

Giver kvinder og mænd forskellige karakterer? Køn og karaktergivning på universitetet

Lotte Bøgh Andersen, ph.d. og professor ved Institut for Statskundskab, Aarhus Universitet.

Heidi Houlberg Salomonsen, ph.d. og lektor ved Institut for Statskundskab, Aalborg Universitet.

Reviewet artikel

Er kvindelige universitetsansatte mere orienteret mod at hjælpe den enkelte studerende end deres mandlige kolleger, og giver de derfor bedre karakterer? Artiklen undersøger dette i en panelundersøgelse af bedømmelsesadfærd på to danske universitetsinstitutter mellem 2004 og 2010 (381 eksaminationer fordelt på 78 universitetslærere). Undersøgelsen viser, at selvom kvindelige bedømmere faktisk er mere orienteret mod den enkelte studerendes ønsker, og selvom dette generelt er forbundet med at give højere karakter, giver kvinderne ikke signifikant anderledes karakterer end deres mandlige kolleger. Disse tilsyneladende modstridende resultater kan skyldes, at kvinder modvirker kønsstereotype forestillinger og/eller, at faglige normer betyder mere end køn. Det er betryggende i den forstand, at studerende ikke bliver behandlet forskelligt alt efter, om de bliver bedømt af en mand eller en kvinde.

Introduktion

Universitetsstuderende har krav på en saglig og ensartet vurdering af deres eksamenspræstationer, og det er yderst uheldigt, hvis bedømmerens personlige karakteristika har betydning. Kvinder er gennemsnitligt mere empatiske end mænd (Nielsen & Salomonsen, 2010), hvilket betyder, at de i højere grad ser tingene fra andres perspektiv og bekymrer sig om andre. Ifølge Cornwell (1974) handler det om psykologisk projektion ind i den rolle, modtageren har: "Putting oneself in the other person's shoes". Empati kan være glimrende i forhold til mange former for undervisning (Rhodes, 2010; Duncan-Andrade, 2011), men i forhold til karaktergivning kan bedømmernes varierende empati give forskellige bedømmelser (Anderson, 2013). Konkret er der således risiko for, at kvindelige universitetslærere giver højere karakterer end mændene. Eftersom empati er knyttet til indlevelse i de studerendes situation, kunne kvindelige bedømmere således tænkes at være mere orienterede mod den enkelte studerendes ønsker (dvs. have højere grad af studenterorientering) og på den baggrund give bedre karakter. Karaktergivning handler om at afveje forskellige hen-

syn, og empati kunne øge prioriteringen af hensynet til den enkelte studerende i forhold til andre hensyn (fx til at undgå karakterinflation). Vores viden om køns effekter på undervisningsområdet er imidlertid begrænset, om end nogle studier har taget spørgsmålet op (Carrington et al., 2008; Dee, 2005; Martin & Marsh, 2005; Andersen & Heinesen, 2013). Især vedrørende medarbejdernes køn og bedømmelsesadfærd på de videregående uddannelser findes der kun få studier (Hoffman & Oreopoulos, 2007, s. 6; Crombie et al., 2003, s. 55; Goodwin & Stevens, 1993, s. 167). I Danmark påviser Hanne Nexø Jensen (2010, s. 19), at mandlige vejledere taler fire femtedele af tiden under specialevejledningerne, mens kvindelige vejledere kun taler knap tre femtedele af tiden. Det kunne tyde på, at kvinderne er mere studenterorienterede i den forstand, at de er mere fokuserede på at bidrage til at opfylde den enkelte studerendes ønsker. Nielsen, Andersen & Pedersen (2012) finder imidlertid, at specialevejllederens køn ikke har nogen betydning for karakteren. Der er ingen danske studier af betydningen af bedømmerens køn for eksaminer, hvor bedømmer og studerende ikke kender hinanden. Den gennemsnitlige forskel i empati mellem kønnene gør det muligt, at der er en kønsforskning i bedømmelserne, og hvis det er tilfældet, må vi have viden om det for at kunne gøre noget ved det. Hvis der ingen kønsmæssig forskel er, vil viden derom kunne forebygge fejlagtige forestillinger om kønnets betydning for karaktergivning. Dette studium begynder at udfylde hullet i den eksisterende forskning ved at lave en undersøgelse af karaktergivningen på to universitetsinstitutter over en årrække. Det er meget relevant at få undersøgt, for hvis der er kønsforskelle i bedømmelsesadfærd, betyder det en ulige behandling af de studerende, og det gør det endnu mere problematisk, at køn er et synligt personligt karakteristika, så de studerende (fx på overbygningsfag) kan vælge underviser efter, om de gerne vil bedømmes af en kvinde. Hvis der faktisk er en kønsforskning, bør den belyses, så den kan håndteres på en ordentlig måde, og hvis der ikke er en kønsforskning, bør eventuelle fordomme afkræftes. Forskningsspørgsmålene i artiklen er således: Er kvindelige universitetsbedømmere mere studenterorienterede end deres mandlige kolleger? Hænger denne studenterorientering sammen med karaktergivningen? Og giver kvinderne som følge heraf bedre karakterer?

Undersøgelsen er foretaget på to statskundskabsinstitutter mellem 2004 og 2010 på baggrund af information fra 381 eksaminer (kombinationer af en given bedømmer og et givent fag) for 78 universitetslærere. Vi kobler objektive karakteroplysninger sammen med spørgeskemaundersøgelser for netop de bedømmere, der har foretaget bedømmelserne. Spørgeskemaundersøgelsen målte blandt andet bedømmernes studenterorientering forstået som deres fokus på at bidrage til opfyldelse af den enkelte studerendes ønsker, og vi har dermed mulighed for at se, om studenterorientering hænger sammen med karaktergivning (og analysere, hvilken rolle køn spiller i den forbindelse). Fordi forskningen viser, at kvinder generelt set har mere empati end mænd, forventes de at være mere studenterorienterede. Det kan gøre en forskel i

bedømmelsessituationer (Anderson, 2013), da indlevelse i den enkelte studerendes situation kan føre til et forstærket ønske om at stille vedkommende bedre (ved at give en bedre karakter). Det diskuteres indgående i næste afsnit, hvorefter vi fremlægger vores forskningsdesign og metoder efterfulgt af vores resultater. Artiklen afrundes med en diskussion af resultaterne, ligesom vi også kommer ind på artiklens begrænsninger og behovet for videre forskning på området.

