Evalueringen af 7-trins-skalaen

Samlet rapport

DANMARKS
EVALUERINGSINSTITUT

INDHOLD

Evalueringen af 7-trins-skalaen

1	**Resumé**	**5**
1.1	Baggrund	5
1.2	Formål med evalueringen	8
1.3	Hovedresultater	8
1.4	Datagrundlag	22
1.5	Læsevejledning	24
2	**Udvikling i karaktergivning**	**25**
2.1	Baggrund og undersøgelsesspørgsmål	25
2.2	Grundskoleområdet følger overordnet den forventede fordeling	26
2.3	De gymnasiale uddannelser følger overordnet den forventede fordeling	27
2.4	Det videregående uddannelsesområde følger ikke den forventede fordeling	29
2.5	Perspektiver på stigning i andelen af karakteren 12	33
3	**Forståelse og anvendelse af skalaen**	**39**
3.1	Baggrund og undersøgelsesspørgsmål	39
3.2	Forståelsen og anvendelsen af skalaen	40
3.3	Målopfyldelse og bedømmelsesværktøjer	47
3.4	Kvalitetssikringen af karaktergivningen (efter karaktergivningen)	57
4	**Karakterskalaens udformning**	**67**
4.1	Baggrund og undersøgelsesspørgsmål	67
4.2	Perspektiver på springene på karakterskala	68
4.3	Brug af og perspektiver på -3 og 00	75

5 Belønning og identifikation af særligt dygtige studerende 78
5.1 Baggrund og undersøgelsesspørgsmål 78
5.2 Udbredelse af 12-tallet på de videregående uddannelser 79
5.3 Perspektiver på muligheder for identifikation og belønning af særligt dygtige studerende 80
5.4 Øvrige muligheder for identifikation af særligt dygtige studerende 83

6 Skalaens internationale anvendelighed 84
6.1 Baggrund – 7-trins-skalaens understøttelse af international anvendelighed 84
6.2 Er spørgsmålet om international anvendelighed stadig relevant? 86
6.3 7-trins-skalaens internationale anvendelighed 91
6.4 De nye anbefalinger i ECTS-systemet og fremtidige perspektiver 94

Appendiks A – Litteraturliste 96

Appendiks B – 7-trins-skalaen og ECTS-skalaen 97

1 Resumé

1.1 Baggrund

7-trins-skalaen blev indført på grundlag af Karakterkommissionens betænkning fra 2004, som indeholder kommissionens analyse af det daværende karaktersystem og anbefalinger om et fremtidigt system (betænkningen kan findes på: http://static.uvm.dk/publikationer/2004/karakterer/karakterer.pdf).

7-trins-skalaen blev taget i brug på de gymnasiale uddannelser i 2006, og i løbet af 2007 blev den indført i de øvrige dele af uddannelsessystemet. 7-trins-skalaen afløste den hidtidige karakterskala, 13-skalaen, som havde været i brug i grundskolen og på de gymnasiale uddannelser siden 1963 og på de videregående uddannelser siden 1971. Den blev gennem tiden anvendt på forskellig måde. Både beskrivelsen af karaktererne og måden at give karakterer på har fx varieret på tværs af uddannelsesområderne, og skalaen blev inden for de enkelte områder justeret og tilpasset flere gange.

Karakterkommissionen pegede på, at en karakterskala efter lang tids brug bliver slidt, og kommissionens analyser viste flere udfordringer med hensyn til 13-skalaen:

- At 13-skalaen ikke var kompatibel med den internationale ECTS-skala (se beskrivelse af ECTS-skalaen i appendiks B) og den nordamerikanske skala, og at den var svær at anvende internationalt, især på grund af undtagelseskarakteren 13.
- At beskrivelserne af skalaens karakterer lagde op til karaktergivning, ud fra hvordan de studerende var rangeret i forhold til hinanden (fx middelkarakter), selvom karaktererne faktisk skulle gives ud fra opfyldelsen af læringsmål (absolut karaktergivning). Dette var særligt en udfordring ved karaktererne 7, 8 og 9, som derfor kunne være svære at adskille.
- At der havde været inflation og uforklarlige forskelle i karaktergivningen på tværs af fag og hovedområder, og at der ikke var en ensartet forståelse af, hvornår de forskellige karakterer skulle bruges.

Karakterkommissionen opstillede fem krav til en moderne skala (se boksen på næste side). Kommissionen vurderede, at en moderne skala skulle være internationalt anvendelig, skulle anvendes ensartet i hele uddannelsessystemet, skulle understøtte, at bedømmelserne udtrykker målopfyldelse, skulle kunne bruges til gennemsnitsberegning og skulle have en klar differentiering mellem karaktertrinene. Kommissionen konstaterede, at 13-skalaen ikke opfyldte de fem krav, og at der derfor var brug for en ny karakterskala.

Karakterkommissionens fem krav til en ny karakterskala

1. International anvendelighed

Den øgede internationalisering stiller krav til karakterskalaen om, at karakterer og eksamensresultater ubesværet skal kunne benyttes internationalt. Internationaliseringen er ikke udelukkende et anliggende for de videregående uddannelser. Således fremgår det af OECD-rapporten *Education at a glance* (OECD, 2004), at internationalisering efterhånden må forventes at omfatte størstedelen af uddannelsesområderne. Det betyder, at karakterskalaen først og fremmest skal være kompatibel med ECTS-skalaen. Det skal således være muligt at oversætte mellem de to skalaer på enkeltkarakterniveau.

2. Skalaen skal anvendes ensartet i hele uddannelsessystemet, og bedømmelsen skal udtrykke graden af målopfyldelse

Overgangen fra stram indholdsstyring til målstyring stiller det krav, at karaktererne skal kunne udtrykke graden af målopfyldelse, altså at karaktererne skal kunne anvendes til absolutte vurderinger i forhold til præcist opstillede mål. Som en konsekvens heraf skal beskrivelsen af de enkelte karakterer relatere sig til målene og ikke til en relativ placering af eksamenspræstationen. Med denne form for karaktergivning sikres en ensartet brug af karakterskalaen, hvilket blandt andet er en forudsætning for at realisere intentionen om at styrke samspillet mellem fagene.

3. Samme skala i hele uddannelsessystemet

Kommissionen finder, at der skal benyttes samme karakterskala overalt i uddannelsessystemet, således at mobiliteten såvel vertikalt som horisontalt mellem uddannelserne understøttes. Derudover fremmer anvendelse af samme skala på de forskellige trin i uddannelsessystemet borgernes kendskab til og accept af karakterskalaen.

4. Klar trindifferentiering

Antallet af trin og springene mellem trinene i karakterskalaen skal være afpasset således, at valget mellem en karakter og nabokarakteren er udtryk for en sikker differentiering. Samtidig skal skalaen være så finkornet, at den kan benyttes i forbindelse med den formative vurdering.

5. Mulighed for gennemsnitsberegning

Anvendelse af karaktergennemsnit er i dag udbredt i dele af det danske uddannelsessystem. Kommissionen finder derfor, at det må stilles som et krav til en dansk karakterskala, at det skal være muligt at beregne gennemsnit. Den enkleste måde at opfylde dette krav på er at benytte talværdier frem for fx bogstaver.

Kommissionen opbyggede derfor en ny skala (7-trins-skalaen), som skulle anvendes på samme måde i hele uddannelsessystemet. Opbygningen af 7-trins-skalaen har taget udgangspunkt i de fem krav på følgende vis:

- Karaktererne skal gives ud fra opfyldelsen af eksplicit formulerede læringsmål (absolut karaktergivning).
- Skalaens karakterer er fastsat, så de følger strukturen i ECTS-skalaen med den hensigt, at de danske karakterer skal kunne oversættes en-til-en til ECTS-skalaen.

- Skalaen har ikke nogen undtagelseskarakter til den ekstraordinært gode præstation, da det er svært at kommunikere til udlandet.
- Skalaens karakterer er tal, som der kan beregnes et gennemsnit af.
- Skalaen har tre brede karakterer, 4, 7 og 10, som skal opsamle hhv. ca. 25 %, ca. 30 % og ca. 25 % af de bestående karakterer og er forskellige nok til at understøtte, at bedømmerne har let ved at vælge, hvilken karakter der er den rigtige.
- Skalaen har ligesom ECTS-skalaen to ikke-bestående karakterer, 00 og -3. Talværdierne er fastsat, så afstanden mellem de ikke-bestående karaktererne er den samme som afstanden mellem de laveste bestående karakterer. Karakterkommissionen fandt, at det ville understøtte, at ca. samme andel ikke bestod deres uddannelse samlet set, sådan som det var tilfældet ved 13-skalaen.
- Karakterkommissionen beskrev også, at der skulle ske en kvalitetssikring af karaktergivningen ved at koordinere målene for faget og prøvens tilrettelæggelse i forhold til den population, der skal bestå prøven. Det skulle understøtte, at de bestående karakterer over tid og for store populationer nogenlunde følger en fast fordeling (se figur 2.1), og at skalaen bruges ens i hele uddannelsessystemet. Dette relative element i karakterskalaen fremgår dog ikke af den juridiske beskrivelse af karakterskalaen i karakterbekendtgørelsen.

Kommissionen valgte desuden, at skalaen skulle følge den danske tradition med talkarakterer, som ikke udtrykker opnåede point eller lignende, og at tallene skulle ligge nogenlunde i samme leje som i 13-skalaen.

Selvom 7-trins-skalaen er implementeret ens på hhv. Undervisningsministeriets og Uddannelses- og Forskningsministeriets område, er der stor forskel på, hvordan der arbejdes med understøttelse og kvalitetssikring af karaktergivningen inden for de to uddannelsesområder.

På grundskoleområdet og på de gymnasiale uddannelser fastlægges de faglige mål for de enkelte fag centralt i Undervisningsministeriet. Ministeriet udarbejder desuden konkrete vurderings-/bedømmelseskriterier for de enkelte prøver, og de skriftlige prøver afholdes som nationale prøver. Derved fastlægges et fælles fagligt niveau centralt i Undervisningsministeriet. Samtidig findes der national statistik for karaktergivning.

På erhvervsuddannelserne fastlægges de faglige mål centralt i læreplaner, uddannelsesbekendtgørelser og uddannelsesordninger. Bedømmelseskriterierne kan være fastlagt både centralt og decentralt. Eksempelvis beskrives bedømmelseskriterier for grundforløbs- og svendeprøver decentralt i et samarbejde mellem de faglige udvalg og skolerne. Uddannelsernes slutmål er centralt beskrevet i uddannelsesbekendtgørelsen for den enkelte erhvervsuddannelse. Målene for grundfagene fastlægges af Undervisningsministeriet, mens de uddannelsesspecifikke mål beskrives i et samarbejde mellem Undervisningsministeriet og uddannelsernes faglige udvalg.

På de videregående uddannelser fastlægges læringsmål for erhvervsakademi- og professionsbacheloruddannelserne og de maritime uddannelser i vid udstrækning centralt i Uddannelses- og Forskningsministeriet via uddannelsesbekendtgørelserne, mens institutionerne, som udbyder de kunstneriske uddannelser, og universitetsuddannelserne, selv sammensætter uddannelsernes struktur og fastlægger læringsmålene. På alle uddannelser sker kvalitetsarbejdet dog decentralt, ligesom alle prøver tilrettelægges decentralt, dog med censur fra et nationalt censorkorps ved de fleste prøver.

