som forskjellsbehandling. Under den påfølgende høringsrunden argumenterte flere institusjoner for at «blind» sensur ville hindre uønsket lojalitet mellom sensorene: «Det er besvarelsen som skal vurderes på nytt, og ikke arbeidet til de første sensorene» (Prop. 59 L (2013-2014) s. 43).

Basert på to års erfaring argumentere de tre juridiske fakultetene for at nyordningen likevel ikke hadde hatt ønsket effekt, noe som ble begrunnet i brev til KD av 29.06.16:

- «Dagens ordning er ikke en reell klageordning etter forvaltningslovens kapittel VI, men en mulighet for studentene til å få en alternativ vurdering. Det er ingenting som tilsier at den nye sensuren er av bedre kvalitet enn den ordinære.
- Klagesensuren har et mindre sammenlikningsgrunnlag. Siden det er studenter med resultater i det lavere sjikt som klager, vil klagesensorene få et skjevt bilde av nivået for det respektive emnet.»

Klagesensur er beheftet med samme utfordringer som ordinær sensur, men kontekstene er ulike. Det samme er sensorenes kunnskaps- og erfaringsbakgrunn. Utfordringene forsterkes ved at klagerne ikke skiller mellom formelle feil og misnøye med utfallet av sensuren. «Blind» sensur har derfor bidratt til uklarhet om kandidatens motivasjon for klagen, og dermed redusert klagesensorenes muligheter til å vurdere dens relevans. En klage på sensurvedtak betyr nemlig ikke nødvendigvis misnøye med det gitte vurderingsuttrykket, slik Universitetet i Oslo skriver til Kunnskapsdepartementet (29.10.13):

Ofte avdekkes det ... at det snarere er snakk om klage på formelle feil enn klage på karakter, for eksempel når studenten klager på misforhold mellom undervisning og eksamensoppgaver. Fordi vi antar at en ordning med blind omsensur innebærer at studentene heller ikke skal levere begrunnelse, mister vi denne muligheten til å få klagen behandlet riktig (<https://www.regjeringen.no/>)

Norsk Studentorganisasjon begrunnet ønsket om «blind» sensur ved at dette er det mest *rettferdige* overfor studentene. I denne studien har vi sett at kandidater som klager i gjen-

nomsnitt presterer svakere sammenliknet med dem som ikke klager (Figur 1 & Figur 2). Litteraturen på dette feltet opererer med ulike aspekter ved rettferdighet; ett som handler om selve vurderingsuttrykket og et annet som relaterer til prosessen: «Distributive justice refers to the extent to which the outcomes received in an allocation decision are perceived as fair [...] whereas procedural justice is concerned with the fairness of procedures used to reach those decisions» (Tata, 1999, s. 265).

Resultatlikhet, prosedyrelikhet og sjanselikhhet kan oppfattes som ulike aspekter ved rettferdighet. Resultatlikhet kan imidlertid ikke forventes i grupper med ulik faglig bakgrunn og motivasjon. De to øvrige aspektene er ønskverdige, men ikke lett å imøtekomme ved klagesensur. Et viktig forutsetning er at vurderingsgrunnlaget skal være mest mulig likt første sensur, noe som i praksis ikke er mulig. I vårt tilfelle var sensurdata fra ordinær sensur ikke tilgjengelig, klagesensorenes bakgrunn og erfaring var ulik, og de delte ikke nødvendigvis det første sensorparets tolkningsrammer for vurderingen. En britisk undersøkelse konkluderte med en rekke begrunnelser for hvorfor praksis avviker mellom ulike sensorer. Blant disse er bruken av ikke offisielle kriterier; kriterier kan ha ulike underkriterier med individuelle tolkningsrom. Sensorene kan ha ulike oppfatninger om krav til faglig nivå og benytte ulik vektning av både gitte og implisitte kriterier (Bloxham, den-Outer, Hudson & Price, 2016).