Hvorfor skulle mænd og kvinder bedømme forskelligt?

Litteraturen om undervisning af voksne ser normalt empati som en positiv egenskab (Rhodes, 2010; Duncan-Andrade, 2011), men nyere studier er begyndt at diskutere, om empati også kan have negative effekter. Potentiel forskelsbehandling ved eksamen er blot en af disse effekter, og Gair (2011, s. 493) nævner også udbrændthed, overføring og 'compassion fatigue'. Når det er specielt relevant at tage betydningen af empati for bedømmelsesadfærden op, skyldes det hensynet til de studerendes retssikkerhed koblet med vores viden om, at køn og empati hænger sammen. Flere studier underbygger, at kvinder tenderer til at have mere empati forstået som evnen til at forstå andre menneskers følelser og hurtigt og automatisk at relatere til disse andre menneskers følelser (Yang et al., 2009, s. 176; Baron-Cohen, 2003; Kanter, 1977; Hale, 1999; Olesen, Aggerholm & Kofoed, 2008; Nielsen & Salomonsen, 2010). Der er uenighed blandt forskerne på området om, hvorvidt kvinders gennemsnitlige højere grad af empati skyldes biologi eller socialisering, og der findes støtte for begge forklaringer (Connellan et al., 2000; Olesen et al., 2008; Kanter, 1977; Hale, 1999). I denne sammenhæng er det vigtigste imidlertid, at kvinder generelt er mere empatiske end mænd. En række undersøgelser viser således, at kvinder og mænd inden for samme faggruppe har forskellig grad af empati (Nielsen, 2013; Kalaian & Freeman, 1994, s. 656; Barber, 2002, s. 392).

Vores argument er, at empati (og dermed også køn) især er relevant for karaktergivningen, fordi den fremmer en bestemt type prosocial motivation kaldet brugerorientering (se fx Andersen, Pallesen & Salomonsen, 2013). Brugerorientering handler i konteksten af voksenundervisning om at hjælpe den enkelte studerende med at opnå det, som den pågældende selv ønsker at opnå. Vi argumenterer for, at universitetslærere, der i høj grad evner at forstå de studerendes følelser og hurtigt og automatisk relaterer til disse følelser (dvs. er meget empatiske), også vil være motiverede for at handle på en måde, der hjælper de studerende som enkeltpersoner. Det er næppe en urealistisk antagelse, at den enkelte studerendes i forbindelse med eksamen ønsker at bestå og få så høj en karakter som muligt. Det er imidlertid ikke givet, at der er en forskel i mænds og kvinders studenterorientering, da selektionen før fastansættelse på universitetet er meget stor, og socialiseringen i forhold til adfærd og motivation er kraftig. Selvom der skulle være en kønsforskell i studenterorientering, påvirker det

ikke nødvendigvis kvinderne og mændenes karaktergivning (som vi kommer ind på nedenfor). Hvis de initiale kønsforskelle i empati ikke udlignes, giver det imidlertid anledning til den forventning, der er udtrykt i hypotese 1:

H1: Kvindelige bedømmere af universitetseksaminer er mere studenterorienterede end deres mandlige kolleger.

Studenterorienteringen (forstået som fokuset på at hjælpe den enkelte studerende med at opnå dennes ønsker) er relevant for karaktergivning, fordi kombinationen af identifikation med de studerende og en orientering mod at gøre godt for konkrete andre mennesker kan få bedømmerne til at tippe opad på karakterskalaen frem for nedad, når præstationen ligger mellem to karakterer. Modsætningen mellem at være orienteret mod at tilgodese hhv. enkeltindviders og det bredere samfunds interesser findes også i andre sammenhænge og bliver i motivationslitteraturen indfanget af begreberne brugerorientering og public service motivation (se fx Andersen, Pallesen & Pedersen, 2011 samt Andersen & Pedersen, 2013). Det viser sig eksempelvis, at mens praktiserende læger med høj grad af brugerorientering ordinerer mere antibiotika per patient end deres kolleger med mindre grad af brugerorientering, holder læger med høj grad af den samfundsrettede public service motivation igen med bredspektret antibiotika, der især indebærer risiko for resistente bakterier (Jensen & Andersen, 2013). Argumentet i nærværende artikel er tilsvarende, at karaktergivning tilgodeser flere typer hensyn, nemlig hensyn til den enkelte studerendes ønsker og hensyn til studenterne som helhed, kommende arbejdsgivere og omdømmet for uddannelsen og karakterskalaen. Mens hensynet til den enkelte studerendes ønsker tilsiger højere karakterer, taler de øvrige hensyn for en fair og saglig karaktergivning, hvor resten af samfundet kan regne med, at en given karakter er udtryk for et givent niveau (Andersen, Pallesen & Salomonsen, 2013). Vores forventning er, at universitetsbedømmere med større studenterorientering tenderer til at give højere karakterer, fordi hensynet til den enkelte studerendes ønsker fylder relativt mere i sammenligning med hensynet til omgivelserne. Denne forventning er udtrykt i hypotese 2, som er underopdelt i to målbare implikationer. Hvis en universitetsbedømmer er meget studenterorienteret, forventes dels et højere karaktergennemsnit, dels en højere beståelsesprocent.