1.2 Formål med evalueringen

I Karakterkommissionens betænkning blev det anbefalet, at 7-trins-skalaen blev evalueret hvert femte år for at følge udviklingen i anvendelsen af den nye karakterskala. Første evaluering af 7-trins-skalaen blev afsluttet i 2013 (EVA, 2013) og viste, at karakterskalaen var implementeret i de relevante dele af uddannelsessektoren. Evalueringen viste også, at karaktererne nogenlunde fulgte den forventede fordeling på alle uddannelsesområder, at karakterskalaen var internationalt anvendelig, og at bedømmerne havde skullet vænne sig til den nye skala, men at de oplevede, at den var anvendelig i bedømmelsessituationen.

Evalueringen pegede også på potentielle udfordringer med skalaen og dens anvendelse. Der var ikke fastlagt faglige læringsmål ved alle prøver. Nogle bedømmere påpegede, at karakteren -3 sender et unødigt negativt signal, og især med hensyn til karakteren 7 fremhævede flere bedømmere, at de gerne ville kunne give mere nuancerede bedømmelser. Evalueringen pegede også på, at der stadig var behov for at understøtte, at skalaen blev brugt og forstået korrekt af både bedømmere og aftagere.

Den første evaluering viste også, at der var udsving i fordelingen af karaktererne og forskelle på tværs af uddannelsesniveauer. Karakterfordelingen blev dog kun belyst på et overordnet niveau, og forskelle mellem fag, prøveformer, faglige uddannelsesområder mv. blev ikke afdækket.

Formålet med denne evaluering er at give et samlet billede af udviklingen i karaktergivningen siden implementeringen af 7-trins-skalaen og at belyse mulige problemstillinger vedrørende 7-trins-skalaens udformning og dens anvendelse på alle uddannelsesniveauer. Evalueringen har fokus på følgende overordnede evalueringsspørgsmål:

1. Hvordan har karaktergivningen udviklet sig siden implementeringen af 7-trins-skalaen, og har der været karakterinflation?
2. Anvendes skalaen, som den var tænkt af Karakterkommissionen, og er kommissionens krav til en moderne karakterskala fortsat aktuelle?
3. Hvad betyder skalaens karakterspring og karakteren -3 for karaktergivningen?
4. Er der tilstrækkelige muligheder for at identificere særligt dygtige elever og studerende og belønne den ekstraordinære præstation?
5. Er 7-trins-skalaen internationalt anvendelig?

1.3 Hovedresultater

1.3.1 Evalueringsspørgsmål 1: Hvordan har karaktergivningen udviklet sig siden implementeringen af 7-trins-skalaen, og har der været karakterinflation?

7-trins-skalaen er en absolut skala, hvor karaktergivningen alene skal foretages ud fra en vurdering af, i hvor høj grad eksamenspræstationen lever op til de mål, der er fastlagt for prøven. Alligevel beskrev Karakterkommissionen, at kvalitetssikringen af karaktergivningen skulle understøtte, at de beståede karakterer over tid og for store populationer følger karakterfordelingen fra ECTS-skalaen, dvs. at det forventes, at 10 % får karakteren 02, 25 % får karakteren 4, 30 % får karakteren 7, 25 % får karakteren 10 og 10 % får karakteren 12.

Karakterudviklingen følger på *grundskoleområdet* aggregeret set den forventede karakterfordeling, når man ser samlet på perioden 2007-16. Gennemsnitskarakteren har været stigende over tid, men det afspejler, at der i starten af perioden blev givet færre 12-taller end forventet. Andelen af 12-taller er steget fra ca. 5 % i 2007 til 11 % i 2016.

Karakterudviklingen følger på *de gymnasiale uddannelser* aggregeret set den forventede karakterfordeling, når man ser samlet på perioden 2007-15. På de gymnasiale uddannelser har andelen af 12-taller også været stigende. I 2007 blev der givet ca. 9 % 12-taller. Dette er i 2015 steget til ca. 13 %, hvilket er en smule over forventningen.

Da der kun er data for en toårig periode, har det ikke været muligt at analysere udviklingen i karaktergivningen på *erhvervsuddannelserne*.

Karaktergivningen på *de videregående uddannelser* følger ikke den forventede fordeling, når man ser på perioden 2007-16, idet der er for stor en andel af høje karakterer (10 og 12) og for lille en andel af lave karakterer (02 og 4) i forhold til det forventede. Andelen af 12-taller har ligget over de forventede 10 % siden indførelsen af skalaen. I 2007 blev der tildelt 13 % 12-taller, og dette er steget til 17 % i 2016. Karaktererne er særligt høje i forbindelse med de afsluttende opgaver, og her ses desuden store forskelle i karaktergivningen på tværs af uddannelsesniveauer og hovedområder.

Splitter vi op på sektorer, ser vi, at erhvervsakademiuddannelserne på et aggregeret plan følger den forventede fordeling. På professionsbacheloruddannelserne bliver der generelt givet for stor en andel af 12-taller og for lille en andel af 4-taller. På universitetsområdet er der stor forskel i karaktergivningen på tværs af bachelor- og kandidatuddannelser. På bacheloruddannelserne ligger karaktergennemsnittet inden for det forventede, om end andelen af 12-taller er for stor i forhold til det forventede niveau. På kandidatuddannelserne bliver der i særlig grad givet for mange 10- og 12-taller i forhold til det forventede, og særligt andelen af 12-taller er støt stigende over tid. På de kunstneriske uddannelser er karakterfordelingen præget af meget færre lave karakterer og mange flere høje karakterer, end den forventede fordeling tilsiger. Dog ser der ikke ud til at have været nævneværdig karakterinflation siden 2011/12 på de kunstneriske uddannelser som helhed. De maritime uddannelser følger den forventede karakterfordeling, samtidig med at der ikke ses en tendens til karakterinflation som på de øvrige videregående uddannelser.

Ændringer i elevsammensætningen og studenterpopulationen kan ikke forklare det stigende karakterniveau. Desuden er der forskelle i karaktergivningen på tværs af fag i grundskolen og på de gymnasiale uddannelser og på tværs af hovedområderne på de videregående uddannelser, som ikke kan forklares ud fra sammensætningen af eleverne/de studerende. Det kan være et tegn på, at skalaen ikke bruges ensartet.

En central årsag til, at karaktererne på de videregående uddannelser afviger fra den forventede fordeling og de uforklarede forskelle i karaktergivningen, kan være, at kvalitetssikringen ikke fungerer, som Karakterkommissionen beskrev den. Forventningen til de videregående uddannelsesinstitutioner er hverken udfoldet klart i kommissionens rapport eller i den efterfølgende implementering af skalaen. Det er derfor ikke nødvendigvis tydeligt for institutionerne, hvad der forventes af dem (jf. afsnit 3.5).

I grundskolen og på de gymnasiale uddannelser sker størstedelen af kvalitetssikringen centralt i Undervisningsministeriet, og både fag, læringsmål og prøver er nationale og understøttes med fælles vejledninger. Det kan have bidraget til at understøtte, at karaktergivningen i højere grad følger det forventede. Også her er der dog tegn på, at skalaen bruges forskelligt inden for forskellige områder.

På alle de tre overstående uddannelsesområder har der været et stigende karakterniveau gennem årene, og derfor er det vigtigt fremover at følge med i udviklingen.

Evalueringen kan ikke give et entydigt svar på, hvorfor der er så store forskelle på tværs af sektorerne på de videregående uddannelser. Det eneste område, hvor der er markante forskelle med hensyn til anvendelsen af skalaen, er i forbindelse med kvalitetssikring. Lederne på erhvervsakademi- og professionsbacheloruddannelser ser, i højere grad end lederne af universitetsuddannelser, kvalitetssikring af karaktergivning, herunder opfølgning på karakterstatistik for de afholdte prøver, som en del af deres arbejde. De inddrager også i højere grad end lederne af universitetsuddannelser statistik over den samlede karakterfordeling på uddannelsen i en given eksamenstermin i kvalitetssikringsarbejdet.

Forskellige perspektiver på karakterudviklingen

Det er fælles for grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser, at særligt mundtlige prøver, projektopgaver (grundskoleområdet) og afsluttende opgaver (videregående uddannelser) afviger særligt meget fra den forventede fordeling. For disse prøveformer bliver der i særligt høj grad givet for mange høje karakterer og for få lave karakterer i forhold til det forventede. I og med at karakterer for skriftlige prøver følger den forventede karakterfordeling, betyder afvigelserne også en skævvridning af den samlede karakterfordeling.

Fokusgruppeinterviews med bedømmere på grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser samt ledere på de videregående uddannelser bibragte også forskellige perspektiver på, hvorfor andelen af 12-taller er øget. En forklaring er, at 12-tallet i starten blev anvendt lige så forsigtigt som topkarakteren på 13-skalaen, og at stigningen derfor skyldes, at bedømmerne nu anvender den oftere, end da skalaen blev indført. En anden forklaring er, at man på særligt grundskoleområdet og de gymnasiale uddannelser kan opleve et pres på karakterniveauet. Det hænger sammen med, dels at den enkelte elevs adgang til ungdomsuddannelse/videregående uddannelse er afhængig af elevens karaktergennemsnit, og dels at karakterer ofte har forvaltningernes fokus og i større eller mindre grad anvendes til at sammenligne og rangere.

Bedømmerne og lederne oplever også, at eleverne/de studerende er mere målrettede og har stort fokus på karakterer og derfor på målopfyldelse i forbindelse med prøver. Der er formentlig flere forhold, der har bidraget til denne udvikling, herunder indførelsen af karakterkrav på erhvervsuddannelserne, et større fokus på karaktergennemsnit og betydning af dette for optagelse på en videregående uddannelse. Indførelse af læringsmål har også gjort det tydeligere for eleverne/de studerende, hvad de bliver bedømt på, og har gjort det muligt for eleverne/de studerende at yde en mere målrettet indsats til prøverne. Dette relaterer sig ikke til karakterskalaen eller anvendelsen af den, men til uddannelsernes læringskultur.

Sidst, men ikke mindst, fremhæves det, at tvivl eller uenighed om karakteren kommer eleven/den studerende til gode, hvorfor der kan være en tendens til at give eleven/den studerende en højere karakter. Dette har formentlig altid gjort sig gældende. Det er ikke muligt inden for denne evalueringsramme at afgøre, hvilken betydning dette har haft for karakterudviklingen, da det ikke er undersøgt, hvor ofte det sker, ligesom der heller ikke er viden om, hvor ofte det gjorde sig gældende under 13-skalaen.

1.3.2 Evalueringsspørgsmål 2: Anvendes skalaen, som den var tænkt af Karakterkommissionen, og er kommissionens krav til en moderne karakterskala fortsat aktuelle?

7-trins-skalaen er en absolut karakterskala. Karaktererne skal derfor gives ud fra opfyldelse af læringsmål og de beskrivelser af karaktertrinene, som findes i karakterbekendtgørelsen (trinbeskrivelserne). Med implementeringen af 7-trins-skalaen blev der på alle uddannelsesområder indført et krav om, at disse læringsmål for alle prøver skulle foreligge i skriftlig form (eksplicit).

I det følgende ser vi på, om bedømmerne giver karakterer på den måde, som det var tænkt af Karakterkommissionen, samt hvordan skalaens trinbeskrivelser, supplerende bedømmelsesværkøjer (kun grundskolen, de gymnasiale uddannelser og erhvervsuddannelserne) samt læringsmålene (kun de videregående uddannelser, da læringsmålsmålene her i højere grad fastlægges decentralt) opleves og anvendes.

Karakterkommissionen beskrev desuden, at der skal ske en kvalitetssikring af karaktergivningen, så prøvernes tilrettelæggelse og læringsmålene over tid og for store populationer understøtter, at de bestående karakterer følger en fast fordeling. Kvalitetssikringen skal desuden understøtte en ensartet brug af karakterskalaen.

