

Et informatikkperspektiv på norsk lovgivning om karakterklager i høyere utdanning

Guttorm Sindre
IDI, NTNU

Abstract

KD har i en nylig proposisjon foreslått endringer av klagesensurordningen, hvor det legges opp til at en modell med blind omsensur, som allerede har vært brukt ved en del læresteder, nå skal gjøres obligatorisk for alle læresteder. I denne artikkelen sammenlignes gammel og ny klagesensurordning med ordningene i en del andre land. Vi diskuterer ulike typer sensurfeil (med eksempel fra programmering) og kvalitetskriterier for klagebehandling, hvorpå norske og en generisk utenlandsk ordning sammenlignes etter disse kriteriene. Sammenligningen er ikke empirisk, kun analytisk basert på "sunn fornuft". Den indikerer at den norske ordningen for klagesensur ikke holder mål mot internasjonal "state-of-the-art", og at den er unødvendig kostbar uten at dette gir seg utslag i høyere kvalitet av sensurresultatene eller muligheter for læring eller prosessforbedring.

1. Introduksjon

Behandling av klager på karakterer krever betydelige ressurser ved mange norske læresteder uten å gi noen økt læring for studentene [1]. Samtidig er klagesensur også en viktig sikkerhetsventil hvis det gis urimelige karakterer – dog kun når det skjer i studentens disfavør, siden ufortjent gode karakterer neppe blir påklaget. Som vist i [2-3] kan sensorreliabiliteten i enkelte emner være lav, så det er hevet over tvil at urettferdige karakterer av og til vil forekomme. Påliteligheten av sensur vil variere mellom ulike fagfelt og evalueringsformer, men IT-utdanninger spenner bredt, fra matematikknære emner hvor hvert delspørsmål har en klar fasit, til emner nært opp til samfunnsfag, med essayoppgaver som kan besvares på mange ulike måter, og store gruppearbeider [4-5] hvor hver gruppe har sitt unike prosjekt. Sist men ikke minst er programmering, som inngår i mange kurs i løpet av et informatikkstudium, utfordrende å rette. Det krever samme presisjon som matematikk, selv ett enkelt tegn (f.eks. $>$ i stedet for \geq) kan være nok til at et program ikke virker som det skal. Men i motsetning til matematikk har ikke oppgavene et fast svar med to streker under – hele besvarelsen må leses like nøye, og løsningsrommet er vesentlig større enn på gjengse matematikkoppgaver. Uansett oppgaveform gjør knappe sensurfrister at store bunker med eksamensbesvarelser må rettes på kort tid, og da er det ikke overraskende at det kan forekomme feil.

Samtidig synes det også hevet over tvil at en betydelig andel av klagen på karakter ved norske læresteder er uten substans. En undersøkelse av Gabrielsen et al. [6] indikerer at rundt 70% ender med samme karakter som opprinnelig gitt. For noen klager har nok studenten konkrete argumenter om feilbedømmelse, etter å ha sett på besvarelse, løsningsforslag og begrunnelse. I andre tilfeller er det bare en generell misnøye med karakteren, uten å kunne vise til spesifikke feil i sensuren. I emner hvor karakter baseres på arbeider underveis pluss avsluttende eksamen, scorer de fleste studenter bra på arbeidene og dårligere på eksamen [7]. Sluttkarakteren vil da komme som en skuffelse etter at arbeidene hadde skapt forhåpninger om noe bedre. En annen plausibel grunn til skuffelse er «over gjennomsnittet»-effekten. Denne gjør seg gjeldende på mange arenaer, f.eks. at 70% av sjåfører tror de er bedre enn snittet. Studier både internasjonalt [8] og ved UiO [9] viser at effekten også er relevant for studenters selvoppfatning. Hvis de fleste mener de er bedre enn snittet, må nødvendigvis mange bli misnøyd når et flertall over tid skal ha C eller dårligere. Sist men ikke minst kan en del karakterklager være rent opportunistiske: Klageren behøver ikke tro at det er

Denne artikkelen ble presentert på NIK 2014. For mer informasjon se <http://www.nik.no/>

noe feil med sensuren, bare at sjansen for bedre karakter ved omsensur er høyere enn risikoen for dårligere karakter. Hvis begrunnelsen f.eks. sier at det var en «sterk C», eller at studenten scoret 75% og dermed C (65-76 etter NTNUs anbefalte skala p.t.)¹, vil de fleste tenke at det skal lite til for å gå opp, men mye til for å gå ned. Mange faglærere har inntrykk av at "nær høyere karakter" er en utbredt motivasjon for å klage. I en undersøkelse av Gynnild [10] fremkom det også at NTNU-studenter som klaget uten å be om begrunnelse først, hovedsakelig hadde fått F (ingen risiko for gå ned), mens de med ståkarakter – hvor en klage ville ha en potensiell risiko – i de aller fleste tilfeller innhentet begrunnelse først.

De siste årene har det vært to ulike ordninger for behandling av karakterklager ved norske læresteder. Den ene ordningen kalles i denne artikkelen «**blind omsensur**». Med denne ordningen får de nye sensorene kun tilgang til samme informasjon som de opprinnelige sensorene fikk (besvarelse, eksamensoppgave, løsning / sensorveiledning). Den andre ordningen, «**informert omsensur**», tilsier at de nye sensorene også får vite opprinnelig karakter og eventuelle begrunnelser fra sensor og student. I følge Gabrielsen et al. [6] var blind omsensur i 2013 i bruk ved 5 universiteter (NTNU, UiA, UiS, UiT, UMB) og 10 høyskoler (bl.a. HiOA, HiG, HiØ), mens informert omsensur var i bruk ved 3 universiteter (UiO, UiB, UiN) og 15 høyskoler (bl.a. NHH, HiST, HiB). For øvrig er de to modellene like. Grad av anonymisering av studenten (som ellers kunne være det man mente med «blind») vil være den samme, likeledes at omsensur gjøres av to nye sensorer, hvorav minst en ekstern, og at ny sensur skal foreligge innen 3 uker fra klagedato, som er samme sensurtid som gjelder for ordinær eksamen.