H2a: Bedømmere af universitetseksaminer, som har større studenterorientering, har en større beståelsesandel af deres bedømte eksaminer

H2b: Bedømmere af universitetseksaminer, som har større studenterorientering, har et højere gennemsnit af deres bedømte eksaminer

Hvis hypotese 1 og hypotese 2 kan bekræftes, er den logiske implikation, at kvindelige bedømmere giver bedre karakterer, fordi de er mere studenterorienterede. Denne forventning er udtrykt i hypotese 3, der ligesom hypotese 2 er opdelt i to testbare implikationer.

H3a: Kvindelige bedømmere af universitetseksaminer har en større beståelsesandel af deres bedømte eksaminer end mandlige bedømmere.

H3b: Kvindelige bedømmere af universitetseksaminer har et højere gennemsnit af deres bedømte eksaminer end mandlige bedømmere.

Selv hvis kvindelige universitetsbedømmere skulle være mere studenterorienterede, er det dog langt fra givet, at de lader det spille en rolle, når de bedømmer eksamensopgaver. Eksempelvis kunne kvindelige bedømmere aktivt modvirke de kønsstereotype forestillinger og søge at lægge sig på præcis det samme niveau som mændene. Kønsstereotyper handler om den måde, hvorpå kvinder og mænd formodes at handle, men det er før set, at undervisere har modvirket disse forestillinger (Sabbe & Aelterman, 2007, s. 530). De undersøgte universiteter har traditionelt været domineret af mænd og mandlige normer, og kvinder, der vil overleve i det system, kan derfor føle en tilskyndelse til at undgå at blive fanget i en feminin rolle. Her kan normerne inden for faget også spille en central rolle. Et argument for, at hverken studenterorienteringen eller køn skulle have en betydning, er således, at adfærd i organisationer bliver reguleret af en række formelle, uformelle og underforståede institutionelle regler (Scott, 2001). Institutioner kan både påvirke adfærd direkte og have betydning for andre forholds betydning for adfærden. Vi har testet begge påvirkningstyper og finder kun den direkte betydning, som fremgår af de viste resultater. Selvom betydningen af institutionerne er interessant i sig selv, er fokus på sammenhængen mellem køn, studenterorientering og karaktergivning. Vi inddrager således styrken af de kollegiale normer, om der er ekstern censor, samt hvor meget stiltiende institutioner (fx underforståede, tavse regler) opleves at regulere karaktergivningen. Alle disse ting kan forventes at modvirke en effekt af studenterorientering og køn, og vi kontrollerer også for, på hvilken organisation universitetsbedømmeren er ansat, da der kan være organisatorisk variation i karaktergivningen. Hvordan disse variable mere specifikt måles, og hvordan studiet i det hele taget er gennemført, diskuteres i næste afsnit.

Forskningsdesign og metoder

Eftersom undersøgelsen her søger at isolere effekten af bedømmernes køn for karaktergivningen, har vi bestræbt os på at holde alle andre variable konstante ved at undersøge to statskundskabsinstitutter over tid. Når vi har valgt at undersøge to institutter frem for et enkelt, skyldes det dels ønsket om at sikre, at sammenhængen ikke kun findes i en enkelt organisatorisk sammenhæng, og dels behovet for et tilstrække-

ligt antal observationer til at drage sikre slutninger. Sidstnævnte tilgodeses også ved at inddrage observationer fra i alt 7 år dvs. 14 eksamensterminer (2004-2010). I analysen af hypotese 1 (sammenhængen mellem køn og studenterorientering) er analyseenhederne dog de enkelte universitetsbedømmere, mens analysen af hypotese 2 og 3 ser på kombinationer af fag og bedømmere i en given termin. Her kan en given bedømmer i faget Videnskabsteori i efteråret 2007 fx have et karaktergennemsnit på 7,02. Det tæller som en observation. På denne måde analyserer vi 381 observationer (kombinationer af en given bedømmer og et givent fag) for 78 universitetslærere. Spørgeskemaundersøgelsen blev udsendt via e-mail, og svarprocenten var 89 (78 ud af 88 mulige respondenter). Appendikset giver overblik over de to datasæt med hhv. bedømmerne og eksaminationerne som analyseenheder.

Vi undersøger udelukkende bedømmelser på de store fag på bacheloruddannelsen, idet der er flere bedømmere på disse fag end undervisere, bedømmerne har begrænset kontakt til de studerende, og de studerendes eksaminer bliver tilfældigt fordelt på bedømmerne. Modsat vælger de studerende sig selv ind på bestemte overbygningsfag med bestemte undervisere og dermed også bedømmere, fordi underviseren i dette tilfælde bedømmer alle opgaverne. For overbygningsbedømmelser ville der i tilgift til bedømmelseeffekten også potentielt være en selektionseffekt og en undervisningseffekt, hvilket er hovedårsagen til alene at undersøge bedømmelser på bacheloruddannelsen. For hovedparten af de undersøgte bedømmelser (75 %) er besvarelserne helt anonyme, mens den sidste fjerdedel af bedømmelserne enten er mundtlige eksaminationer eller skriftlige opgaver, hvor den studerendes navn står på opgaven. Hvis analyserne gennemføres alene med anonyme eksaminer, gør det ingen substantiel forskel på resultaterne. Vi har information om, hvorvidt der er ekstern censur, og hvilket køn den eksterne censor har, men inddragelse af denne variabel viser sig hverken at have betydning for karaktergivningen eller for sammenhængen mellem kønnet på den interne bedømmer og karaktergivningen. Den indgår derfor ikke i de viste resultater.