Kvalitetssikringsarbejdet er organiseret forskelligt på tværs af uddannelsesområderne, og der er ikke beskrevet krav med hensyn til dette i karakterbekendtgørelsen. I det følgende ser vi derfor nærmere på, hvordan arbejdet fungerer i praksis, og hvordan de forskellige aktører ser deres egne roller.

Skalaen anvendes som en absolut skala af bedømmerne

Karakterkommissionen lagde til grund for brugen af 7-trins-skalaen, at karaktergivning skal være absolut:

> En absolut karaktergivning af en eksamenspræstation foretages alene ud fra en vurdering af, i hvor høj grad eksamenspræstationen lever op til de mål, der er fastsat for prøven. For at kunne foretage en absolut karaktergivning af en eksamenspræstation er det derfor en forudsætning, at der er opstillet præcise mål og vurderingskriterier for præstationer ved den pågældende prøve. Karaktergivningen foretages helt uafhængigt af kendskab til og sammenligning med andre eksamenspræstationer.
>
> Betænkning om indførelse af en ny karakterskala til erstatning af 13-skalaen. Afgivet af Karakterkommissionen, november 2004

Der må altså i bedømmelsessituationen ikke tilstræbes en bestemt fordeling af karaktererne.

Fokusgruppeinterviews viser, at størstedelen af bedømmerne på tværs af grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser i overensstemmelse med forventningen og lovgivningen anvender skalaen absolut.

På grundskoleområdet peger nogle bedømmere dog på, dels at de møder bedømmere, som stadig har den opfattelse, at man skal tilstræbe en bestemt fordeling af karaktererne, dels at karakterskalaens relative element indirekte har betydning for deres karaktergivning, idet de oplever, at en del af bedømmelsesværktøjerne (bedømmelsesværktøjerne uddybes i afsnit 3.3) udformes ud fra ønsket om at ramme den intenderede fordeling. Det ligger dog uden for denne evaluerings rammer at afgøre, hvilken betydning dette har for anvendelsen af skalaen.

Lederne og censorformændene er i spørgeskemaundersøgelsen blevet bedt om at vælge tre styrker ved 7-trins-skalaen ud af ni mulige. Med undtagelse af på grundskoleområdet er "At bedømmelser sker ud fra eksplicitte læringsmål" blandt de tre mest valgte styrker på alle områder. 52 % af lederne på de gymnasiale uddannelser, 50 % af lederne på erhvervsuddannelserne, 70 % af lederne på de videregående uddannelser og 68 % af censorformændene har valgt dette som en styrke. Blandt lederne på de videregående uddannelser er det den mest valgte styrke. Dette tyder på, at karakterkommissionens krav om, at bedømmelsen skal være absolut frem for relativ, stadig er aktuel.

Trinbeskrivelsen opleves som anvendelig, men der er ikke noget entydigt billede af, om trinbeskrivelsens fokus på mangler er uhensigtsmæssigt

Helt overordnet er det værd at bemærke, at bedømmerne, dvs. lærere, undervisere og censorer på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser, er meget enige i deres oplevelse af 7-trins-skalaen. Det samme gælder deres forståelse og anvendelse af skalaen. Spørgeskemaundersøgelsen viser, at bedømmerne på tværs af områderne oplever, at de har tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin på 7-trins-skalaen (94-95 % af bedømmerne er meget enige eller enige heri), og at de har et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer (92-95 % af bedømmerne er meget enige eller enige heri). Bedømmerne oplever også, at deres medbedømmere forstår karakterskalaen på samme måde som dem (mellem 79-85 % af bedømmerne er meget enige eller enige heri).

Beskrivelsen af de forskellige trin på 7-trins-skalaen (trinbeskrivelsen) lægger vægt på mangler i præstationen. Fx beskrives karakteren 7 som følger: "Karakteren 7 gives for den gode præstation, der demonstrerer opfyldelse af fagets mål, med en del mangler." 7-trins-skalaen betegnes derfor ofte som en mangelskala.

Størstedelen af bedømmerne er enige i, at de forskellige karakterer er klart beskrevet i trinbeskrivelsen, og at beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning. Evalueringen tegner et billede af, at fokus på mangler i beskrivelsen af de enkelte karaktertrin er uhensigtsmæssigt. 40-50 % af bedømmerne er enige i, at trinbeskrivelserne ud fra et læringsperspektiv har for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), frem for hvad den studerende har lært (mestring). Omvendt er omkring en fjerdedel af bedømmerne (21-25 %) uenige eller meget uenige i dette. Et lignende billede tegner sig, når bedømmerne spørges om, hvorvidt beskrivelsen af karaktertrinene i for høj grad lægger op til, at bedømmelsen foretages på baggrund af mangler. Det er 39-50 % af bedømmerne enige i, mens 24-31 % er uenige eller meget uenige.

I fokusgrupper og interviews er det et udbredt perspektiv, at det ud fra et læringsperspektiv ikke er meningsfuldt at skulle bedømme eleverne/de studerende ud fra mangler. Dels fordi mangler i en læringssituation, ifølge bedømmerne, ikke er entydigt negativt, men også kan rumme et udviklingspotentiale, og dels fordi bedømmerne beskriver, at det, når det drejer sig om at bygge mennesker op, er centralt at bygge på det, de kan, og ikke alene på det, de ikke kan.

Evalueringen peger på, at der derfor kan være behov for at se nærmere på trinbeskrivelsen, herunder det hensigtsmæssige i, at denne har fokus på mangler. Spørgsmålet om, hvorvidt trinbeskrivelsen skal ændres, relaterer sig i høj grad til, hvilken læringskultur man ønsker at fremme.

Variation i vurderingen og anvendelsen af bedømmelsesværktøjerne på grundskoleområdet, de gymnasiale uddannelser og erhvervsuddannelserne.

På *grundskoleområdet* og *de gymnasiale uddannelser* udarbejder Undervisningsministeriet en række bedømmelsesværktøjer til de skriftlige og mundtlige prøver til understøttelse af karaktergivningen.

På *erhvervsuddannelserne* anvendes en kombination af centrale og lokale bedømmelsesværktøjer. I grundfag er rammerne for faget og prøven beskrevet i den centralt udarbejdede læreplan med faglige mål, kernestof, eksaminationsgrundlag og bedømmelsesgrundlag, men opgaven til prøven stilles af skolen på grundlag af den prøvebeskrivelse, som skolen inden for rammerne af den centralt udarbejdede læreplan har udfoldet i den lokale undervisningsplan. De afsluttende prøver udformes typisk i et samarbejde mellem skolerne og de faglige udvalg. Uddannelsernes slutmål er fastlagt i uddannelsesbekendtgørelserne, og den konkrete tilrettelæggelse af prøven og udformningen af censorvejledning sker i et samarbejde mellem skolerne og de faglige udvalg.

Evalueringen viser, at bedømmerne på de tre ovenstående uddannelsesområder i varierende grad anvender de centrale bedømmelsesværktøjer. På grundskoleområdet er de mest anvendte centrale bedømmelsesværktøjer rettevejledningen, prøvevejledningen og vurderingsskemaerne. På det gymnasiale område er de mest anvendte centrale bedømmelsesværktøjer fagets læreplan, vejledningen til læreplanen, vejledningen til censorer og bedømmernes egne retteark til den skriftlige prøve. På erhvervsuddannelserne er de mest anvendte centrale bedømmelsesværktøjer fagbekendtgørelserne.

Bedømmerne på grundskoleområdet og de gymnasiale uddannelser beskriver i fokusgruppeinterviews, at bedømmelsesværktøjerne særligt kommer i spil, hvis der opstår tvivl eller uenighed med hensyn til bedømmelsen af en elev, og at bedømmelsesværktøjerne i de situationer bidrager til at understøtte en ensartethed i bedømmelsen.

Bedømmernes vurdering af de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, varierer. 82 % af bedømmerne på grundskoleområdet er meget enige eller enige i, at bedømmelsesværktøjerne, der stilles til rådighed centralt eller af uddannelsen, giver dem et solidt udgangspunkt for bedømmelsen. Det samme gælder 57 % og 60 % af bedømmerne på hhv. de gymnasiale uddannelser og erhvervsuddannelserne.

I og med at bedømmelsesværktøjerne har en stor betydning for anvendelsen af skalaen, og at evalueringen peger på, at der er visse udfordringer med bedømmelsesværktøjerne, kan der være behov for fremover at se nærmere på bedømmelsesværktøjerne på de gymnasiale uddannelser og erhvervsuddannelserne.

Læringsmål kan udfordre på de videregående uddannelser

På de videregående uddannelser fastlægges mange af læringsmålene decentralt på institutionerne. For erhvervsakademi- og professionsbacheloruddannelserne samt de maritime uddannelser fastlægges læringsmål for de fælles elementer dog centralt i Uddannelses- og Forskningsministeriet via uddannelsesbekendtgørelserne. Det er derfor relevant at belyse, hvordan arbejdet med og kvaliteten af læringsmålene opleves.

Bedømmerne på *de videregående uddannelser* er blevet spurgt om, hvorvidt de generelt mener, at de faglige mål er formuleret tydeligt og er anvendelige i bedømmelsessituationen. Under halvdelen af bedømmerne (44 %) mener, at det er tilfældet. Der er ikke noget, der indikerer, at det har nogen betydning, om læringsmålene er centralt eller decentralt formuleret.

Censorformændene på de videregående uddannelser er samtidig blevet bedt om at vurdere den aktuelle status med hensyn til institutionernes arbejde med opstilling af læringsmål. 89 % af censorformændene er meget enige eller enige i, at der er opstillet læringsmål for alle faglige prøver. 74 % er meget enige eller enige i, at de faglige læringsmål er formuleret, så de understøtter bedømmelser ud fra målopfyldelse. 20 % af censorformændene er meget enige eller enige i, at der er stor forskel på kvaliteten af de faglige læringsmål inden for respondentens censorområde. Fokusgrupper med ledere og bedømmere på de videregående uddannelser viser desuden, at læringsmålene

mange steder revideres ofte, hvilket alt andet lige vanskeliggør en ensartet bedømmelse, da bedømmerne ofte skal bedømme ud fra nye mål. Desuden giver fokusgrupperne og de individuelle interviews et indtryk af, at der er forskellige hensyn. I en bedømmelsessituation skal målene helst være så konkrete som muligt, men for ikke at revidere dem hele tiden eller låse undervisningen for meget skal målene helst være på et vist abstraktionsniveau. Desuden viser evalueringen, at ændring af læringsmål oftest er drevet af andre hensyn, fx aftagernes behov, og ikke udvikling i karaktergivning. Derudover angiver blot 8 % af lederne på de videregående uddannelser, at de faglige læringsmål bliver justeret, hvis der er problemer med hensyn til karaktergivningen. Ledernes typiske reaktion er i stedet at drøfte det med underviserne/de interne bedømmere.

Bedømmerne på de videregående uddannelser beskriver i fokusgruppeinterviews og individuelle interviews, at læringsmålene særligt kommer i spil, hvis der opstår tvivl eller uenighed med hensyn til bedømmelsen. Læringsmålene kan derfor bidrage til at understøtte en ensartethed i bedømmelserne.

Læringsmål har altså stor betydning for skalaens anvendelse på de videregående uddannelser. Da evalueringen peger på, at der er visse udfordringer med læringsmålene, kan der være behov for fremover at fokusere mere på dette område.