Gabrielsen et al. [6] sammenligner også de to ordningene når det gjelder utfall av omsensuren, basert på (ikke helt komplette) tall fra 30 norske læresteder for semestrene høst 2008 til og med vår 2010. Noen av resultatene er her syntetisert i Tabell 1. Som man kan se fører blind omsensur til at flere klager får forbedret sin karakter, men også til at vesentlig flere (en tredobling) får dårligere karakter. Noen klager gir også flere karaktertrinn endring, netto karakterforbedring per klage er i begge tilfeller ca. 0,17 karaktertrinn. Dette betyr at en student i snitt må levere 6 klager i løpet av et studium for å forbedre karakteren sin med ett trinn i ett emne. Med rundt 20 emnekarakterer i et bachelorstudium, vil netto forbedring av en bachelorkandidats karaktersnitt per klage være omtrent 0.01 karaktertrinn. Med unntak av situasjoner hvor snittet havner helt på vippen mellom to karakterer må man således ha levert et betydelig antall klager for at dette skal kunne påvirke hovedkarakteren. Dette viser at studenter sjelden oppnår mye med å klage, og at klagebehandling bruker forholdsvis store ressurser på noe som ikke gir særlig betydningsfulle resultater.

Tabell 1: Resultater ved omsensur (syntetisert og omregnet fra [6])

Ordning	Gunst	Ugunst	Uendret	+ per klage
Informert	19,3%	3,1%	77,4%	0.17
Blind	23,8%	9,8%	66,3%	0.17

I sin Prop 59L (2013-2014)² foreslår KD å endre Universitetsloven slik at blind omsensur skal være det eneste tillatte: «*Departementet mener [...] at blind sensur er mer i overensstemmelse med lovgivers intensjon om at ny sensur skal være så lik en vanlig førstegangssensur som mulig. Det er denne modellen som gir best sikkerhet for*

¹ <https://innsida.ntnu.no/wiki/-/wiki/Norsk/Prosentvurderingsmetoden>

² <http://www.regjeringen.no/nb/dep/kd/dok/regpubl/prop/2013-2014/Prop-59-L-20132014/5.html?id=754661>

en upartisk vurdering av eksamen, og det er denne modellen som får støtte fra de fleste av høringsinstansene.» (seksjon 5.3.4 i proposisjonen som det er vist til i fotnoten).

I denne artikkelen stiller vi følgende forskningsspørsmål: (1) Hvordan er klagesensurordningene i andre land som det er naturlig å sammenligne oss med? (2) Hvilken ordning (blind omsensur, informert omsensur, eller noe annet – f.eks. inspirert fra utlandet) gir størst sannsynlighet for å rette opp eventuelle feil i opprinnelig sensur? (3) Hvilken ordning (blind / informert / annet) gir best læring for studentene, og for fagmiljøene. (4) Hvilken ordning (blind / informert / annet) er den mest kostnadskrevende? Med unntak av spørsmål 1 vil spørsmålene bli undersøkt med en analytisk tilnærming preget av «sunn» fornuft, av mangel på empiriske data.

Resten av artikkelen er strukturert som følger: Seksjon 2 sammenligner den norske klageordningen med tilsvarende ordninger fra andre land. Seksjon 3 trekker opp noen kvalitetskriterier for klagesensur. Seksjon 4 sammenligner klageordninger etter kriteriene fra seksjon 3. Seksjon 5 gir en avsluttende diskusjon og konklusjon.

2. Sammenligning med andre land

En inngående analyse av klageordninger i et bredt utvalg av land er hinsides denne artikkelens målsetning, men Tabell 2 oppsummerer funnene fra et drøyt dagsverk med søk på nettet. For alle landene gjelder at klage må være begrunnet, så dette er ikke nevnt eksplisitt for hver rad. I kolonnen "Kilde" i tabellen er det angitt hvorvidt ordningen er funnet beskrevet i landets lovgivning, eller bare på universitetsnettsider. I det siste tilfellet er ordningen neppe representativ for alle læresteder i landet, men er i det minste et eksempel på en klageordning som er akseptabel i henhold til landets lovgivning.

Tabell 2: Klageordninger i andre land

Land	Ordning	Kilde
Sverige	Student kan be faglærer revurdere karakteren, den kan ikke overprøves av andre. Eventuell endring kun i favør av studenten. Uppsala har en ordning med karakterombudsmann som kan anbefale tiltak men ikke selv overprøve karakteren som faglærer har satt.	L ³ , W ⁴
Danmark	Klage på karakter går først til opprinnelige sensorer for begrunnelse (og evt. løsning hvis sensorer innser feil, eller studenten av begrunnelsen ser at karakter var ok). Student kan deretter klage videre til nemnd. Nemnda får alle dokumenter og kan avvise klagen hvis den ikke har tilstrekkelig substans, eller ved medhold tilby ny prøve eller ny sensur.	L ⁵
England	Kan i utgangspunktet bare påklage formelle feil ved eksamen, ikke sensors akademiske skjønn.	L ⁶
Tyskland	Karaktersetting kan påklages skriftlig til faglærers overordnede, som diskuterer med faglærer. Faglig skjønn overprøves vanligvis ikke, annet enn i eksepsjonelle tilfeller.	W ⁷
USA	Varies fra universitet til universitet. Typisk må student diskutere med	W ⁸

³ "Finner en examinator att ett beslut om betyg är uppenbart oriktigt på grund av nya omständigheter eller av någon annan anledning, skall han eller hon ändra beslutet, om det kan ske snabbt och enkelt och om det inte innebär att betyget sänks." Höskoleförordning (1993:100), §24

⁴ <http://www.uu.se/student/rattigheter/betygsombudsman/>

⁵ <https://www.retsinformation.dk/Forms/R0710.aspx?id=163987>: *Bekendtgørelse om eksamen og censur ved universitetsuddannelser (eksamensbekendtgørelsen)*, §§34-37.