De to afhængige variabler er hhv. en given bedømmers karaktergennemsnit i et givent fag i en given eksamenstermin og den tilsvarende beståelsesprocent. Den officielle konvertering er anvendt til at kombinere eksaminer før og efter ændringen af karakterskalaen i september 2007, men stadigvæk inddrager vi en variabel for, om karaktergivningen fandt sted under anvendelse af den gamle eller den nye skala. Vi inddrager også separate tidstrends for perioden før og efter ændringen i karakterskalaen. For karaktergennemsnit inddrages kun observationer, hvor mindst 5 studerende blev bedømt for at undgå, at ekstreme observationer skævvrider analyserne. Ud over bedømmernes køn inddrager vi også deres alder (og for analysen med bedømmere som enheder også antallet af år siden kandidatgraden), da yngre/mindre erfarne bedømmere kan tænkes at have forskellig bedømmelsesadfærd og en forskellig

grad af studenterorientering. For at måle studenterorienteringen stillede vi fire spørgsmål (se tabel 1), der handlede om, hvorvidt underviseren var motiveret for at bidrage til, at den studerendes ønsker blev tilgodeset. Svarene blev afgivet på en fem-punkts Likertskala. En faktoranalyse af spørgsmålene (tabel 1) tyder på, at spørgsmålene dækker over én bagvedliggende dimension. Statistisk sammenhængsanalyse viser også, at sammenhængen mellem de anvendte spørgsmål er stor.

Ordlyd i spørgeskemaet	Loadings i principal component analysis
Hensynet til den enkelte studerende er for mig altid vigtigere end hensynet til de formelle regler	0,702
Det giver mig energi at vide, at jeg har medvirket til, at den enkelte studerende lærer noget	0,579
Det er vigtigt at sætte den enkelte studerende i centrum	0,797
De studerendes tilfredshed er meget vigtig for, om jeg anser opgaven for at være løst på en god måde	0,605

Tabel 1: Items til at måle studenterorientering Note: Cronbachs alfa: 0,60

De spørgsmål, vi bruger til at måle de institutionelle variable med, findes i det tidligere omtalte appendiks. Disse spørgsmål blev gentaget i skemaet, således at vi fik en separat måling for hvert af de fag, bedømmerne havde vurderet eksaminer inden for.

Analyserne er random effects panel regressioner og (for beståelsesandel) tobit panel regressioner. En tobit regression er en statistisk metode, der tager højde for, at der kun er to mulige udfald af beståelsesbeslutningen: bestået og ikke-bestået. Da langt de fleste studerende består, har de fleste bedømmere en beståelsesprocent omkring 100, og tobit regression korrigerer for, at det hverken er muligt at have mere end 100 % beståede eksaminer eller mindre end 0 % beståede eksaminer (se Cameron & Trivedi, 2009 s. 521-524 for en diskussion af tobit regression).

Resultater

I dette afsnit præsenterer vi først resultaterne fra analyserne af sammenhængen mellem bedømmernes køn og deres studenterorientering (hvor bedømmerne som individer er analyseenheder) og derefter analyserne af hhv. beståelsesprocent og karaktergennemsnit.

Tabel 2 viser, at der er de forventede kønsforskelle i studenterorientering. De signifikante og positive regressionskoefficienter for 'kvindelige bedømmere' betyder, at kvinderne er systematisk mere studenterorienterede sammenlignet med deres mandlige kolleger (som forventet i hypotese 1). Kvinderne scorer næsten 9 point højere

end mændene, hvilket på en skala fra 0 til 100 må siges at være en substantiel forskel. Denne sammenhæng er gældende, selvom der kontrolleres for betydning af institutionelle normer og kollegainstitutioner.

	Model 2-1	Model 2-2	Model 2-3
Kvindelige bedømmere	8,805*	8,996*	8,949*
	(0,032)	(0,047)	(0,049)
Universitetsinstitut A		-3,815	-4,088
		(0,396)	(0,367)
Alder		0,086	0,658
		(0,593)	(0,399)
År siden kandidatgrad			-0,586
			(0,453)
Konstant	68,556	65,320	49,910
	(0,000)	(0,000)	(0,025)
N	78	75	75
R ²	0,059	0,058	0,072
Adj. R ²	0,047	0,018	0,013

Tabel 2: OLS regressioner af studenterorientering. (skaleret fra 0-100). Ustandardiserede regressionskoefficienter. 2010. Note: * $p < 0.05$. Analyseenhederne er universitetsbedømmere fra to statskundskabsinstitutter.

I eksamensbedømmelser handler den mest fundamentale beslutning om, hvorvidt en given studerende skal bestå eller dumpe, og tabel 3 belyser dette ved en række tobit regressioner af beståelsesandelene for en given eksamen for en given bedømmer i et givent semester. Tabellen viser, at regressionskoefficienterne for 'kvindelig bedømmer' ikke er signifikant forskellige fra nul, hvilket peger på, at der ingen systematiske kønsforskelle er i beståelsesprocenterne. Det kunne skyldes begrænsninger i antallet af observationer, men eftersom regressionskoefficienterne også er ubetydeligt små (svarende til en kønsforskel på få procent i beståelsesandel) taler resultaterne for at afvise forventningen om en sammenhæng mellem køn og beståelsesprocent. Modsat forventningen i hypotese 2a lader kvinderne altså ikke systematisk deres studerende bestå i større omfang end de mandlige bedømmere. Derimod lader alder til at have en systematisk negativ effekt. De negative regressionskoefficienter for alder i model 3-1, 3-2 og 3-3 betyder, at ældre bedømmere har større tendens til at dumpe studerende end yngre undervisere. Det skyldes næppe, at ældre undervisere underviser dårligere end yngre, eller at ældre undervisere systematisk har studerende, der laver mindre. Som hovedregel er der mange bedømmere og mere end én underviser per fag på bachelorårgangene. Et bud er snarere, at bedømmere med alderen bliver mere sikre på deres bedømmelse og derfor i højere grad tør tage den ubehagelige beslutning at dumpe studerende. Det støttes af, at alder ikke har nogen direkte effekt, når niveauet af generel institutionalisering inddrages i model 3-4. Denne variabel måler netop, hvor sikker individet føler sig på sin karaktergivning. Hvis en bedømmer oplever en høj grad af generel institutionalisering i denne forstand, tenderer vedkommende mod at have en lav beståelsesprocent. Med andre ord udligner faste normer

for karaktergivning på et fag betydningen af alder. På den måde kan institutionalisering være med til at sikre en ensartet bedømmelse på tværs, som vi kommer nærmere ind på i diskussionsafsnittet.