Understøttelse af bedømmerne bygger i høj grad på en-til-en-oplæring

Evalueringen viser, at understøttelsen af bedømmerne er individfokuseret og ikke nødvendigvis en systematiseret praksis. Det er således især en-til-en-oplæring eller sparring fra kollegaer, overværelse af en kollegas prøveafholdelse og afholdelse af de første prøver sammen med erfaren censor/eksaminator, som lederne tilbyder nye bedømmere såvel som erfarne bedømmere, mens kurser eller seminarer er mindre udbredte. Sidemandsoplæring har mange fordele, men også en indbygget risiko for, at misforståelser med hensyn til, hvordan skalaen skal anvendes, og lignende, videreføres fra én bedømmer til en anden.

Spørgeskemaundersøgelsen blandt bedømmere peger ikke på noget større behov for mere understøttelse. Størstedelen af bedømmerne på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser føler sig godt klædt på til at bedømme prøverne. Dog er 31-44 % af bedømmerne enige i, at det vil være nyttigt at have mere fokus på bedømmelsespraksis på deres skole/institution, og 16-34 % er enige i, at der er behov for, at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne. Fokusgrupperne giver indtryk af, at bedømmerne særligt oplever, at der kan være behov for en større systematik med hensyn til at klæde nye bedømmere og censorer på til karaktergivning, mens de erfarne bedømmere oplever, at netop erfaring, herunder særligt erfaring som censor, klæder dem godt på. Hvad angår censorer på det videregående uddannelsesområde, afholder 21 % af censorformændene i dag kurser for nye censorer, mens 14 % afholder kurser for erfarne censorer. Uddannelses- og Forskningsministeriet arbejder dog på nuværende tidspunkt på et fælles censorkursus for alle censorer på det videregående uddannelsesområde.

Evalueringen peger på, at der kan være behov for at se nærmere på understøttelse og oplæring af nye bedømmere.

Ledelsernes fokus på kvalitetssikring varierer meget på tværs af sektorer

Karakterkommissionen beskrev, at 7-trins-skalaen skal bruges ensartet i hele uddannelsessektoren. Da karaktergivningen sker på grundlag af målopfyldelse, kræver det, at det faglige niveau koordineres. Kommissionen beskrev samtidig, at det skulle sikres, at de bestående karakterer over tid og for store populationer nogenlunde skulle følge en fast, forventet fordeling (10 % 02-taller, 25 % 4-taller, 30 % 7-taller, 25 % 10-taller og 10 % 12-taller). Det skulle ske gennem kvalitetssikring af ka-

raktergivningen, så prøverne og over tid også de faglige mål blev tilpasset til de studerende/eleverne. Den forventede fordeling skulle ikke realiseres ved den enkelte prøve eller det enkelte år, men for store populationer og over tid.

Kvalitetssikringsarbejdet er forankret forskelligt på de forskellige uddannelsesområder. Kvalitetssikringsopgaven er kun beskrevet af Karakterkommissionen og indgår ikke i karakterbekendtgørelsen. Det er derfor relevant at afdække, hvordan kvalitetssikringsopgaven forstås og håndteres i praksis.

Kvalitetssikring på grundskoleområdet, de gymnasiale uddannelser og erhvervsuddannelserne

På *grundskoleområder og de gymnasiale uddannelser* er kvalitetssikringen todelt. Dels er der den interne kvalitetssikring på den enkelte skole/institution, herunder hvordan skolen/institutionen og skolens ledelse arbejder med at sikre kvalitet og ensartethed i karaktergivningen, dels er der den centrale kvalitetssikring af karaktergivningen, herunder hvordan Undervisningsministeriets læringskonsulenter, fagkonsulenter, opgaveudvalg og opgavekommissioner arbejder med kvalitetssikringen, når de producerer og kvalitetssikrer prøveopgaver og gennemfører forcensur i forbindelse med prøveafholdelsen. Med hensyn til den centrale kvalitetssikring indberetter skolerne på *grundskoleområdet* og *det gymnasiale område* prøvekaraktererne til Styrelsen for It og Læring. Undervisningsministeriet, herunder læringskonsulenter, fagkonsulenter, opgaveudvalg og opgavekommissioner, samler op på karakterfordelingen til prøverne og på censorernes erfaringer med opgavesættene og med elevernes forståelse og besvarelse heraf. Opsamlingen på censorernes erfaringer og karakterfordelingerne kan give anledning til fremover at justere opgavernes indhold, sværhedsgrad og formulering. Samtidig kan man justere i vejledninger, hvis fx nogle elementer i prøven og faget viser sig vanskelige eller uklare for bedømmerne at bedømme. Opsamlingen på karakterfordelinger kan ifølge kvalitetssikringsaktørerne være et vigtigt redskab til at opnå viden om, ved hvilke typer opgaver der er særlig brug for opmærksomhed.

På *erhvervsuddannelserne* er der ekstern censur i forbindelse med alle prøver. Undervisningsministeriet beskikker særlige censorer i udvalgte grundfag med henblik på stikprøvevis at kvalitetssikre prøveafholdelsen på erhvervsskolerne. De særlige censorer afrapporterer til ministeriets fagkonsulenter. Undervisningsministeriet afholder desuden et årligt censormøde, hvor resultatet af den særlige censur drøftes. De afsluttende prøver gennemføres ligeledes med ekstern censur, oftest med censorer fra erhvervslivet. Prøverne gennemføres typisk på baggrund af detaljerede prøvebeskrivelser og censorvejledninger med tydelige bedømmelseskriterier.

Evalueringen viser, at størstedelen af lederne på *grundskoleområdet, de gymnasiale uddannelser* og *erhvervsuddannelserne* i meget høj grad eller i høj grad ser det at understøtte, at bedømmerne bruger karakterskalaen på en korrekt og ensartet måde, og at bedømmerne har et ensartet niveau i forståelsen af kravene i de faglige læringsmål, som en del af deres arbejde med karaktergivningen. De fleste ledere på *grundskoleområdet* og *de gymnasiale uddannelser* anser også det at følge op på karakterstatistik for de afholdte prøver og evaluering af karakterpraksis ved afholdte prøver som en del af deres arbejde med karaktergivningen. Kvalitetssikringen både centralt og på de enkelte skoler har primært fokus på prøvekarakterer frem for standpunktskarakterer. Evalueringen viser dog, at 42 % af bedømmerne på grundskoleområdet, 31 % af bedømmerne på de gymnasiale uddannelser og 44 % af bedømmerne på erhvervsuddannelserne er enige i, at det vil være nyttigt at have mere fokus på bedømmelsespraksis på deres skole/institution, og hhv. 34 %, 16 % og 34 % er enige i, at der er behov for, at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne.

Kvalitetssikring på de videregående uddannelser

På *de videregående uddannelser* foregår kvalitetssikring primært lokalt på de enkelte institutioner. Derudover har censorformændene en rolle, idet de som led i kvalitetssikringen af uddannelserne blandt andet skal afgive en årlig beretning og medvirke til en løbende dialog om uddannelsens, fagets eller fagområdets udvikling ved mindst hvert andet år at afholde censormøder med censorerne i censorkorpset og kontaktmøder mellem uddannelsesinstitutionerne og censorerne, jf. eksamensbekendtgørelserne på det videregående uddannelsesområde. Det er dog ikke beskrevet i lovgivningen, hvad kvalitetssikringsopgaven indebærer for ledelserne, ligesom der ikke lovgivningsmæssigt eller i vejledninger og opgavebeskrivelser til fx censorformændene er fastsat rammer for, hvem der skal monitorere karaktergivningen eller hvordan.

Evalueringen viser, at størstedelen af lederne på de videregående uddannelser ser det at følge op på evt. tilbagemeldinger fra censorsystemet angående karaktergivningen som en del af deres arbejde med karaktergivningen, mens omkring halvdelen ser det at følge op på karakterstatistik for de afholdte prøver som en del af deres arbejde med karaktergivningen, og 42 % af lederne ser det at understøtte, at der evalueres på karaktergivningspraksis ved de afholdte prøver, som en del af deres arbejde med karaktergivningen. Der er interessante forskelle mellem ledere af hhv. erhvervsakademi-, professionsbachelor- og universitetsuddannelser, hvor lederne af universitetsuddannelserne er mindre tilbøjelige til at se kvalitetssikring af karaktergivning som en del af deres arbejde.

Selvom opfølgning på karaktergivning ikke specifikt er beskrevet som en opgave for censorformændene, viser evalueringen, at størstedelen af censorformændene (71 %) følger op på udviklingen i karaktergivningen inden for censorområdets uddannelser. Det er dog under halvdelen (42 %) af censorformændene, der har adgang til karakterstatistik for alle bedømmelser på korpsets uddannelser.

Den største afvigelse fra den forventede karakterfordeling finder vi på de videregående uddannelser, herunder særligt universitetsuddannelserne. I og med at kvalitetssikring i modsætning til de øvrige områder foregår decentralt, kan der være behov for at sætte fokus på de videregående uddannelsers kvalitetssikring af karaktergivningen, herunder om denne er i tråd med Karakterkommissionens tanker.

1.3.3 Evalueringsspørgsmål 3: Hvad betyder skalaens karakterspring og karakteren -3 for karaktergivningen?

Med overgangen fra 13-skalaen til 7-trins-skalaen reducerede man antallet af trin fra ti til syv, blandt andet med det formål at skabe en klarere trindifferentiering. En af begrundelserne for at indføre større spring var at understøtte en tydelig differentiering mellem karaktererne og dermed en større sikkerhed i karaktergivningen. Karakterkommissionens begrundelse for reduktionen af antallet af trin var, at kommissionen i forbindelse med interviews og samtaler ofte blev mødt med det synspunkt, at det kan være svært at afgøre, om en given præstation skal vurderes til fx 7 eller 8 eller fx 8 eller 9. Karakterkommissionen fremhævede dog samtidig, at 7-trins-skalaen skulle kunne benyttes i forbindelse med den løbende vurdering.

De store spring (4-7-10) skal afspejle, at de tre karakterer er brede (dvs. at de opsamler hhv. 25 %, 30 % og 25 % af bedømmelserne). Dette står i kontrast til 13-skalaen, hvor midterkaraktererne numerisk lå mere tæt, herunder de tre midterkarakterer 7, 8 og 9.

Af de 7 trin er to ikke-bestået: 00, der gives for den utilstrækkelige præstation, der ikke demonstrerer en acceptabel grad af opfyldelse af fagets mål, og -3, der gives for den helt uacceptable præstation. De to karakterer skulle give mulighed for at sende et signal til den studerende om, hvor

langt præstationen var fra at være bestået. Karakterernes talværdier er desuden fastsat, så de understøtter, at andelen, der ikke består deres uddannelse samlet set, ligger på nogenlunde samme niveau som ved 13-skalaen.

De store spring på skalaen har deres svagheder

Evalueringen viser, at de store spring for en stor del af bedømmerne ikke understøtter en større sikkerhed i karaktergivning. På tværs af grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser angiver et mindretal af bedømmerne (28-40 %), at de er meget enige eller enige i, at de brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om den rigtige karakter. Særligt bedømmerne på grundskoleområdet oplever dette. Samtidig vurderer de færreste bedømmere (fra 20 % (videregående uddannelser) til 28 % (grundskole)), at det ikke har nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinene forskellige steder på skalaen.

Spørgeskemaundersøgelsen blandt bedømmerne viser, at flertallet oplever, at der er områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have. Det gælder særligt springene mellem 4, 7 og 10. Dette handler om, at lærerne/underviserne ikke ønsker at trække elevens/den studerendes gennemsnit ned, men også om, at bedømmerne oplever, at 7-tallet skal rumme for forskellige præstationer.