⁶ Kan bare klage på formelle feil ved eksamen, ikke på sensors skjønn: "A complaint which falls within subsection (1) is not a qualifying complaint to the extent that it relates to matters of academic judgment" (Higher Education Act 2004, Part II, §12 (2)) <http://www.legislation.gov.uk/ukpga/2004/8/part/2>:

⁷ http://www.fu-berlin.de/en/sites/fubest/doku/Academic_Regulations_and_Code_of_Conduct.pdf

	eksaminator / faglærer, deretter evt. klage, noen steder til nemnd, noen steder til overordnet (instituttleder, dekan). Uenighet om faglig skjønn normalt ikke tilstrekkelig klagegrunn, må dreie seg om dokumentert urettferdighet (f.eks i forhold til andre studenter), formelle feil eller diskriminering.	
Irland	Klage på karakter går til nemnd, må angi presist hvilke deler man mener er feilbedømt. Det sies eksplisitt at det at man var poengmessig nær en bedre karakter, <u>ikke</u> er en akseptabel klagegrunn som sådan.	W ⁹
NL	Prøver først å løse saken uformelt ved kontakt med faglærer, kan deretter klage videre til nemnd med detaljert skriftlig begrunnelse for klagen. Nemnd kan overprøve karakter.	W ¹⁰
Canada	Først uformell kontakt med faglærer, så eventuelt klage. Mange universiteter krever et klagegebyr (typisk 20-50 CAD) som blir refundert hvis studenten får medhold av ny sensor, som har alle dokumenter tilgjengelig.	W ¹¹
Australia	Først uformell kontakt med faglærer, så eventuelt klage. Klagegebyr vanlig (typisk 20-50 AUD) som blir refundert hvis studenten får medhold av ny sensor, som har alle dokumenter tilgjengelig.	W ¹²
Sveits	Først kontakt med faglærer, denne kan selv be om endring av karakter hvis feil oppdages. Hvis fortsatt uenighet kan student sende en klage til nemnd. Presis liste over hva som er mulige klagegrunner. Saksbehandlingsgebyr som refunderes hvis medhold, studenter uten midler kan søke om dekning av omkostninger.	W ¹³

Selv om den norske debatten om klagesensur kan synes å dreie seg om preferanser for blind eller informert omsensur, er det vel så interessant å observere fra tabellen at begge disse norske ordningene er uvanlige i et internasjonalt perspektiv. Norsk høyere utdanning avviker her fra vanlig internasjonal praksis på en rekke punkter:

- 1. I Norge har studenten lov til å klage uten først å være i dialog med faglærer / opprinnelig sensor.** I de fleste andre land må studenten først få forklart hvorfor en karakter ble gitt, og eventuelt søke en minnelig løsning hvis denne forklaringen ikke synes tilfredsstillende. En klage kan bare reises hvis studenten etter denne dialogen fortsatt finner grunn til å klage. I Norge *kan* studenten be om begrunnelse først, men er ikke pålagt å gjøre det.
- 2. I Norge kreves det ikke at studenten begrunner klagen.** Forfatteren har ikke klart å finne noe annet land hvor dette er tilfelle. I alle landene i Tabell 2 må en klage eller anmodning om revurdering av karakter begrunnes.
- 3. I Norge er alle klager på karakter garantert å få omsensur av to nye sensorer.** Forfatteren har ikke klart å finne noe annet land hvor dette er tilfelle. Andre steder vil klager bli avvist uten realitetsbehandling hvis de enten mangler begrunnelse eller har en begrunnelse som ikke er gangbar (f.eks. at man var nær en høyere karakter). De fleste steder vil klager også kunne avvises uten omsensur på tross av en begrunnelse, hvis klagenemnd finner at begrunnelsen ikke har tilstrekkelig substans.
- 4. I Norge praktiserer mange læresteder blind omsensur, og dette vil snart bli gjeldende standard for alle lærestedene.** I andre land hvor det fins muligheter

⁸ f.eks., <http://catalog.arizona.edu/2011-12/policies/gradappeal.htm>, http://web.mit.edu/registrar/reg/grades/grade_change.html, http://registrar.berkeley.edu/grade_appeal.html

⁹ http://www.ucd.ie/appeals/exam_appeal.htm

¹⁰ <http://gsss.uva.nl/about/rules-regulations/appeals-and-complaints.html>

¹¹ http://umanitoba.ca/student/resource/student_advocacy/final_grade_appeal.html,

¹² <http://www.newcastle.edu.au/current-students/learning/assessments-and-exams/appealing-a-final-result>

¹³ http://myunisg.ch/images/shsg/services/rekursberatung/RekursberatungLeitfaden_EN.pdf

for omsensur, sier reglene gjennomgående at nye sensorer eller klagenemnd skal ha alle sakens dokumenter. Noen steder holder klagenemndene også muntlige høringer hvor student og faglærer hver for seg må stille opp og forklare sitt syn på saken. Generelt fokuserer klagebehandlingen i andre land på å opplyse sakene best mulig, mens man i Norge etterstreber at klagesensuren skal være så lik en vanlig førstegangssensur som mulig, som vist i sitatet fra KDs proposisjon.

Klageordningen ved norske læresteder var altså allerede annerledes i internasjonal sammenheng, og blind omsensur som det eneste tillatte vil gjøre Norge enda mer avvikende i denne sammenhengen. Å ha en ordning som er ulik fra resten av verden er ikke nødvendigvis dumt, det kan også hende at norske myndigheter har tenkt smartere enn alle andre når det gjelder dette sakskomplekset. I neste seksjon vil vi trekke fram noen prinsipper for klagebehandling som kan bidra til å belyse dette spørsmålet.