Den positive sammenhæng mellem studenterorientering og beståelsesprocent i tabel 3 er som forventet i hypotese 3a. Det vil sige, at undervisere, der synes, at det er vigtigt at sætte den studerendes ønsker i centrum, er mere tilbøjelige til at lade deres studerende bestå eksamen.

	Model 3-1	Model 3-2	Model 3-3	Model 3-4
Alder	-0,005*** (0,00)	0,002* (0,03)	-0,002* (0,03)	-0,001 (0,41)
Kvindelig bedømmer	-0,038 (0,23)	0,014 (0,62)	-0,004 (0,90)	-0,000 (1,00)
År (2008=0)	0,015** (0,00)	-0,029** (0,01)	-0,029** (0,01)	-0,024* (0,03)
År*ny karakterskala		0,017 (0,34)	0,017 (0,35)	0,009 (0,62)
Ny karakterskala		0,136*** (0,00)	0,141*** (0,00)	0,125*** (0,00)
Universitetsinstitut A		-0,149*** (0,00)	-0,143*** (0,00)	-0,133*** (0,00)
Ekstern bedømmer		-0,038 (0,07)	-0,038 (0,07)	-0,028 (0,16)
Studenterorientering			0,001* (0,04)	0,002* (0,02)
Kollegainstitutioner				0,000
General				-0,002* (0,03)
institutionalisering				
Konstant	1,267** * (0,00)	1,098*** (0,00)	0,996*** (0,00)	1,046*** (0,00)
N	381	380	380	362
sigma_e	0,121	0,120	0,120	0,116
sigma_u	0,112	0,077	0,074	0,067
rho	0,460	0,293	0,272	0,251

Tabel 3: Tobit panel regressions af beståelsesprocenter. (skaleret fra 0-100). Ustandardiserede regressionskoefficienter (p-værdier i parentes). 2004-2010. Note: Analyseenheden er en eksamination af et givent fag i et givent semester udført af en given bedømmer * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Referencekategorien er mandlig bedømmer og ingen ekstern censor.

Tabel 4 viser random effect regressioner af karaktergennemsnittene i de undersøgte eksaminer (en observation er gennemsnittet for en given bedømmer i et givent semester). Tabel 4 tyder på, at studenterorientering fører til højere karaktergennemsnit (som forventet i hypotese 3b), og det er også tilfældet, hvis vi alene beregner gennemsnittet af de beståede eksaminer (denne analyse er ikke vist). Derimod finder vi ingen direkte kønsforskel på karaktergennemsnittene, idet regressionskoefficienterne for 'kvindelig bedømmer' både er små og statistisk insignifikante. Regressionskoefficienter på under 0,2 peger således på, at en eventuel forskel mellem mænd og kvin-

der er under en femtedel karakterpoint. Derudover er koefficienterne i tre ud af fire modeller i tabel 4 negative (hvilket betyder, at kvinder om noget skulle have et lavere karaktergennemsnit end mænd). Det taler for en afvisning af hypotese 2b.

	Model 4-1	Model 4-2	Model 4-3	Model 4-4
Alder	-0,017*	-0,006	-0,005	-0,001
	(0,03)	(0,48)	(0,52)	(0,87)
Kvindelig bedømmer	-0,164	0,045	-0,134	-0,186
	(0,43)	(0,83)	(0,53)	(0,38)
År (2008=0)	0,201***	-0,073	-0,077	-0,107
	(0,00)	(0,25)	(0,22)	(0,10)
År*ny karakterskala		0,129	0,133	0,156
		(0,18)	(0,16)	(0,10)
Ny karakterskala		0,879***	0,888***	0,947***
		(0,00)	(0,00)	(0,00)
Universitetsinstitut A		-0,725***	-0,660***	-0,545**
		(0,00)	(0,00)	(0,01)
Ekstern bedømmer		-0,236	-0,236*	-0,273*
		(0,05)	(0,05)	(0,02)
Studenterorientering			0,015**	0,013*
			(0,01)	(0,02)
Kollegainstitutioner				0,001
				(0,88)
Generel institutionalisering				-0,005
				(0,34)
Konstant	7,127***	6,263***	5,220***	5,464***
	(0,00)	(0,00)	(0,00)	(0,00)
R ² overall	0,120	0,239	0,249	0,259
R ² between	0,150	0,256	0,289	0,278
R ² within	0,149	0,195	0,195	0,201
N	367	366	366	350
sigma_e	0,732	0,713	0,713	0,715
sigma_u	0,947	0,865	0,839	0,767
rho	0,626	0,595	0,581	0,535

Tabel 4: Random effects regressioner af bedømmernes gennemsnit af karakterer. standardiserede regressionskoefficienter (*p*-værdier i parentes). 2004-2010. Note: Analyseenheden er en eksamination af et givent fag i et givent semester udført af en given bedømmer * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Referencekategorien er mandlig bedømmer og ingen ekstern censor.

Sammenfattende har vi fundet, at kvindelige universitetsbedømmere er mere studenterorienterede, men at de ikke giver bedre karakterer, selvom bedømmere med høj grad af studenterorientering giver bedre karakterer. Kvindelige bedømmere er (om ikke blødere overfor den enkelte studerende, så i hvert fald) mere orienteret mod at bidrage til at opfylde den enkelte studerendes ønsker. Kvindelige bedømmere giver imidlertid ikke systematisk anderledes karakterer end mændene. Hvordan den uoverensstemmelse skal forstås, diskuteres i næste afsnit.