Dog oplever en gruppe bedømmere, at de brede karaktertrin har gjort det lettere at opnå enighed mellem censor og eksaminator. De efterspørger ikke at kunne differentiere karaktergivningen yderligere med flere trin på skalaen.

Kun 10 og 12 betragtes som gode karakterer

Bedømmerne fortæller i fokusgruppeinterviewene og de kvalitative interviews, at eleverne/de studerende har en tendens til kun at betragte 10 og 12 som gode karakterer, mens 7-tallet ikke opleves som en god karakter. Blandt gymnasieeleverne er der desuden et øget fokus på karaktergennemsnit og dermed på de enkelte gymnasiekarakterer, da de har en oplevelse af, at det har stor betydning med hensyn til at komme ind på en videregående uddannelse. Det betyder ifølge bedømmerne på *de gymnasiale uddannelser*, at en del elever er villige til at gøre en ekstra og meget målrettet strategisk indsats for at forbedre deres karakterer. Lærerne oplever desuden karaktergivningen som et stort ansvar, fordi karaktererne betyder så meget for eleverne, og de store spring kan gøre en stor forskel for elevernes gennemsnit. Mellem 40 % (erhvervsuddannelserne) og 53 % (videregående uddannelser) angiver da også, at de store spring mellem 4, 7 og 10 jævnligt fører til en højere bedømmelse for at undgå at give eleven/den studerende en for lav karakter. Dette er ikke direkte relateret til skalaen, men det har betydning for anvendelsen af den i de konkrete bedømmelsessituationer og kan have haft betydning for karakterinflationen i løbet af de seneste år.

Samlet set peger evalueringen på, at den tydelige differentiering i form af de store spring på midten af skalaen for størstedelen af bedømmerne ikke har medført en større sikkerhed i karaktergivningen. For en betydelig gruppe af bedømmere har de store spring ført til, at de giver en højere karakter for ikke at give en for lav karakter. Dette hænger formentlig sammen med det store fokus, der er på karaktergennemsnit i disse år, og det kan have haft betydning for karakterinflationen.

Svært at anvende skalaen til løbende vurdering

Karakterkommissionen angiver i sin rapport, at det forventes, at skalaen kan anvendes til løbende vurdering. Bedømmerne på *grundskoleområdet* og *de gymnasiale uddannelser* oplever, at skalaen er svær at anvende til løbende vurdering. De brede karaktertrin kan gøre det svært for eleverne at rykke sig fra én karakter til en anden, fordi trinene er så brede. I fokusgrupper og interviews peger bedømmerne på, at dette gør det svært at synliggøre elevernes progression i forbindelse med lø-

bende standpunktskarakterer, hvilket kan skade elevernes motivation. Langt størstedelen af bedømmerne (73-82 %) er da også enige eller meget enige i, at karaktererne på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne.

Lederne bakker op om dette, idet 56-74 % angiver, at det, at karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne, er en væsentlig svaghed ved skalaen.

Evalueringen viser altså, at skalaen i dag ikke opleves som særligt anvendelig til løbende vurdering, da særligt det brede 7-tal skal rumme for forskellige præstationer og besværliggør synliggørelse af progression.

-3 er en undtagelseskarakter med stor symbolsk betydning

På tværs af grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser fortæller bedømmerne, at de meget sjældent anvender karakteren -3. Denne pointe understøttes af registerundersøgelsen, som viser, at under 1 % af de karakterer, som er blevet givet siden 2007 på grundskoleområdet og på de gymnasiale uddannelser, har været -3. På de videregående uddannelser gør det sig gældende for ca. 3 % af karaktererne.

Kun få bedømmere angiver springet mellem 00 og -3 som et af de steder på skalaen, de har svært ved at bruge. Alligevel angiver en overvejende del af bedømmerne i spørgeskemaet, at de oplever, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende (53- 72 % af bedømmerne). Særligt bedømmerne på grundskoleområdet oplever, at -3 sender et unødigt negativt signal. Dette perspektiv vedrører både skalaen i sig selv og anvendelsen af skalaen.

Blandt bedømmerne på de gymnasiale og videregående uddannelser, der deltog i fokusgruppeinterviews, findes også det perspektiv, at anvendelsen af -3 har til formål at sende et signal til eleven/den studerende om, at det ikke er muligt at bestå faget, og at eleven/den studerende måske endda bør genoverveje sit studievalg. Dette ligger i tråd med Karakterkommissionens forventninger, som netop var, at de to ikke-beståede karakterer skulle sende forskellige signaler til eleven/den studerende.

Hvad angår grænsen mellem bestået og ikke-bestået, svarer 65-72 % af bedømmerne i spørgeskemaundersøgelsen, at denne er tilstrækkeligt tydelig på skalaen. Dette understøttes i fokusgruppeinterviewene, hvor bedømmerne fortæller, at de som regel ikke er i tvivl om, hvorvidt en elev/studerende skal dumpe til prøven eller ej.

Samlet set viser evalueringen, at mange bedømmere, især på grundskoleområdet, anser karakteren -3 som svær at anvende, fordi den er negativ, hvilket også afspejler sig i statistikken, der viser, at -3 sjældent anvendes. Der er dog en anden ikke-bestået karakter, 00, så dette vurderes ikke at have stor betydning for brugen af skalaen samlet set.

1.3.4 Evalueringsspørgsmål 4: Er der tilstrækkelige muligheder for at identificere særligt dygtige studerende på de videregående uddannelser og belønne den ekstraordinære præstation?

7-trins-skalaen har ikke en undtagelseskarakter til den ekstraordinære præstation som 13-skalaens 13-tal. 13-tallet blev givet for den usædvanlig selvstændige og udmærkede præstation, mens 12-tallet gives for den fremragende præstation, der demonstrerer udtømmende opfyldelse af fagets

mål med ingen eller få uvæsentlige mangler. Det forventes, at omkring 10 % af alle bestående karakterer er 12.

Det er et bevidst valg, at 7-trins-skalaen for at gøre skalaen mere internationalt anvendelig ikke har nogen undtagelseskarakter som det gamle 13-tal. Der er internationalt kun få karakterskalaer, som har en undtagelseskarakter. En Eurydice-rundspørge i de europæiske lande foretaget af Styrelsen for Forskning og Uddannelse peger på, at topkarakteren på de udenlandske skalaer ofte bruges ved 13-30 % af de bestående prøver, og tal fra USA peger på, at andelen kan være endnu større (se "7-trins-skalaen og ECTS-skalaen" – appendiks B). 13-tallet var derfor svært at forklare for personer i udlandet. Samtidig pegede Karakterkommissionen på, at der på tværs af sektorer, fag og hovedområder var stor forskel på, hvor ofte bedømmerne gav 13.

Det er dog blevet påpeget, at ændringen gør det sværere for aftagerne at identificere de dygtigste danske studerende, og at det er uhensigtsmæssigt, at skalaen ikke giver mulighed for at belønne den ekstraordinære præstation.

Uddannelses- og Forskningsministeriet har implementeret en talentbekendtgørelse, som blandt andet giver institutionerne mulighed for at belønne særligt dygtige studerende, men muligheden for belønning er ikke indbygget i karakterskalaen.

Stigning i brugen af 12-tallet

Siden indførelsen af 7-trins-skalaen har der været en stigning i brugen af 12-tallet på de videregående uddannelser samlet set og på tværs af de tre store sektorer. Udviklingen med hensyn til 12-tallet har været stort set identisk på professionsbachelor- og universitetsuddannelserne, hvor brugen er steget fra knap 13 % 12-taller i 2007 til omtrent 18 % 12-taller i 2016. Niveauet af uddelte 12-taller på erhvervsakademiuddannelserne er lavere, men til gengæld har stigningen været mere markant samlet set. Man er således gået fra at tildele et 12-tal ved omtrent 7 % af alle eksamener i 2007 til at tildele 12-taller ved 15 % af alle eksamener i 2016. Andelen af 12-taller er især stor i forbindelse med de afsluttende opgaver.

Efterspørgsel efter at kunne belønne den ekstraordinære præstation

Størstedelen af lederne og bedømmerne på *de videregående uddannelser* angiver, at de er meget enige eller enige i, at skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation. I spørgeskemaundersøgelsen svarer 64 % af bedømmerne, at fokus på mangler gør det sværere at belønne det ekstraordinære. Bedømmerne beskriver i fokusgruppeinterviewene 12-tallet som en meget bred karakter, der kan dække over meget forskellige præstationer. De savner at kunne differentiere og nuancere bedømmelsen af de studerende, der toppræsterer, sådan at man kan skelne mellem de studerende, der får 12, fordi de ikke havde væsentlige fejl og mangler og lever op til målbeskrivelsen, og de studerende, der får 12, fordi de leverede en ekstraordinær præstation.

Både lederne på de videregående uddannelser og censorformændene peger på, at det er en svaghed ved skalaen, at den ikke i tilstrækkelig grad rummer mulighed for at honorere den ekstraordinært gode præstation. Fokusgruppeinterviews med ledere og bedømmere viser samtidig, at der ikke er enighed om, hvorvidt der skal være en exceptionel karakter på karakterskalaen som 13-tallet på den gamle skala. Der er på den ene side en gruppe, som ønsker 13-tallet tilbage, og på den anden side en gruppe, som ønsker at belønne den ekstraordinære præstation, men som ikke ønsker en exceptionel karakter. Lederne og bedømmerne er samtidig opmærksomme på, at det gamle 13-tal stillede de studerende dårligere, når de ville studere i udlandet.

Interviews med arbejdsgiverorganisationerne viser, at selvom ansættelse ikke blot er baseret på karakterer, så spiller karakterer en rolle – særligt ved ansættelse af nyuddannede – og særligt i virksomheder, hvor man har mange ansøgere til en stilling. Arbejdsgiverorganisationerne ønsker ikke umiddelbart 13-tallet tilbage, men med stigningen i andelen, som får 12, oplever de, at virksomhederne har stadig sværere ved at identificere de ekstraordinært dygtige studerende. De peger derfor på, at det er centralt for dem, at inflation i karaktererne modarbejdes, så 12-tallet reelt kun tildeles i omegnen af 10 % af de bestående præstationer, som Karakterkommissionen beskrev det. Det gælder også i forbindelse med de afsluttende opgaver.

Spørgsmålet om, hvorvidt skalaen bør give mulighed for at belønne den ekstraordinære præstation, er relateret til selve skalaens udformning, da 7-trins-skalaen netop ikke indeholder en karakter for den ekstraordinære præstation. Evalueringen viser, at der på de videregående uddannelser er et udbredt ønske om at kunne belønne den ekstraordinære præstation samt at fremme en læringskultur, hvor der er fokus på mestring frem for mangler og på, at de studerende yder deres bedste. Det ligger dog ikke inden for denne evaluerings ramme at vurdere, om dette betyder, at der er behov for en ændring af skalaen, eller om andre mekanismer kan opfylde dette ønske.

1.3.5 Evalueringsspørgsmål 5: Er 7-trins-skalaen internationalt anvendelig?

13-skalaens opbygning blev oplevet som en udfordring for danske studerendes mulighed for at blive optaget på gode uddannelsesinstitutioner i udlandet. Der var stor forskel på de karakterskalaer, der blev brugt verden over, men kun få havde en undtagelseskarakter. Især det danske 13-tal blev derfor oplevet som en udfordring. Et centralt krav til 7-trins-skalaen var, at den skulle være internationalt anvendelig. Derfor opbyggede Karakterkommissionen 7-trins-skalaen, så den var kompatibel med den europæiske ECTS-skala, som indgik i ECTS-systemet som et værktøj til oversættelse af karakterer. Danmark gik derfor fra en karakterskala med ti trin til en skala med syv med en fast forventet fordeling over tid og uden en undtagelseskarakter.