NSO knyttet spørsmålet om prosedyreliket til rettferdighet, og studentorganisasjonens ønsker dannet, i hvert fall delvis, beslutningsgrunnlag for innføring av «blind» sensur. Studentenes stemme er viktig i dette spørsmålet, men bedre evalueringsteoretiske begrunnelser ville styrket beslutningsgrunnlaget for valg av klagesensurordning. I skrivende stund har regjeringen vedtatt en tredje sensur der det er to eller flere trinns endring fra ordinær sensur til klagesensur. Dermed er prinsippet om «blind» sensur igjen forlatt ved at man vender tilbake til en form for åpen sensur, slik ordningen var tidligere.

## Oppsummering og konklusjon

Kandidater som klager på sensurvedtak, presterer noe svakere sammenliknet med sine medstudenter som ikke klager, og denne tilbøyeligheten er upåvirket av sensurordningen. Klage tilbøyeligheten har økt gjennom hele perioden fra 2006 til 2017, altså ikke vesentlig påvirket av sensurordningen. Etter at «blind» sensur ble innført, har relativt færre kandidater beholdt karakteren fra ordinær sensur, mens relativt flere har blitt endret både i positiv og negativ retning. Omfanget av endringer er størst ved SV og HF og UV, og minst ved MN. «Blind» sensur har ført til større spredning og større avvik i forhold til opprinnelig karakter, men også til jevnere fordeling mellom endringer i positiv og negativ retning.

«Blind» klagesensur ble innført blant annet for å hindre lojalitetsbaserte beslutninger ved klagesensur. Klage på sensurvedtak fulgte etter dette ikke klageprosedyrer kjent fra forvaltningsretten, men ble gjennomført som en helt ny sensur. I verste fall bidrar denne ordningen til at en «helt riktig» vurdering ved ordinær sensur ender opp som en som er «helt feil» ved ny sensur. Blant de kjente feilkildene ved vurdering er klagesensorene om mulig i en enda mer utsatt posisjon enn sensor(ene) ved ordinær sensur. Her kan spesielt nevnes systematiske feil på grunn av sampling av oppgaver, som nye sensorer ikke har mulighet til å kjenne til eller kompensere for. Det framstår også som en urimelig oppgave å vurdere en klage uten begrunnelse. Ingen lærer noe av dette, samtidig som opplevelsen av vilkårlighet og mistillit til vurderingsarbeidet forsterkes ved flere og større sprik i vurderingene.

I tilbakeblikk er det lett å dele ønsket om likebehandling og rettferdighet, men noen ganger kan ønsket om det beste ende som det godes fiende. «Blind» klagesensur kan være et

eksempel, der nyordningen imøtekom studentenes ønsker, men der begrunnelsene viste seg å være utilstrekkelige. Dette reiser i sin tur spørsmål ved forholdet mellom myndighetsutøvelse og forskningsbasert innsikt i forbindelse med reformarbeid i høyere utdanning.

Sensur forutsetter utøvelse av faglig skjønn, og det er lang tradisjon for at dette blir organisert som en samtale mellom bedømmere der argumenter blir prøvd i forhold til hverandre. «Blind» klagesensur kan derfor oppfattes som grunnleggende uakademisk med mistro til klagesensorenes faglige og personlige integritet. All faglig kritikk er basert på *begrunnelser*, og det er gjennom denne prosessen faglige standarder blir etablert og bekreftet. Klagesensur uten kjennskap til grunnlaget for klagen og til begrunnelsen for karakteren strir mot akademiske idealer og løsning av uenigheter. Kandidatens begrunnelse sammen med en beskrivelse av kriteriene ved ordinær sensur kunne vært til hjelp, men det beste ville trolig være et felles møte mellom alle parter med vekt på begrunnelser og læring.

## Erkjentlighet

Takk til student Yngvild Hamre ved NTNU for bistand i forbindelse med de statistiske analysene i artikkelen.