3. Kvalitetskriterier for klagebehandling

Det fins relativt lite akademisk litteratur om klagebehandling og prinsipper for dette. Gynnild har sammenlignet klagebehandling ved NTNU og i Australia [10] og UiO vs U.C.Berkeley [1]. Noen amerikanske studier [11-12] diskuterer hvordan antall klager kan reduseres ved tydelig kommunikasjon og klare sensurkriterier, men det fins en grense for hva man kan oppnå med dette. Enkelte fag er av en slik natur at eksamensoppgaver må ha et stort løsningsrom hvis de skal teste kreative og konstruktive evner (som f.eks. i informatikk, hvis man skal lære å lage programmer, databaser, IT-arkitekturer osv.), og det er umulig å spesifisere på forhånd alle mulige riktige måter en oppgave kan løses på, eller alle mulige feil som kan gjøres og hvor mye som bør trekkes for hver av disse.

Kravitz et al. [13] trekker opp noen prinsipper for rettferdighet i klagebehandling:

- **konsistens:** prosedyrer bør være de samme fra klage til klage og over tid.
- **upartiskhet:** beslutningstagere har ikke egeninteresse i saken, og er ikke skjevt påvirket av personer med slik egeninteresse.
- **nøyaktighet:** beslutninger skal være basert på så nøyaktig informasjon som mulig.
- **korrigerbarhet:** prosedyrene bør muliggjøre korreksjon av feil
- **representativitet:** synspunktene til alle involverte parter skal tas i betraktning
- **etisk holdbar:** prosedyrene må være i tråd med de involvertes etiske standard

To viktige mål for all karaktersetting er pålitelighet (samme karakter uavhengig av hvem som sensurerer, rekkefølge i bunken, osv.) og gyldighet (karakter er i samsvar med oppnåelse av læringsmål), og dette bør også gjelde klagesensur. Mange av prinsippene listet opp ovenfor kan nettopp sies å bidra til pålitelighet og gyldighet (upartiskhet, nøyaktig informasjon, korreksjon av feil, representativitet). Helst bør klagesensur være mer pålitelig enn førstegangssensur, ellers kan den like gjerne som å rette opp en feil, skape en feil hvor den opprinnelige karakteren var grei. Som basis for analysen er det derfor interessant å sette opp noen kategorier for mulige sensurfeil:

- **uærlig sensur.** Faglærer / sensor har bevisst gitt studenten en ufortjent karakter, enten for god (f.eks. pga. bestikkelser, utpressing, vennetjenester, personlige sympatier) eller for dårlig (f.eks. pga. personlige antipatier eller fordommer basert på rase, kjønn, religion eller lignende).

- **administrative feil.** Faglærer / sensor har gitt, eller ment å gi, en karakter, men i FS står en annen karakter. Eksempel (anekdotisk): En faglærer blir kontaktet av en student som er overrasket over å ha fått D i Programmering etter å ha følt at eksamen gikk greit. Faglærer sjekker regnearket hvor han hadde karakterene og finner at kandidaten står med B. Tydeligvis har det skjedd en feilregistrering på et senere tidspunkt i prosessen, enten ved overføring av karakterer fra regneark til sensurliste på papir, eller ved studieadministrasjonens påfølgende overføring av karakterer fra disse papirlistene og tilbake til et annet datasystem. Slike feil kan selvsagt like gjerne gå i kandidaters favør, men i så fall blir de neppe påklaget.
- **faktaoverseelser.** I motsetning til forrige kategori, har studenten fått den karakteren sensor mente å gi, men feilaktige fakta er lagt til grunn i bedømmelsen – betydelige nok til at studenten ville ha fått en annen karakter hvis overseelsene ikke hadde skjedd. Eksempel (disfavør): På en innledende programmeringseksamen tror sensor at kandidaten ikke har besvart 3b fordi det er blankt mellom 3a og 3c. Men 3b sto noe bortgjemt midt inni et annet delsvare og var greit løst, så studenten burde hatt bra score her, ikke null. På 4a har studenten brukt rekursjon i stedet for løkker, enda rekursjon ikke var pensum i det innledende kurset det er snakk om. Den som retter, ser i farten ikke at løsningen er korrekt pga stort avvik fra løsningsforslaget, og gir lav score. Eksempel i studentens favør: Studenten har skrevet kode som ved første øyekast ser riktig ut, men som egentlig inneholder en del feil, slik som manglende initialisering av en variabel, en løkke som avslutter en runde for tidlig, et metodekall med feil parametre. Sensor oppdager ikke feilene og gir full score.
- **avvikende skjønn.** Eksempel: Studenten er blitt trukket uforholdsmessig mye for småfeil (i programmering f.eks. små syntaksfeil), kanskje vesentlig mer enn det andre kandidater fikk i trekk for lignende feil. Eller: Studenten har fått urimelig god uttelling for et svar som egentlig viser lav forståelse for programmering (f.eks. hvis «åpen bok»-eksamen, en besvarelse som består av kodefragmenter fra boka på en måte som ikke helt passer til det stilte spørsmålet – men hvor studenten gis en god del uttelling da han tross alt åpner en fil og går gjennom en løkke som leser inn data fra fila. Avvikende skjønn kan lett oppstå i store emner med mange ulike sensorer, men kan også forekomme hvis samme sensor svinger over tid, f.eks. avhengig av humør, besvarelsens plassering i bunken, etc.