Diskussion og konklusion

Der kan være flere grunde til, at kvinder tilsyneladende ikke som forventet giver bedre karakterer, selvom de faktisk lader til at være mere studenterorienterede. Når denne orientering ikke lader til at føre til en bruttokønsforskel, kan det som nævnt i teori afsnittet skyldes, at kvinderne selv modvirker kønsstereotyper (Sabbe & Aelterman, 2007, s. 530). Hvis kvinderne søger at undgå at blive fanget i en feminin rolle, kan det gøre, at de afstår fra handlinger, som får dem til at fremstå som bløde over for kolleger og studerende i karaktergivningssituationen (Kanter, 1977). Med henblik på at undgå følelsesmæssig motivation såvel som at fremstå som værende følelsesmæssigt motiveret (Meier, Mastracci & Wilson, 2006, s. 900) er det muligt, at nogle kvindelige bedømmere er hårdere i deres bedømmelse end mænd med tilsvarende studenterorientering. Denne fortolkning understøttes af en kønsforskel i universitetslærernes beskrivelser af deres følelser ved at give lave karakterer. Mens hele 25,4 % af mændene erklærede sig helt eller delvist enige i, at de var ubehageligt til mode, når de gav en lav karakter, gjorde det sig kun gældende for 11,1 % af de kvindelige bedømmere. Ud over resultaterne vedrørende køn finder vi også, at alder har betydning for karaktergivningen. Ældre bedømmere er således mindre i tvivl, når de bedømmer, og tager i højere grad den ubehagelige beslutning om at dumpe en studerende. Det viser relevansen af at kigge på personlige karakteristika, men det illustrerer også, at der stadig er behov for at forske videre inden for dette område.

Et af undersøgelsens vigtige resultater er, at studenterorientering hænger signifikant sammen med karaktergivning, således at bedømmere med høj grad af studenterorientering lader en større andel af de studerende bestå og har et højere karaktergennemsnit. Det er problematisk af den årsag, at de studerende bliver forskelsbehandlet alene ud fra deres bedømmers studenterorientering. Vi vil imidlertid argumentere for, at det næppe er hensigtsmæssigt at arbejde for, at universitetsansatte generelt får en lavere grad af studenterorientering. Denne orientering kunne sagtens tænkes at hænge positivt sammen med graden af studentercentrering i undervisningssituationer. Studentercentrering handler jf. Trigwell, Prosser & Waterhouse (1999, s. 58) om, at underviserens handlinger i undervisningen er rettet mod at ændre/udvikle de studerendes forståelse. Denne tilgang, hvor den studerende er aktiviteternes omdrejningspunkt, har en positiv effekt på de studerendes læring (ibid., s. 66). Derfor kunne en normativ implikation være at søge at forøge de mindre studenterorienterede bedømmers studenterorientering således, at vi af ad den vej fik en ens bedømmelsespraksis uden at give afkald på de positive effekter knyttet til studenterorientering. En anden måde at sikre ensartet bedømmelsespraksis på kunne være også at styrke de faglige normer for karaktergivning yderligere. Den i artiklen beskrevne undersøgelse tyder på, at de faglige normer tilsyneladende er stærke nok til at modvirke en kønsforskel, mens der stadig er en systematisk sammenhæng mellem karaktergivning og

studerorientering. De faglige normer kunne bl.a. styrkes ved øget faglige diskussion i forbindelse med for- og eftercensur.

Studiet her har en række begrænsninger, som fremtidige undersøgelser vil kunne komme ud over. Undersøgelsen er foretaget på statskundskabsstudier og kan ikke uden videre overføres til andre kontekster. Ikke desto mindre er resultaterne interessante og kan efterforskes i andre sammenhænge. Mens karaktererne er indsamlet over en årrække, er der kun indsamlet data om underviserens studenterorientering én gang. For at styrke analysens resultater kunne man med fordel indsamle underviserens svar over tid. Med forbehold for disse begrænsninger er det centrale resultat i artiklen, at selvom studenterorienterede bedømmere giver bedre karakterer, og selvom kvinder er mere studenterorienterede end mænd, er der ikke systematisk forskel i bedømmelserne mellem mandlige og kvindelige bedømmere.

Lotte Bøgh Andersen, ph.d. og professor ved Institut for Statskundskab, Aarhus Universitet samt KO-RA, Det Nationale Institut for Kommuners og Regioners Analyse og Forskning. Forsker i offentlige ledere og medarbejdere med særligt fokus på ledelse, motivation, performance og køn.

Heidi Houlberg Salomonsen, ph.d. og lektor ved Institut for Statskundskab, Aalborg Universitet. Forsker i offentlig forvaltning og organisering med særligt fokus på relationen mellem politikere og embedsmænd, offentlig ledelse og kommunikation.

Litteratur

- Andersen, L.B. & L.H. Pedersen (2013). Does ownership matter for employee motivation when occupation is controlled for? *International Journal of Public Administration*, 36(12), s. 840-856.
- Andersen, L. B., Pallesen, T. & Salomonsen, H. (2013). Doing good for others and/or for society? The relationships between public service motivation, user orientation and university grading. *Scandinavian Journal of Public Administration*. 17(3), s. 23-44.
- Andersen, L. B. & Heinesen, E. (2013). *Gender and examinations marks: Why do girls excel in language and boys in Science?* Paper presented at Workshop on Economics of Successful Children 10-13. Januar 2013, Aarhus, Danmark.
- Andersen, L. B., Pallesen, T. & Pedersen, L. H. (2011). Does Ownership Matter? Public Service Motivation among Physiotherapists in the Private and Public Sectors in Denmark. *Review of Public Personnel Administration*, 31, s. 10-27.
- Anderson, A. (2013). "I Was That Student": Exploring Instructor Experience of Empathy Amid Identities in the Classroom
<http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/39601/AndersonAnna2013.pdf?sequence=1> (tilgået 25. januar 2014).
- Barber, T. (2002). A special duty of care: exploring the narration and experience of teacher caring. *British Journal of Sociology of Education*, 23, s. 383-395.
- Baron-Cohen, S. (2003). *The Essential Difference: The Truth about the Male and Female*