Stadig flere danske studerende rejser ud som en del af deres uddannelsesforløb. Danmark modtager også stadig flere internationale studerende, og karakterer udgør fortsat en central del af adgangskravene i forbindelse med både udvekslingsophold og hele uddannelser i Danmark såvel som i udlandet. Det er derfor stadig relevant at tilstræbe, at de danske karakterer er internationalt anvendelige. I udlandet er der ofte også fokus på karakterer i forbindelse med opnåelse af stipendier, og der er i højere grad praksis for at overføre karakterer til eksamensbeviserne, end der er i Danmark. De danske karakterers anvendelighed i udlandet er derfor også et parameter, når det handler om at understøtte et godt uddannelsessamarbejde med udlandet.

Der er ikke sket en harmonisering af karakterskalaerne internationalt; de europæiske skalaer er stadig meget forskellige. Mange lande har skalaer med bogstavkarakterer, men disse benyttes ofte forskelligt med forskellige forståelser eller krav med hensyn til at opnå de enkelte karakterer. Den danske skala skiller sig blandt andet ud ved at bruge talkarakterer, som ikke har en eksplicit kobling til antallet af point, som er opnået i bedømmelsen.

Overordnet set peger evalueringen på, at 7-trins-skalaen fungerer internationalt, også efter at ECTS-skalaen er udgået af ECTS-systemets anbefalinger. 7-trins-skalaen hindrer ikke mobilitet eller indgåelse af udvekslingsaftaler. Evalueringen peger dog også på, at skalaen kræver formidling i udlandet. Det skyldes, både at den ikke ligner de typiske skalaer i andre lande, og at opkoblingen til ECTS-skalaen kan give anledning til misforståelser i lande, der bruger bogstavkarakterer.

Formidling er centralt, og ECTS-opkoblingen kan udfordre

Evalueringen peger på, at international anvendelighed ofte handler mere om formidlingsarbejde end om, hvordan karakterskalaen konkret er opbygget.

Både internationale medarbejdere og studerende, der rejser ud, peger på, at Uddannelses- og Forskningsministeriets officielle engelsksprogede beskrivelser af 7-trins-skalaen ofte ikke er tilstrækkelige. Beskrivelserne formår ikke i tilstrækkelig grad at bygge bro mellem på den ene side den danske skala og den danske karaktergivningspraksis og på den anden side skalaerne i de lande, som Danmark især sender studerende til, herunder USA og Australien. Evalueringen finder dog, at det kun i mindre grad er en udfordring for danske studerende, der rejser ud. Det kan dog være en udfordring for udenlandske studerende, som skal have danske karakterer med hjem. Det kan kræve supplerende formidling og dialog med de udenlandske institutioner om, hvad karakterne betyder.

Inden for Europa opleves ECTS-skalaen stadig som en hjælp, men uden for Europa kan skalaen være en udfordring. Evalueringen peger på, at ECTS-skalaens bogstaver også kan forvirre, fordi et C (oversættelsen af et dansk 7-tal til ECTS-skalaen) fx anses for at være en dårlig karakter i det amerikanske og kinesiske uddannelsessystem. De danske karakterer 4 og 02 oversættes til hhv. D og E, som i andre skalaer er ikke-bestået eller betegner dumpegrænsen. Da oversættelsen indgår i de danske eksamensbeviser og de officielle engelsksprogede beskrivelser af 7-trins-skalaen, kan institutionerne og de studerende ikke bare styre uden om ECTS-skalaens bogstavkarakterer, når de formidler de danske karakterer.

Hovedparten af de internationale medarbejdere peger på, at det vil være en udfordring for deres arbejde, hvis 7-trins-skalaen ændres. Skiftende karakterskalaer kræver et stort formidlingsarbejde. Det gælder både med hensyn til relationen til samarbejdspartnere i udlandet og med hensyn til at få etableret en forståelse af den danske skala i udlandet, så de udenlandske institutioner behandler danske studerende retfærdigt, når de søger om optagelse på egen hånd. Nogle repræsentanter peger på, at det kunne være en fordel, hvis den danske skala knyttede sig til en meget brugt udenlandsk skala. Drøftelserne i begge grupper peger dog på, at det også ville kunne skabe udfordringer, da skalaerne ofte bruges forskelligt, og det ikke er klart, hvilken skala der ville være mest relevant.

De interviewede internationale medarbejdere peger også på, at bestået/ikke-bestået karakterer er en udfordring. De accepteres ofte ikke i udlandet og kan gøre det svært for danske studerende at dokumentere deres kompetencer.

Internationale anbefalinger om brug af statistik afspejles ikke i Danmark

I Europa er der et stigende fokus på at understøtte, at statistik om karaktergivning kan bruges i forbindelse med oversættelse af karakterer. Det fordrer en skala med et vist antal bestået karaktertrin og en gennemsigtighed med hensyn til, hvordan karakterskalaerne faktisk bruges. ECTS-systemets nye anbefalinger er derfor, at der skal offentliggøres engelske beskrivelser af, hvordan de nationale skalaer faktisk bruges, samt statistik om karaktergivningen (karakterfordelinger for de bestået karakterer samt andele af ikke-bestået). Det anbefales også, at oversættelse og omregning af karakterer skal ske direkte mellem de nationale skalaer ved hjælp af karakterfordelinger.

Evalueringen peger dog på, at omregningerne ikke altid er retvisende, hvis de gøres 100 % mekanisk. Det faglige niveau kan variere selv mellem lande, der bruger ens karakterskalaer, og nogle gange også mellem institutioner og uddannelser inden for samme land. Forskellene kan være så markante, at der er behov for at foretage tilpasninger af omregningsmetoden, hvis den skal være

rimelig. Det påpeger de internationale medarbejdere, men det påpeges også i litteraturen på området, og metoden går igen i flere af de optagelsessystemer, som Styrelsen for Forskning og Uddannelse har undersøgt.

Karakterfordelinger er en hjørnesten i ECTS-systemets tilgang til gennemsigtighed om karaktergivning. Gymnasiers og grundskolers karakterstatistik er offentligt tilgængelig, og andre lande bruger den, når de omregner danske karakterer i forbindelse med optag. De videregående uddannelser skal offentliggøre karakteroplysninger – gennemsnit og karakterer for afsluttende opgaver – men de er ikke forpligtede til at offentliggøre karakterfordelinger i en form, der kan bruges til karakteroversættelser, selvom det er en del af anbefalingerne i ECTS-systemet. Der er derfor i mange tilfælde ikke gennemsigtighed med hensyn til den faktiske karaktergivning, og karakteromregning efter ECTS-systemets metode er derfor ikke mulig.

Selvom de videregående uddannelsesinstitutioner typisk ikke arbejder med officielle oversættelser af karakterer fra udlandsophold, viser evalueringen, at de har brug for at oversætte og vurdere karakterer fra udlandet i flere sammenhænge. Det gælder fx ved vurdering af studerende, der søger om optagelse, understøttelse af egne dimittender, der søger videre, tildeling af udlandsophold til egne studerende ud fra tidligere karakterer etc. Evalueringen peger på, at institutionernes praksis kan være meget forskellig i den sammenhæng. Nogle arbejder systematisk med opsamling af karakterstatistik og systemunderstøttelse af karakteroversættelser, mens andre laver ad hoc-vurderinger, når det er relevant.

Fokusgruppeinterviews med de internationale medarbejdere og enkeltinterviews med studerende peger på to overordnede holdninger med hensyn til større transparens i forbindelse med karaktergivningen. Den ene holdning er, at offentliggørelse af karakterstatistik og fælles omregningssystemer kan være en hjælp og understøtte en mere ensartet behandling af karakterer fra udlandet. Den anden holdning er, at det vil være for ressourcekrævende at bruge og ikke nødvendigvis vil være retvisende.

De internationale medarbejdere peger generelt på, at krav om at oversætte karakterer og faste regler og systemer for det kan være ressourcekrævende at administrere, og de peger på, at det kan være svært at opnå stor præcision i omregningerne, fordi karaktergivningen kan variere inden for lande og nogle gange også inden for institutioner.

1.4 Datagrundlag

Rapporten bygger på følgende metoder:

- Deskriptive og regressionsbaserede analyser af data for alle tilgængelige afgivne karakterer i perioden 2007-16 (2007-15 for de gymnasiale uddannelser)

  Erhvervsuddannelserne er ikke repræsenteret i de deskriptive og regressionsbaserede analyser, da der kun forelå data for en toårig periode, og det dermed ikke var muligt at analysere udviklingen i karaktergivningen.

- Spørgeskemaundersøgelse blandt bedømmere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser
- Spørgeskemaundersøgelse blandt ledere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser samt censorformændene på de videregående uddannelser.

I de kvalitative analyser var der fokus på grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser, idet der kun var fokus på de områder, som indgik i evalueringens fase 1. I og med at de deskriptive og regressionsbaserede analyser af data om alle tilgængelige afgivne karakterer i perioden 2007-16[1] viste, at der særligt på de videregående uddannelser er sket en karakterinflation, er der særligt fokus på dette område, og derfor er der gennemført flere fokusgrupper, enkeltinterviews samt fokusgruppeinterviews med ledere. Følgende kvalitative interviews er gennemført:

- Tre fokusgruppeinterviews med samlet set 15 bedømmere inden for fagene dansk, engelsk og matematik på grundskoleområdet
- Fire fokusgrupper med samlet set 22 bedømmere inden for fagene dansk, engelsk og matematik på de gymnasiale uddannelser
- Ti fokusgrupper med samlet set 66 bedømmere på de videregående uddannelser
- To fokusgrupper med internationale medarbejdere på de videregående uddannelsesinstitutioner
- Fire fokusgruppeinterviews med samlet set 30 ledere på de videregående uddannelser
- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på grundskoleområdet
- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på de gymnasiale uddannelser
- Fire dobbeltinterviews med bedømmere på de videregående uddannelser
- Fem enkeltinterviews med bedømmere på de videregående uddannelser
- Fem enkeltinterviews med opgavekommissionsformænd (tre) og læringskonsulenter (to) på grundskoleområdet
- Fem enkeltinterviews med opgavekommisionsformænd (to) og fagkonsulenter (tre) på stx
- Otte enkeltinterviews med studerende, som har læst hele eller dele af deres uddannelse i udlandet
- Rundspørge om karaktergivning i udlandet via Eurydice-netværket.

Mere information om datagrundlaget findes i bilagsrapporterne *Kortlægning og analyse af karaktergivningen – Baggrundsrapport til Evaluering af 7-trins-skalaen*, *Bedømmernes perspektiv – Bilagsrapport til Evalueringen af 7-trins-skalaen*, *Ledernes perspektiv – Bilagsrapport til Evalueringen af 7-trins-skalaen*, *Internationale medarbejdere og studerendes perspektiv på 7-trins-skalaens internationale anvendelighed – Bilagsrapport til Evaluering af 7-trins-skalaen* og 7-trins-skalaen og ECTS-skalaen – appendiks B.

[1] 2007-15 for de gymnasiale uddannelser.

1.5 Læsevejledning

Kapitel 2 præsenterer de analyser, der er relateret til evalueringsspørgsmål 1: "Hvordan har karaktergivningen udviklet sig siden implementeringen af 7-trins-skalaen, og har der været karakterinflation?".

Kapitel 3 ser nærmere på evalueringsspørgsmål 2: "Anvendes skalaen, som den var tænkt af Karakterkommissionen, og er kommissionens krav til en moderne karakterskala fortsat aktuelle?".