Av disse fire mulige årsakene til urettferdige karakterer vil vi anta at de to første er sjeldne, slik at urettferdige eksamenskarakterer i all hovedsak skyldes faktaoverseelser og avvikende skjønn. Ikke desto mindre er de to første årsakene alvorlige for studenter som blir rammet av det hvis det skulle forekomme, så vi tar alle fire med i den videre analysen. En god klageordning bør da ha følgende egenskaper:

- *Høy sannsynlighet for å korrigere evt. uærlig sensur*
- *Høyt avskrekkingspotensial for å begå uærlig sensur*
- *Høy sannsynlighet for å korrigere evt. administrative feil*
- *Høy sannsynlighet for å korrigere evt. faktaoverseelser fra opprinnelig sensur*
- *Høy sannsynlighet for å korrigere evt. avvikende skjønn fra opprinnelig sensur*
- *Lav sannsynlighet for nye faktaoverseelser (som opprinnelig sensur ikke begikk)*
- *Lav sannsynlighet for nye skjønnsfeil (som opprinnelig sensur ikke begikk)*
- *Klageprosessen bør fremme studenters læring*
- *Klageprosessen bør fremme lærestedets prosessforbedring (finne årsaken til feil, finne tiltak for å redusere risiko for at lignende feil skal skje igjen)*

- *Klageprosessen bør være kostnadseffektiv* (dvs. ikke bruke ressurser på en måte som ikke gir økt kvalitet i beslutningene)

Når det gjelder å korrigere for evt. feil i opprinnelig sensur, må det bemerkes at klageordningen primært vil adressere feil i studentens uheld, siden tilfeller med betydelige feil i studentens fordel sjelden blir gjenstand for klage.

4. Analytisk sammenligning av sensurordninger

I det følgende vil vi sammenligne tre forskjellige klageordninger i forhold til kriteriene listet opp over. De tre ordningene er blind omsensur, informert omsensur, og hva vi med et samlebegrep kaller «utenlandsk» ordning. Sistnevnte antas å ha noe i retning av følgende egenskaper: Studenten må først ta uformell kontakt med faglærer for å få forklaring på karakteren og prøve å finne en minnelig løsning, og kan bare levere klage hvis dette ikke fører fram. Klagen må i så fall begrunnes, med konkret henvisning til tilfeller av feilbedømmelse, og de som skal vurdere klagen får tilgang til all relevant informasjon. Opprinnelig sensur overprøves bare hvis man finner at den hadde konkrete feil (f.eks. faktaoverseelser) eller var urimelig.

Tabell 3-5 viser sammenligning av de tre sensurordningene etter nevnte kriterier. Hvit rute betyr "bedre", jo mørkere gråtone, jo dårligere fremstår ordningen sammenlignet med konkurrentene. Det er generelt verdt å merke seg at "Lav" betyr lavere enn konkurrentene, ikke nødvendigvis lav sannsynlighet i absolutt forstand (dvs. sannsynlighet for suksess kan godt være 90% men likevel bli karakterisert som lav hvis konkurrentene anses å ha 95 eller 99%). I tabell 3 ser vi på kriteriene som omhandler uærlig sensur og administrative feil. Disse to feilkategoriene anses som de minst sannsynlige men er likevel viktig å begrense så mye som mulig. Blind omsensur kan anses å ha en fordel når det gjelder å korrigere uærlig sensur, siden førstegangs sensor, ikke har noen formelle muligheter til å påvirke omsensuren med en tendensiøs begrunnelse¹⁴. Men på de to andre kriteriene kommer blind omsensur dårligere ut. Omsensuren kommer da bare med en ny karakter (forhåpentligvis riktigere), de nye sensorene har ingen formening om at den avvek fra forrige karakter, og administrasjonen som mottar rapport om ny karakter - og vil se et eventuelt avvik - har på den annen side ingen forutsetninger for å vurdere faglige årsaker til dette. Med informert omsensur, hvor faglærer må produsere en begrunnelse som skal stå seg mot studentens tilsvarende overfor de nye sensorene, vil eventuelt tendensiøs begrunnelse ha større risiko for å vekke oppsikt. Dette blir enda mer påfallende med "utenlandsk" ordning, hvor faglærer må forsvare begrunnelsen overfor en overordnet instans (klagenemnd, dekan). Mangelen på kobling til opprinnelig sensur ved blind omsensur reduserer også sjansen for å korrigere en administrativ feil. La oss si at studenten fikk E ved opprinnelig sensur men pga en registreringsfeil får F i studiesystemet. Hvis studenten klager uten å be om begrunnelse (som er vanlig når man får F), og vi antar at utfallet av den blinde omsensuren fordeler seg rundt den intenderte karakteren fra første sensur slik som Gabrielsen et al. [6] indikerte, er det fortsatt nær 10% sjanse for kandidaten får F også ved omsensuren. Det kan selvsagt hende ved et sammentreff at

¹⁴ Dog kan denne fordelingen være tvilsom, for blind omsensur innebærer at heller ikke studenten har mulighet til å komme med innspill til de nye sensorene. En uærlig faglærer kan i fravær av formelle muligheter til å påvirke klagesensuren også ha uformelle kanaler til disposisjon (f.eks. ved kjennskap til hvem klagesensorene er), eller kan ha andre triks på lager (produsere en sensorveiledning som er strengere enn opprinnelig rettespråk). Studenten har kanskje bare de formelle påvirkningsmulighetene. Således kan det hende at blind omsensur forskyver balansen i en uærlig faglærers favør, og at studenten ville ha vært bedre tjent med et mer transparent system hvor begge fikk uttale seg til klagesensorene.

dette var en "riktigere" karakter, men det essensielle poenget er at kandidaten neppe ville ha klaget på E i utgangspunktet – og aldri skulle ha behøvd å klage til omsensur, da administrative feil burde løses på andre og enklere måter.