## Litteratur

- Aalde, G. (2001). *Saksdokumenter ved ny sensur som følge av klage over karakterfastsetting*. Oslo.
- Bloxham, S., den-Outer, B., Hudson, J., & Price, M. (2016). Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, 41(3), 466–481. DOI: <https://doi.org/10.1080/02602938.2015.1024607>.
- Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London: Kogan Page.
- Høringssvar fra Det juridiske fakultet ved Universitetet i Bergen til forslag om endringer i universitets- og høyskoleloven (16.09.13). Hentet fra: [https://www.regjeringen.no/contentassets/6d2ca9cd2d174b7ab056253ae9f0f085/uiib\\_jur\\_fak.pdf?uid=Universitetet\\_i\\_Bergen,\\_Det\\_juridiske\\_fakultet\\_](https://www.regjeringen.no/contentassets/6d2ca9cd2d174b7ab056253ae9f0f085/uiib_jur_fak.pdf?uid=Universitetet_i_Bergen,_Det_juridiske_fakultet_).
- Høringsnotat – forslag til endringer i universitets- og høyskoleloven. (2013). Oslo. Hentet fra: <https://www.regjeringen.no/contentassets/a304f8804ea54b4a9d6230a143896b09/hoeringsnotat.pdf>.
- Kjellemo, B.T. (1999). *Saksdokumenter ved ny sensur som følge av klage over karakterfastsetting*. Oslo.
- Kunnskapsdepartementet (1988). *Lov om universiteter og vitenskapelige høyskoler (universitetsloven)*. (Ot.prp. nr. 62, 1988–89).
- Kunnskapsdepartementet (2013). *Endringer i universitets- og høyskoleloven. (Prop. 59 L, 2013–2014)*. Hentet fra: [https://www.regjeringen.no/no/dokumenter/Prop-59-L-20132014/id754603/?q=Prop%2059%20L%20\(2013-2014\)](https://www.regjeringen.no/no/dokumenter/Prop-59-L-20132014/id754603/?q=Prop%2059%20L%20(2013-2014)).
- Kunnskapsdepartementet (2017). *Kultur for kvalitet i høyere utdanning. (Meld. St. 16, 2016–2017)*. Oslo. Hentet fra: <http://www.regjeringen.no>
- Mortensen, T. (2015). Studenter taper på blind sensur. *Universitas*. Hentet 14.04. fra: <http://universitas.no/nyheter/60460/studenter-taper-pa-blind-sensur>.
- Norsk studentorganisasjon (NSO) (2013). *Brev til Kunnskapsdepartementet om klagesensur*. Hentet 17.06. fra: <https://student.no/content/uploads/2014/07/Innspill-til-Kunnskapsdepartementet-om-klagesensur.pdf>
- Oftedal, T. (2013). *Brev til kunnskapsdepartementet om klagesensur*. Hentet fra: <https://www.student.no/content/uploads/2014/07/Innspill-til-Kunnskapsdepartementet-om-klagesensur.pdf>.
- Rust, C. (2002). The impact of assessment on student learning. *Active Learning in Higher Education*, 3(3), 145–158.
- Sadler, D.R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13(2), 191–209.

- Sadler, D.R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175–194.
- Sadler, D.R. (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 34(7), 807–826.
- Sadler, D.R. (2010). Assessment in higher education. I P. Peterson, E. Baker & B. McGaw (Red.), *The international encyclopedia of education* (Vol. 3, s. 249–255). Oxford: Elsevier.
- Sadler, D.R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, 67(3), 273–288. DOI: <https://doi.org/10.1007/s10734-013-9649-1>.
- Snyder, B.R. (1971). *The hidden curriculum*. New York: Knopf.
- Tata, J. (1999). Grade Distributions, Grading Procedures, and Students' Evaluations of Instructors: A Justice Perspective. *The Journal of Psychology*, 3, 261–273.
- Universitetsloven (1990). *Lov av 16. juni 1989 nr. 77 om universiteter og vitenskapelige høyskoler (LOV-1989-06-16-77)*. Oslo: Grøndahl.
- Universitetsloven (1997). *Lov av 12. mai 1995 nr. 22 om universiteter og høyskoler*. Oslo: Cappelen Damm Akademisk.
- Universitets- og høyskoleloven (2005). *Lov om universiteter og høyskoler. (LOV-2005-04-01-15)*. Hentet fra <https://lovdata.no/dokument/NL/lov/2005-04-01-15>.