Kapitel 4 handler om evalueringsspørgsmål 3: "Hvad betyder skalaens karakterspring og karakteren -3 for karaktergivningen?".

Kapitel 5 præsenterer analyserne i forbindelse med evalueringsspørgsmål 4: "Er der tilstrækkelige muligheder for at identificere særligt dygtige elever og studerende og belønne den ekstraordinære præstation?". Bemærk, at der kun er fokus på de videregående uddannelser.

Kapitel 6 præsenterer analyserne i forbindelse med evalueringsspørgsmål 5: "Er 7-trins-skalaen internationalt anvendelig?".

Datagrundlaget i afrapporteringen

I denne rapport præsenteres udvalgte resultater fra den samlede evaluering af 7-trins-skalaen. Bilagsrapporterne og tabelrapporterne viser analyser og data fra den samlede dataindsamling.

Figurerne viser andelen af bedømmere og ledere, som svarer "Enig" eller "Meget enig" eller "I høj grad" eller "I meget høj grad" på et givent spørgsmål. Der er dog også en væsentlig andel af bedømmerne og lederne, som benytter sig af "Hverken enig eller uenig"-kategorien. I de tilfælde præsenteres andelen, som er uenige eller meget uenige, i tekstform.

Forskelle på tværs af fag, hovedområder og sektorer (videregående uddannelser) præsenteres, i det omfang der er tale om signifikante forskelle.

Citater anvendes til at eksemplificere eller illustrere pointer fra analysen.

2 Udvikling i karaktergivning

2.1 Baggrund og undersøgelsesspørgsmål

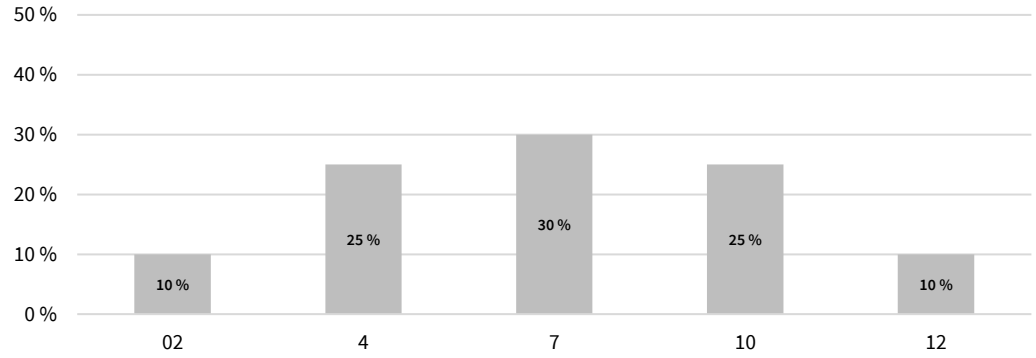
Karakterkommissionens betænkning anbefalede at anvende karakterskalaen ensartet i hele uddannelsessystemet. 7-trins-skalaen er konstrueret efter ECTS-skalaen, hvortil der på forhånd er knyttet en fast procentvis fordeling af de bestående karakterer, jf. figuren nedenfor.

Selvom 7-trins-skalaen er en absolut skala med karaktergivning ud fra målopfyldelse, havde Karakterkommissionen en forventning om, at de bestående karakterer over tid og for store populationer ville sprede sig nogenlunde normalfordelt. Det relative element i karakterskalaen betyder, at karaktergivningen over tid og for store populationer bør følge karakterfordelingen fra ECTS-skalaen, som fremgår af figur 2.1. Det gælder ned på de enkelte fag på grundskoleområdet og de gymnasiale uddannelser (fx danskfaget på de gymnasiale uddannelser på tværs af typer), men er ikke nærmere specificeret på de videregående uddannelser.

Der er ingen forventninger til brugen af *ikke-bestående karakterer,* ligesom der ikke er forventninger om en normalfordeling inden for en specifik prøveform.

FIGUR 2.1

Den forventede karakterfordeling



I dette kapitel sætter vi fokus på følgende evalueringsspørgsmål: ”Hvordan har karaktergivningen udviklet sig siden implementeringen af 7-trins-skalaen, og har der været karakterinflation?”

Kapitlet bygger på:

- Deskriptive og regressionsbaserede analyser af data om alle tilgængelige afgivne karakterer i perioden 2007-16 (2007-15 for de gymnasiale uddannelser)

Erhvervsuddannelserne er ikke repræsenteret i de deskriptive og regressionsbaserede analyser, da der kun forelå data for en toårig periode, og det dermed ikke var muligt at analysere udviklingen i karaktergivningen.

- Fokusgruppeinterviews med bedømmere på grundskoleområdet, de gymnasiale uddannelser og videregående uddannelser samt ledere på de videregående uddannelser.

2.2 Grundskoleområdet følger overordnet den forventede fordeling

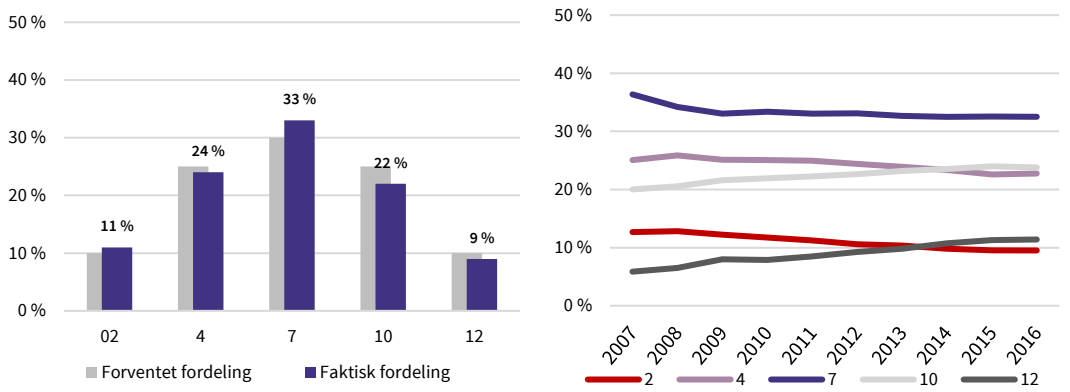
Karaktergivningen på grundskoleområdet *følger den forventede fordeling* på et aggregeret niveau over tid, jf. figur 2.2. Gennemsnitskarakteren har været stigende, men det afspejler, at der i starten af perioden blev givet færre 12-taller end forventet. Samlet set følger karaktergivningen på grundskoleområdet dermed forventningerne, og i det seneste år (2016) ser den ud til at være fordelt jævnt i forhold til den forventede normalfordeling.

Figur 2.2 viser desuden, at der siden indførelsen af 7-trins-skalaen har været en *opadgående trend med hensyn til brugen af 12-taller*, så der nu er den forventede andel af 12-taller. 12-andelen er steget fra ca. 5 % i 2007 til ca. 11 % i 2016.

Der er forskelle mellem prøve- og standpunktskarakterer. Det viser sig eksempelvis, ved at der bliver givet for høje karakterer i forhold til den forventede fordeling ved mundtlige afgangsprøver, hvorimod standpunktskarakterer og skriftlige prøvekarakterer ligger tættere på det forventede niveau.

FIGUR 2.2

Fordeling af og udvikling i karakterer i grundskolen, 2007-16, %



Gennemsnit: 6,8 12-andel: 9 % Fejlprocent: 9 % Største afv.: 7 og 10

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til Evalueringen af 7-trins-skalaen*.
Note: Figuren til venstre viser den forventede fordeling af alle bestående karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle bestående karakterer ved grundskolen – alle karakterer. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin for karakterer i grundskolen – alle karakterer i perioden 2007-16. N = 19.671.118.

Engelsk adskiller sig fra de øvrige fag, idet karaktererne i faget allerede lå på den forventede fordeling i 2007, og siden er karakterniveauet steget. Det betyder, at andelen af 12-taller i engelsk ligger over det forventede niveau i 2016. En dybere analyse afslører, at dette skyldes andelen af 12-taller, der bliver givet i de mundtlige afgangseksamener, der i 2016 nåede et niveau på 26 %.

Det er blevet undersøgt, om udviklingen i brugen af 12-taller kan skyldes, at elevpopulationen generelt er blevet dygtigere siden indførelsen af 7-trins-skalaen. Nærmere regressionsanalyser af sammenhængen mellem udviklingen i brugen af 12-taller over tid og udviklingen i diverse baggrundsvariable[2] relateret til den studerende og dennes forældre viser, at ændringen i elevpopulationen ikke kan forklare udviklingen i brugen af 12-taller i perioden 2010-16.

Andelen af ikke-beståede karakterer er i hele perioden på ca. 3 %. Følger man udviklingen over tid, ses et tydeligt fald i andelen af karakteren 00, hvorefter andelen har ramt et nogenlunde stabilt niveau på omkring 2 %. Karakteren -3 har ligget stabilt lavt i alle årene.

2.3 De gymnasiale uddannelser følger overordnet den forventede fordeling

Analysen af de gymnasiale uddannelser viser, at den gennemsnitlige fordeling af alle tildelte karakterer for hele perioden 2007-15 *følger den forventede normalfordeling*, jf. figur 2.3. Billedet ser dog noget anderledes ud, når vi splitter op på år, fag og prøvetyper.

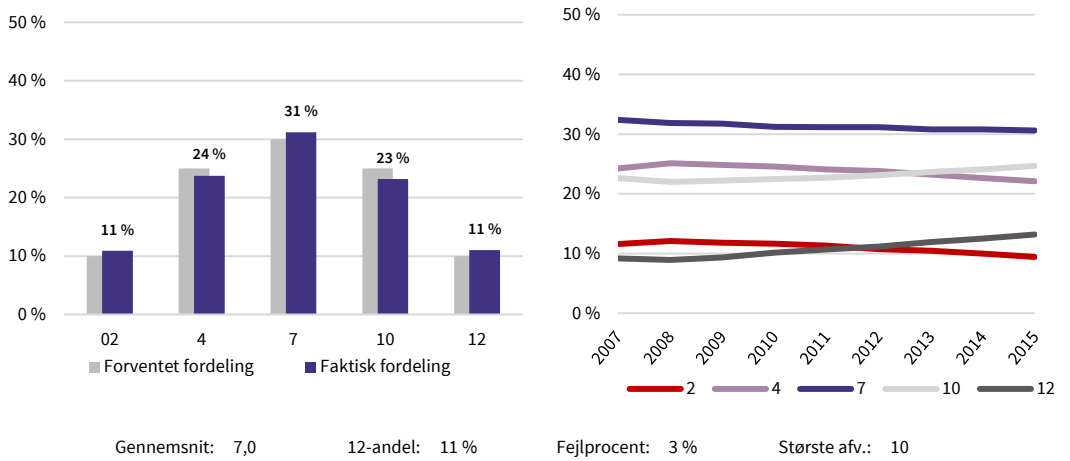
Figuren viser, at 12-andelen i begyndelsen af perioden lå på et niveau omkring det forventede, hvorefter andelen af 12-taller er steget til et højere niveau end det forventede, fra 9 % i 2007 til 13 % i 2015.

Der er forskel på tværs af karaktertyper (års- og eksamenskarakterer). Dette viser sig, ved at der ved skriftlige eksamener bliver givet for lave karakterer, mens der ved mundtlige eksamener bliver givet for høje karakterer, i forhold til den forventede fordeling. Både mundtlige og skriftlige årskarakterer har omvendt en tendens til at lægge sig tæt op ad den forventede fordeling.