Tabell 3: Sammenligning av tre klageordninger: uærlighet og administrative feil

Kriterium	Blind omsensur	Informert omsensur	"Utenlandsk"
Korrigere uærlighet	Høy?? (Men se fotnote 14 på forrige side)	Middels?	Middels?
Avskrekke uærlighet	Lav (klagebehandlingen vurderer ikke kvaliteten av første sensur, kun besvarelsen på nytt, uærlighet blir dermed ikke eksponert)	Middels (klagebehandling vil i det minste se misforhold ved opprinnelig sensur, men primært mandat er vurdering av besvarelsen)	Høy (i et system hvor klage går til dekan, overordnet nemnd e.l., vil det føles verre å skulle forsvare en uærlig sensur, og en faglærer som blir en gjenganger i slike saker kan få et dårlig rykte)
Korrigere administrativ feil	Lav (oppdager ikke at adm feil var årsaken, korreksjon forutsetter da at klagesensur gir samme resultat som intendert i opprinnelig sensur, dette skjer bare i ca. 70% av tilfellene)	Middels (hvis første sensor må skrive begrunnelse ved klage, vil denne med høy sannsynlighet oppdage at en adm. feil har skjedd.)	Høy (adm feil kan oppdages i uformell dialog mellom student og faglærer og saken trenger aldri gå til noen omsensur)

I Tabell 4 sammenlignes ordningene med hensyn på sjansen til å korrigere eller unngå faktaoverseelser og avvikende skjønn. For feil fra opprinnelig sensur som var i studentens disfavør, kan vi se tilbake på eksempler fra seksjon 3: en deloppgave som feilaktig ble antatt blank fordi den var bortgjemt i besvarelsen, en rekursiv løsning som fikk dårlig uttelling fordi den var helt ulik løsningsforslag, eller små syntaksfeil som det ble trukket uforholdsmessig mye for. Med blind omsensur er det fullt mulig at man gjør samme overseelser som førstegangs sensur. Ved informert omsensur kunne studenten begrunne sin klage ved henvisninger til faglærers begrunnelse: "Jeg har fått 0 poeng for 3b. Denne er imidlertid besvart midt på ark 7, og på en måte som burde gi god uttelling. ... 4a. Min løsning er annerledes enn LF fordi jeg har brukt rekursjon, ikke pensum men det sies ikke noe sted at det ikke er lov. Hvis dere taster inn og kjører funksjonen vil dere se at den virker perfekt.."

Med mindre klagesensorene overhodet ikke gjør jobben sin (f.eks. lar være å lese studentens begrunnelse for klagen), må det anses langt mindre sannsynlig at de skal kunne overse de nevnte feilene fra første sensur etter å ha lest begge parters begrunnelser enn hva som kunne skje ved blind omsensur der de bare har oppgavetekst, besvarelse og løsningsforslag. Det ville også være merkelig å anta at klagesensorene pga. kollegial "lojalitet" til faglærer skulle unnlate å ta studentens argumenter til følge i en situasjon hvor de er lett etterprøvbare og åpenbart riktige. Hvis faglærer skulle være en person klagesensorene har stor faglig respekt for, vil de med stor sannsynlighet tolke de påviste tilfellene ved opprinnelig sensur som en glipp, ikke som noe faglærer faktisk har ment (dvs., faglærer har neppe ment at et delsvar som er plassert på et uventet ark skal ha null poeng, eller at en perfekt fungerende rekursiv løsning skal gi dårlig uttelling). I den grad "lojalitet" overhodet skulle være en faktor, ville klagesensorene da tenke: Hva ville faglærer selv ha ønsket her? (hvis det f.eks. var en annen sensurprosess hvor faglærer selv skulle revurdere karakter ved denne typen innspill fra studenter). Det

overveiende mest sannsynlige er vel i så fall at faglærer selv også ville ha ønsket at feilene ble korrigert og at studenten fikk den fortjente uttellingen.

Tabell 4: Sammenligning av tre klageordninger: feil mhp fakta og skjønn

Kriterium	Blind omsensur	Informert omsensur	"Utenlandsk"
Korrigere faktafeil og skjønnfeil som var i kandidats disfavour	Lav (studentens begrunnelse kunne ha påpekt slike feil i første sensur, men noen slik begrunnelse får disse klagesensorene ikke se)	Høyere (studentens begrunnelse for å klage kan påpeke slike faktafeil)	Høyere (faglærer / første sensor har rettetreningen og sammenligningsgrunnlaget, og har da høy kompetanse til å sjekke om innspill fra student indikerer at det er begått en feil)
Korrigere faktafeil og skjønnfeil som var i favør	Høyere (flere går ned i karakter med blind omsensur enn med informert omsensur fordi blind sensur har generelt høyere varians. Liberal klageordning «forleder» en del studenter til å klage i risikable situasjoner?)	Middels (få går ned i karakter. Kan tenkes at faglærers begrunnelse blir en sovepute her? Nye sensorer ser ikke etter andre svakheter i tillegg til de som faglærer har påpekt, og studenten påpeker dem neppe heller)	Lav (i mange land tilsier ordningen at karakter ikke kan settes ned. Høyere krav til studenten i forbindelse med klage gjør dessuten at færre klager uten å ha en god sak.
Unngå nye faktafeil	Lav: (i) ingen hjelp fra observerte fakta i første sensur; (ii) ingen hjelp fra studentens påstand om fakta, og (iii) liten rettetrening for dette oppgavesettet	Middels. Har hjelp av argumenter fra faglærer og student (hvor sistnevnte vil fokusere på fakta i sin favør, mens sistnevnte også dekker fakta i disfavour)	Høy: (i) vet hva det er trukket for, (ii) får innspillene fra student, (iii) har erfaring i å rette oppgavesettet. Ingen grunn til at faglærer skal gjøre nye feil som ikke ble gjort første gang.
Unngå nye skjønnfeil	Lav: (i) ingen hjelp fra forklaring av skjønn i opprinnelig sensur, (ii) ingen hjelp om student har argument relatert til skjønn, (iii) mindre sammenligningsgrunnlag enn førstesensorene, (iv) mindre detaljert kjennskap til emnet enn faglærer og tilsynssensor har.	Middels: Har hjelp av argumenter om skjønn fra førstesensor og student. Men fortsatt mindre grunnlag for sammenligning enn førstesensorene, og kanskje mindre detaljert kjennskap til emnet enn faglærer og tilsynssensor har.	Høy: Den generelle holdningen er at man ikke skal overprøve faglærers skjønn, annet enn i ekstreme tilfeller.