- Brain*. New York: Perseus Books Group.
- Cameron, A. C. & Trivedi, P. K. (2009). *Microeconometrics Using Stata*. Texas: Stata Press.
- Cornwell, R. C. (1974). The Student, the Teacher, and Empathy. *The Journal of Business Education*, 50(3), s. 123-124.
- Crombie, G., Pyke, S. W., Silverthorn, N., Jones, A. & Piccinin, S. (2003). Students' Perceptions of Their Classroom Participation and Instructor as a Function of Gender and Context. *Journal of Higher Education*, 74(1), s. 51-76.
- Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A. & Ahluwalia, J. (2000). Sex Differences in Human Neonatal Social Perception. *Infant Behaviour and Development*, 23, s. 113-118.
- Dee, T. S. (2005). A Teacher Like Me: Does Race, Ethnicity or Gender Matter? *American Economic Review*, 95(2), s. 158-165.
- Duncan-Andrade, J. M. R. (2011). The principal facts: New directions for teacher education. I: A. F. Ball & C. A. Tyson (red.). *Studying Diversity in Teacher Education*. Lanham, MD, Rowman & Littlefield Publishers, Inc. (tilgæet på http://iseeed.org/wp-content/uploads/2012/12/JDA_The-Principal-Facts.pdf 25. januar 2014).
- Gair, S. (2011). Creating spaces for critical reflection in social work education: learning from a classroom-based empathy project. *Reflective Practice: International and Multidisciplinary Perspectives*, 12(6), s. 791-802.
- Goodwin, L. D. & Stevens, E. A. (1993). The influence of Gender on University Faculty Members' Perception of "Good" Teaching. *Journal of Higher Education* 64(2), s. 166-185.
- Hale, M. (1999). He Says, She Says: Gender and Worklife. *Public Administration Review*, 59, s. 410-424.
- Hoffman, F. & Oreopoulos, P. (2007). A professor like me: The influence of instructor gender on college achievement. *NBER Working Paper 13182* <http://www.nber.org/papers/w13182> (June 22 2011).
- Jensen, H. N. (2010). Det lukkede rum - en dør på klem til specialevejledning. *Dansk Universitetspædagogisk Tidsskrift*, 8, s. 17-22.
- Jensen, U. T. & Andersen, L. B. (2013). *When Doing Good for the Individual Can Harm Society and Vice Versa: Trade-offs Between Public Service Motivation and User Orientation*. Konferencepapir præsenteret på International research conference: One step beyond – refining public service motivation theory and research methods, Utrecht, Holland 14-16. november 2013.
- Kalaian, H. A. & Freeman, D. J. (1994). Gender differences in self-confidence and educational beliefs among secondary teacher candidates. *Teaching and Teacher Education*, 10, s. 647-658.
- Kanter, R. M. (1977). Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women. *The American Journal of Sociology*, 82, s. 965-990.

- Meier, K., Mastracci, S. H. & Wilson, K. (2006). Gender and Emotional Labor in Public Organizations: An Empirical Examination of the Link to Performance. *Public Administration Review*, 66, s. 899-909.
- Nielsen, V. L., Andersen, L. B. & Pedersen, M. B. (2012). Specialet som glasloft: Hvorfor overhaler de mandlige studerende kvinderne ved specialet?" *Dansk Universitetspædagogisk Tidsskrift* 7(12), s. 58-69.
- Nielsen, V. L. (2013). Gender and differences in personal attributes: Myth or just irrelevant to job behavior because of selection and self-selection mechanisms! *Department of Political Science, Aarhus University, Denmark.*
- Nielsen, V. L. & Salomonsen, H. H. (2010). Mænd og kvinders adfærd i den offentlige sektor. *Økonomi og Politik*, 1, s. 65-79.
- Olesen, J., Aggerholm, K. & Kofoed, J. (2008). Flere end to slags børn – en rapport om køn og ligestilling i børnehaven. København: Ministeriet for Ligestilling.
- Rhodes, T. (2010). *Assessing Outcomes and Improving Achievement: Tips and Tools for Using Rubrics*. Washington, DC: Association of American Colleges and Universities. Tilgået 25. januar 2014 på <http://www.aacu.org/value/rubrics/InterculturalKnowledge.cfm>.
- Sabbe, E. & Aelterman, A. (2007). Gender in teaching: a literature review. *Teachers and Teaching*, 13, s. 521-538.
- Scott, R. S. (2001). *Institutions and Organizations*, 2nd ed. Thousand Oaks, London, New York: Sage.
- Trigwell, K., Prosser, M. & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education*, 37, s. 57-70.
- Yang, C. Y., Decety, J., Lee, S., Chen, G. & Cheng, Y. (2009). Gender differences in the Mu rhythm during empathy for pain: An electroencephalographic study. *Brain Research*, 1251, s. 176-184.

Appendiks

Variable	Definition	Gennemsnit	Std.	Observeret minimum	Observeret maksimum	n
Kvindelig bedømmer	Kønnet på den interne bedømmer (0= mand, 1=kvinde)	0,23	0,42	0	1	78
Alder	Bedømmerens alder, da spørgeskemaet blev udsendt	43,27	11,60	26	68	75
År siden kandidatgrad	Antal år siden bedømmerens kandidatgrad, da spørgeskemaet blev udsendt	15,81	11,56	1	42	75
Studenter- orientering	Motivation til at bidrage til at opfylde den enkelte studerendes ønsker (se tabel A1 for ordlyd i spørgeskemaspørgsmål). Teoretisk variation ml. 0 og 100	70,59	15,36	31,25	93,75	78
Kollega institutioner	Graden, hvormed formelle og uformelle regler blandt kolleger opfattes som regulerende for adfærden. (se appendiks for ordlyd i spørgeskemaspørgsmål).	52,97	18,88	0,00	83,33	74
General institutionalisering	Underforståede regler for bedømmelser. (se appendiks for ordlyd i spørgeskemaspørgsmål).	72,61	14,73	33,33	100,00	74