Et annet viktig ankepunkt mot blind omsensur er at det kan ha større risiko for å introdusere nye feil i sensuren. Anta igjen eksempler fra seksjon 3, en student har skrevet kode som på overflaten kan se riktig ut, men som ved nærmere inspeksjon er belemret med en del feil (manglende initialisering av variabel, løkke som går en runde for lite, metodekall med feil parametre). Anta at førstegangssensur oppdaget alle disse feilene. Ved blind omsensur er klagesensorene avhengig av å finne alle feilene i besvarelsen selv, da de ikke har noe materiale som gjør oppmerksom på allerede detekterte feil. Å rette store bunker med eksamener er læringsprosess hvor man etter hvert blir gradvis flinkere til å finne selv de mest subtile feilene. Mange faglærere (forfatteren inkludert) går gjerne tilbake og ser over de 10 første besvarelsene på nytt

når man er igjennom hele bunken – fordi man helt i starten var mer usikker. Selv om klagesensorene kan ha høy kompetanse på retting innen faget generelt, har de ikke samme trening på å rette det spesifikke oppgavesettet. Ved noen læresteder som bruker blind omsensur, prøver man å kompensere for dette ved å legge ved noen besvarelser i tillegg til den / de som er påklaget, og hvor karakter på disse andre opplyses. Dette gir dog ingen garanti for å unngå tabber i klagesensuren. I programmering eller lignende emner er det mange ulike feil som kan gjøres i en besvarelse, og det skal mye til at de vedlagte eksempelbesvarelsene har gjort akkurat samme slags feil som de påklagde.

Tabell 5: Sammenligning av tre klageordninger: kostnad, læring, prosessforbedring

Kriterium	Blind omsensur	Informert omsensur	"Utenlandsk"
Kostnad	Høy (omsensur x2) uavhengig av klagens årsak og substans	Høy (omsensur x2), men kan spare litt hvis faglærer er enig med studenten?	Lavere (færre klager å behandle: noen saker løses uformelt, og klager uten substans kan avvises)
Læring for student	Ingen (studenten trenger ikke begrunne klagen)	Middels (lærer av å måtte begrunne klagen sin; om begrunnelse ikke skulle være påkrevd ved lov kan det være de facto nødvendig for å lykkes ved informert omsensur, for å påpeke evt. feil i faglærers begrunnelse)	Høy (lærer av uformell dialog med faglærer, samt av å begrunne klagen sin)
Prosessforbedring	Ingen (blind omsensur vurderer ikke den første sensuren overhodet, det bare kommer en ny karakter som erstatter den gamle. Man har da ingen anelse om det var den gamle eller nye karakteren som var "feil", og om forskjell skyldtes adm feil, faktaoverseelser eller avvikende skjønn.)	Mulig (ingen garanti for at prosessforbedring skjer, men det er iallfall mulig. Klagesensorene ser den opprinnelige begrunnelsen og kan eventuelt bemerke hva som er årsak til diskrepans i karakter og formidle dette videre til rette vedkommende som har mulighet til å redusere risiko for lignende feil i fremtiden.)	Bedre: mandat for klagebehandler er i større grad å vurdere kvaliteten av opprinnelig sensur, ikke kun å gjøre en ny sensur. Ved å behandle færre klager (kun de med substans) muliggjøres grundig og prestisjetung behandling av disse får (nemnder / dekan), som innebærer større press på fagmiljø om å redusere risiko for lignende feil i fremtiden

Tabell 5 ser på kostnader, læring og prosessforbedring. Understøttning av læring er minimal så lenge studenten verken trenger å ha noen dialog med faglærer eller begrunne sin klage. Som påpekt av Gynnild [1], virker det som man ved amerikanske universiteter har mye større fokus på læringsaspektet ved klagerutinene, og at det har vært overraskende lite interesse for dette i Norge. Å understøtte prosessforbedring er heller ikke lett når klagebehandlingen ikke ser noen relasjon til den opprinnelige sensuren, så man ikke kan vite hva som eventuelt var galt med den og hvorfor. Det ligger i den norske ordningens natur at den er kostbar, siden den pålegger universitetene å gjøre omsensur med to nye sensorer for absolutt alle karakterklager som kommer, enten de måtte være substansielle eller ikke. Ved informert omsensur kan man muligens spare noe i forhold til blind omsensur i tilfeller hvor faglærer etter å ha blitt kontaktet av kandidaten, er enig med dennes innvendinger (f.eks., faglærer ser at han har begått en glipp, som å overse at oppgave 3b var besvart på et annet ark, eller ser at han har trukket

denne kandidaten vesentlig mer for noen små syntaksfeil enn det som ble gjort med andre kandidater). Det beste i så fall ville jo være om faglærer selv kunne ta initiativ til å endre karakteren (a la utenlandsk ordning) så den ikke trenger å gå til noen omsensur, men i den grad omsensur er påkrevd hvis klagen først er reist, kan faglærer i det minste spare litt arbeid for de nye sensorene ved å selv sende inn en revidert begrunnelse hvor han sier seg enig i studentens anførsler.

4. Diskusjon og konklusjon

Alt i alt indikerer analysen av at blind omsensur kommer dårligere ut enn informert omsensur, og at begge de norske ordningene kommer dårligere ut enn "utenlandsk" ordning – alt dette vel å merke basert på forfatterens subjektive "forsøk på sunn fornuft"-argumenter heller enn noen empirisk sammenligning av de tre sensorordningene. Denne sammenligningen kan selvsagt virke noe bombastisk, for eksempel sett i forhold til sammenligningen som gjøres av Gabrielsen et al. [6], som nøkternt skriver at tallene deres kan tolkes på ulike måter: «*De som foretrekker den nye ordningen, vil gjerne hevde at de påviste forskjellene skyldes at klagesensorene i den gamle ordningen i for stor grad er påvirket av førstesensorenes begrunnelse og karaktersetting. Forsvarerne av den gamle ordningen vil på den andre siden kunne si at når mer enn hver tredje klage etter den nye ordningen resulterer i karakterendring, har det sin bakgrunn i at klagesensorene der mangler det viktige sammenligningsgrunnlaget av andre eksamensbesvarelser som førstesensorene har.*» [6] (s.20).

Konklusjonen i denne artikkelen er at Gabrielsen et al. er inne på noe i den siste setningen sitert over, men at det nevnte sammenligningsgrunnlaget ikke er den eneste utfordringen for blind omsensur, og kanskje ikke heller den verste. Hovedproblemet er at blind omsensur har høyere sannsynlighet enn andre ordninger for å introdusere nye feil i sensuren som ikke ble begått første gang, både faktaoverseelser (fordi klagesensorene ikke har vært igjennom læringsprosessen fra å rette et stort antall besvarelser for samme oppgavesett) eller avvikende skjønn (dels sammenligningsgrunnlaget, men også fordi de neppe har like detaljert kjennskap til eksakt pensum og hvordan dette har vært vektlagt i undervisningen som opprinnelige sensorer hadde). Siden man tillater ubegrunnede klager, noen motivert bare ut fra at man strøk eller var nær en høyere karakter, er det stor risiko for at mange som hadde en rettferdig karakter fra første sensur, får en mindre rettferdig karakter (til gunst eller ugunst) ved klagen.

Mest bekymringsfullt er kanskje holdningene KD kommuniserer i sin preferanse for klageordning. Andre lands ordninger er preget av en grunnleggende tillit til faglærer altså en tro på at denne selv har kompetanse til, og ønske om, å rette opp sensurfeil hvis studenter påpeker slike. Den norske klageordningen gjennomsyres derimot av en grunnleggende mistillit til faglærere. Når proposisjonen sier at blind omsensur gir best garanti for upartisk vurdering av en klage, kan dette vanskelig tolkes på annen måte enn at faglærer antas å være partisk, og at klagesensorer også vil bli partiske hvis de ser faglærers argumenter. Forvaltningsloven er imidlertid klar på at man kun er part i en sak hvis man har en personlig interesse av utfallet, og en NAV-ansatt som innvilger eller avslår en søknad om stønad, regnes ikke som part i saken, heller ikke om søkeren deretter klager vedtaket inn for Fylkesmannen (som selvsagt også ser hva det opprinnelige vedtaket var, uten å bli regnet som partisk av den grunn). Faglærere har normalt ingen personlig interesse av hvilken karakter en student får, og en begrunnelse vil inneholde både positive og negative synspunkter på en besvarelse (hva man har fått uttelling for, hva det trekkes for) og er således en redegjørelse for en beslutning, ikke et partsinnlegg for at studenten skal få dårligst mulig karakter. I Norge er frykten for

partiske faglærere tydeligvis så stor at man heller foretrekker en klageordning med økt vilkårlighet som følge av mangelfull opplysning av sakene.

Noe som overraskende nok ikke ser ut til å ha vært diskutert i høringsrunden, er en mulig kombinasjon av blind og informert sensur, a la vitenskapelig review (hvor man får se andres etter at man har lagt inn sin egen). Hvis klagesensorene først gjør sin egen vurdering og sender inn denne, unngår man at funn fra første sensur blir en sovepute. Hvis de deretter får se faglærers og studentens synspunkter, kan de sammenholde alle opplysninger og vurdere hva som blir riktigst konklusjon. Eventuelle feil i opprinnelig sensur kan da også rapporteres til rette vedkommende for å bidra til prosessforbedring. Men man ville fortsatt behandle et stort antall klager, med mindre man kan forlange begrunnelser med en viss substans.

Referanser

1. Gynnild, V., *Student appeals of grades: a comparative study of university policies and practices*. Assessment in Education: Principles, Policy & Practice, 2011. **18**(1): p. 41-57.
2. Bjølseth, G., A. Havnes, and P. Lauvås, *Lavt sensorsamsvar-kan det bedres?* UNIPED, 2011. **34**(4).
3. Rye, J.F., *Konsistente karakterer? En undersøkelse av sensorreliabilitet i sosiologifaget*. UNIPED, 2014. **37**(3).
4. Andersen, R., et al., *Project courses at the NTH: 20 years of experience*, in *Software Engineering Education*, J. Díaz-Herrera, Editor. 1994, Springer Berlin Heidelberg. p. 177-188.
5. Jaccheri, L. and G. Sindre. *Software engineering students meet interdisciplinary project work and art*. in *Information Visualization, 2007. IV'07. 11th International Conference*. 2007: IEEE.
6. Gabrielsen, E., I.Å. By, and M. Randen, *Klagesensurordningene - hvordan fungerer de?* UNIPED, 2012. **35**(1): p. 5-21.
7. Kristiansen, M., T. Heine Nätt, and C.F. Heide, *Kvantitativ undersøkelse av mulige sammenhenger mellom vurderingsform og karakterer i høyere utdanning*. UNIPED, 2013. **36**(2).
8. Visser, B.A., M.C. Ashton, and P.A. Vernon, *What makes you think you're so smart? Measured abilities, personality, and sex differences in relation to self-estimates of multiple intelligences*. Journal of Individual Differences, 2008. **29**(1): p. 35.
9. Kanten, A.B. and K.H. Teigen, *Better than average and better with time: Relative evaluations of self and others in the past, present, and future*. European Journal of Social Psychology, 2008. **38**(2): p. 343-353.
10. Gynnild, V. and L. Therry. *An investigation of student appeals and critical review*. In *International Conference on Engineering Education (ICEE)*. 2003.
11. Germaine, R., et al., *Lessons learned through a review of grade appeal data*. Publication of National University, 2011: p. 39.
12. Pulich, M.A., *Student Grade Appeals Can be Reduced*. Improving College and University Teaching, 1983. **31**(1): p. 9-12.
13. Kravitz, D., E. Stone-Romero, and J. Ryer, *Student Evaluations of Grade Appeal Procedures: The Importance of Procedural Justice*. Research in Higher Education, 1997. **38**(6): p. 699-726.