Tabel A1: Variable i datasættet med universitetsbedømmere som analyseenheder

Variable	Definition	Gen-nemsnit	Std.	Observeret minimum	Observeret maksimum	n
Karaktergennemsnit	Sum af karakterer givet af en bedømmer i en given eksamenstermin i et givent fag divideret med antal studerende	6,149	1,28	1,8	10,21	473
Beståelsesandel	Andel studerende, der bestod en given eksamen ud af alle studerende, der pågældende bedømmer havde oppe til eksamen i det pågældende fag	0,941	0,11	0,286	1	493
Kalenderår	Året, hvori bedømmelsen fandt sted (2004,5 er fx sommeren 2004)	2007,4	1,96	2004	2010,5	494
År (2008=0)	År siden ny skala trådte i kraft	-0,618	1,96	-4	2,5	494
Ny karakterskala	Er karakteren givet med ny skala (alle karakterer er omregnet til ny skala via officiel omregningsfaktor)	0,498	0,50	0	1	494
Institut	Aarhus =0, Aalborg=1	0,331	0,47	0	1	496
Ekstern censor	Var der ekstern censur (0=nej, 1=ja)	0,488	0,50	0	1	496
Alder	Alder på bedømmer i det givne semester	43,41	11,47	25	68	381
Kvindelig bedømmer	Kønnet på den interne bedømmer (0= mand, 1=kvinde)	0,203	0,40	0	1	497
Studenter- orientering	Motivation til at bidrage til at opfylde den enkelte studerendes ønsker (se tabel A1 for ordlyd i spørgeskema-spørgsmål). Teoretisk variation ml. 0 og 100	69,31	15,21	31,25	93,75	410
Kollega-institutioner	Graden, hvormed formelle og uformelle regler blandt kolleger opfattes som regulerende for adfærden. (se appendiks for ordlyd i spørgeskemaspørgs-	54,60	16,83	0	83,33	367

	mål).						
General institutionalisering	Underforståede regler for bedømmelser. (se appendiks for ordlyd i spørgeskemaspørgsmål).	75,45	17,17	33,33	100	372	

Tabel A2: Variable fra datasættet med eksaminationer som analyseenheder 2004-2010

Ordlyd	Loadings i principal component analysis
Jeg er sjældent i tvivl om, hvilken karakter jeg skal give en besvarelse	0,696
Jeg er sjældent uenig med censor om, hvilken karakter vi skal give en besvarelse	0,818
Der er stor enighed blandt mine kollegaer om, hvad der kendetegner en god besvarelse	0,842

Tabel A3: Måling af general institutionalisering. Note: Cronbachs alfa 0,68. De individuelle universitetsbedømmere er analyseenheder.

	Full Danish version	Antal point i formative indeks
	Der er forcensur, hvor bedømmerne har læst den eller de samme besvarelser	
Formelle ex ante kollegainstitutioner	Der er forcensur, hvor bedømmerne diskuterer besvarelserne uden at have læst de samme besvarelser	3
	Der er en anden form for forcensur	2
	Ingen forcensur	1
	Der er eftercensur i form af et møde (evt. efter forudgående e-mail korrespondance)	4
Formelle ex post kollegainstitutioner	Der er eftercensur alene i form af e-mail korrespondance om bedømmelserne (fx karakterniveau, andel af dumpede, spredning mv.)	3
	Der er en anden form for eftercensur	2
	Der er ikke eftercensur	1
Grad af uformel kollegadiskussion	Jeg diskuterer løbende besvarelserne med undervisere fra kurset i forbindelse med min bedømmelse af besvarelserne	1-5 (5= max. Enighed)

Tabel A4: Items til måling af kollegainstitutioner. Ordlyd

Ordlyd i spørgeskemaet	Loadings i principal component analysis
Hensynet til den enkelte studerende er for mig altid vigtigere end hensynet til de formelle regler	0,702
Det giver mig energi at vide, at jeg har medvirket til, at den enkelte studerende lærer noget	0,579
Det er vigtigt at sætte den enkelte studerende i centrum	0,797
De studerendes tilfredshed er meget vigtig for, om jeg anser opgaven for at være løst på en god måde	0,605

Tabel A5: Items til at måle studenterorientering. Note: Cronbachs alfa: 0,60.

	Gennemsnit	Beståelsesandel	År (2008=0)	År*Ny skala	Age	Kvindelig bedømmer	Institut	Ekstern censor	Studerorientering	Koll. Institutioner	Generel inst.
Gennemsnit af karakterer	1,000										
Beståelsesandel	0,6428	1,0000									
År (2008=0)	0,3293	0,1191	1,0000								
År*Ny skala	0,3089	0,1213	0,8059	1,0000							
Age	-0,1270	-0,2705	-0,1107	-0,0963	1,0000						
Kvindelig bedømmer	0,0174	0,0420	0,0702	0,0167	-0,2966	1,0000					
Institut	-0,2887	-0,4504	-0,1845	-0,1621	0,2926	0,1067	1,0000				
Ekstern censor	-0,0794	-0,0846	-0,0234	0,0648	0,1386	-0,1548	-0,1666	1,0000			
Studerorientering	0,1489	0,1950	0,0253	0,0155	-0,1780	0,3240	-0,0330	-0,1122	1,0000		
Koll. Institutioner	0,0924	0,1454	0,0667	0,0340	-0,2197	0,2417	-0,2504	0,0510	-0,0105	1,0000	
Gen. inst.	-0,2230	-0,2865	-0,1991	-0,1807	0,4835	-0,0974	0,3135	0,0224	-0,0925	-0,0336	1,0000
Ny skala	0,3823	0,1664	0,8594	0,6477	-0,0653	0,0557	-0,1316	0,1136	0,0310	0,0814	-0,1807

Tabel A6: Korrelationer mellem undersøgte variable. Pearson's r.