2 Baggrundsvariable dækker over evner (på baggrund af tidligere uddannelsesforløb), køn, etnicitet, forældres uddannelsesniveau og forældres socioøkonomiske status.

FIGUR 2.3

Fordeling af og udvikling i karakterer på de gymnasiale uddannelser, 2007-15, %



Gennemsnit: 7,0 12-andel: 11 % Fejlprocent: 3 % Største afv.: 10

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen.*
Note: Figuren viser den forventede fordeling af alle beståede karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle beståede karakterer på gymnasiale uddannelser ud fra data for hvert enkelt år i perioden 2007-15. N = 9.620.249.

Der er forskelle mellem *de forskellige fag* inden for de gymnasiale uddannelser. For dansk og engelsk ser karakterfordelingen ud til at have nærmet sig den forventede fordeling over tid, mens karakterfordelingen for matematik og fysik bevæger sig længere og længere væk fra den forventede fordeling over tid. I fysik er det andelene af 4- og 7-taller, som afviger mest fra normalfordelingen. I matematik er det stort set alle karaktertrin, som afviger fra den forventede normalfordeling. Her er det anvendelsen af karakteren 02, der har den største afvigelse fra det forventede, mens anvendelsen af karaktererne 4, 10 og 12 ligeledes afviger fra den forventede fordeling. 12-andelen er således større end forventet, mens andelen af 4- og 10-taller har været mindre end de forventede 25 %. Karakterfordelingen i samfundsfag følger normalfordelingen, dog med en overvægt af 7-taller.

Karaktererne ved de mundtlige eksamener skiller sig generelt ud med en større andel af 12-taller i forhold til det forventede niveau. De mundtlige eksamener for næsten alle fag er karakteriserede ved en betydelig større andel af 12-taller. Omvendt bliver der givet en overvægt af lave karakterer ved skriftlige prøver, hvor særligt matematik og fysik på hf skiller sig ud med et lavt gennemsnit.

Ændringer i elevgruppens sammensætning kan ikke forklare den generelle stigning i andelen af 12-taller over tid. For de fleste fag kan inddragelsen af baggrundsoplysninger kun forklare en smule af udviklingen. Regressionsanalyser af sammenhængen mellem udviklingen i brugen af 12-taller over tid og udviklingen i diverse baggrundsvariable[3] relateret til den studerende og dennes forældre viser, at ændringen i elevpopulationen ikke kan forklare stigningen i brugen af 12-taller i perioden 2007-15.

3 Baggrundsvariable dækker over evner (på baggrund af tidligere uddannelsesforløb), køn, etnicitet, forældres uddannelsesniveau og forældres socioøkonomiske status.

Andelen af ikke-bestående karakterer er i hele perioden på ca. 4 %. Andelen af -3 er meget lille og konstant over tid, mens andelen af 00 er faldet i løbet af den samlede periode. Når vi opdeler på fag, kan vi se, at andelen af ikke-bestående karakterer er særligt stor på matematik, hvor den ligger på ca. 9 %.

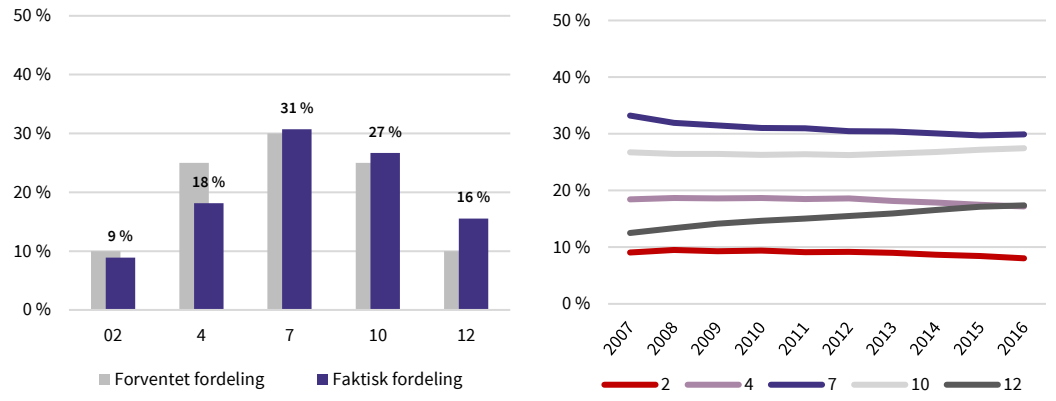
2.4 Det videregående uddannelsesområde følger ikke den forventede fordeling

Hovedresultaterne for det videregående uddannelsesområde er, at karaktergivningen for de bestående karaktertrin *ikke følger den forventede fordeling*, jf. figur 2.4. Vi ser i stedet en generel overvægt af høje karakterer og en undervægt af de lavere karakterer. Der er tale om en tendens, der går på tværs af samtlige uddannelsesniveauer med undtagelse af erhvervsakademiuddannelserne og de maritime uddannelser, hvor fordelingerne i højere grad svarer til de forventede.

Figuren viser desuden, at der har fundet karakterinflation sted siden 7-trins-skalaens indførelse, idet særligt andelen af 12-taller er steget over årene. I hele perioden 2007-16 har 15,5 % af de tildelte bestående karakterer været 12-taller, hvilket ligger over den forventede andel i henhold til normalfordelingen (10 %).

FIGUR 2.4

Fordeling af og udvikling i karakterer på alle videregående uddannelser, 2007-16, %



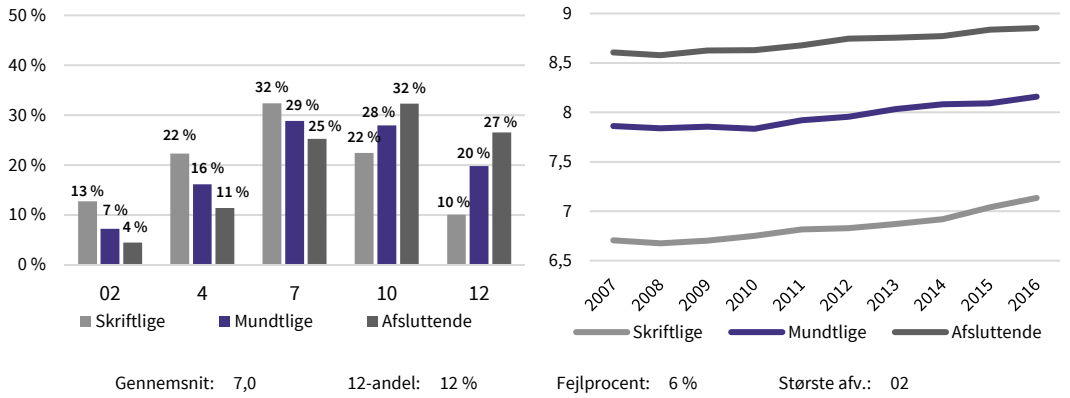
Gennemsnit: 7,6 12-andel: 16 % Fejlprocent: 22 % Største afv.: 4

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen*.
Note: Figuren til venstre viser den forventede fordeling af alle bestående karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle bestående karakterer på alle videregående uddannelser. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin på alle videregående uddannelser i perioden 2007-16. N = 7.248.401.

Der er en markant forskel på fordelingen af karakterer ved hhv. skriftlig, mundtlig og afsluttende eksamen, jf. figur 2.5. Karaktergivningen ved skriftlige eksamener følger den forventede fordeling aggregeret set i perioden 2007-16, mens vi ser en overvægt af høje karakterer ved mundtlige prøver i forhold til det forventede. Der er dog størst overvægt af høje karakterer i forbindelse med afsluttende prøver. Dette gælder på tværs af sektorer, institutioner og hovedområder.

FIGUR 2.5

Karakterfordeling og udvikling i gennemsnitskarakterer opdelt på prøveformer, 2007-16, %, alle videregående uddannelser



Gennemsnit: 7,0 12-andel: 12 % Fejlprocent: 6 % Største afv.: 02

Kilde: Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen.
Note: Figuren til venstre viser karakterfordelingen for samtlige videregående uddannelser opdelt på eksamensform. Figuren til højre viser udviklingen i gennemsnittet i hele perioden opdelt på eksamensform. N = 4.913.025.

Der er ikke tegn på, at karakterinflationen skyldes en udvikling i studenterpopulationen. Dette kunne have været tilfældet, hvis uddannelserne havde optaget flere dygtigere studerende, som forventeligt ville opnå flere høje karakterer. Derimod har det øgede optag på de videregående uddannelser i de seneste ti år medført et øget optag af studerende, der har klaret sig mindre godt på deres gymnasiale uddannelse, og som har en mindre stærk social baggrund, hvilket ser ud til at have lagt en dæmper på karakterinflationen.

Forskellen mellem uddannelsesinstitutioner med hensyn til andelen af tildelte 12-taller kan i visse tilfælde forklares med forskelle i institutionernes studenterpopulationer. Dette er særligt tydeligt inden for universitetssektoren, hvor eksempelvis forskellen mellem 12-andelen på hhv. Københavns Universitet og Syddansk Universitet på kandidatuddannelserne kan forklares ved at korrigere for forældre- og gymnasiebaggrund. Dette billede er mindre tydeligt, når vi ser på forskelle mellem erhvervsakademier og professionshøjskoler.

Den største afvigelse fra det forventede finder vi med hensyn til karaktertrinet 4. Der gives markant færre 4-taller end forventet, og dette gør sig gældende på tværs af uddannelsesniveauer og hen over alle ti år – med enkelte undtagelser.

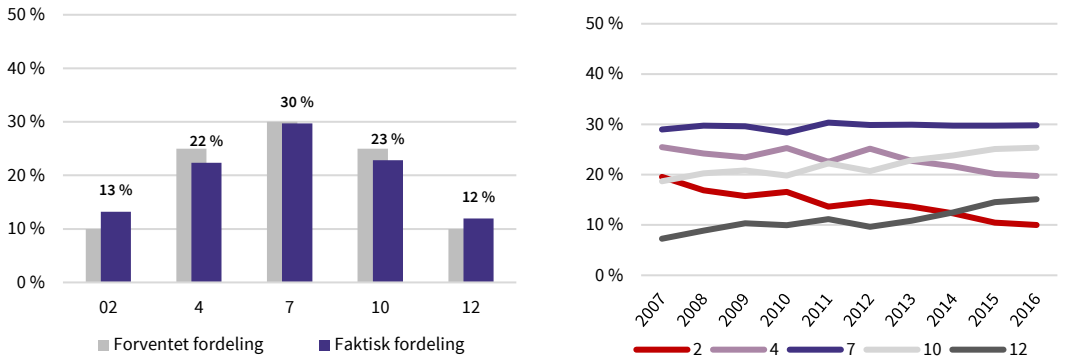
Analysen af ikke-beståede karakterer viser, at der i løbet af perioden har været en tendens til, at der på erhvervsakademiuddannelser bliver givet færre ikke-beståede karakterer. Karakteren 00 bliver benyttet på tværs af alle de videregående områder, mens karakteren -3 i særlig grad gives på områder, hvor der er en tradition for at aflevere blankt til eksamen.

Erhvervsakademiuddannelser følger den forventede fordeling, mens professionsbachelor-, bachelor- og kandidatuddannelser afviger

Figur 2.6 viser, at karaktergivningen på *erhvervsakademiuddannelserne* følger den forventede fordeling på et aggregeret niveau. Dette dækker dog over store forskelle mellem institutioner og hovedområder. Disse forskelle kan generelt ikke forklares ud fra forskelle med hensyn til studentergrundlag.

FIGUR 2.6

Fordeling af og udvikling i karakterer på erhvervsakademiuddannelser, 2007-16, %



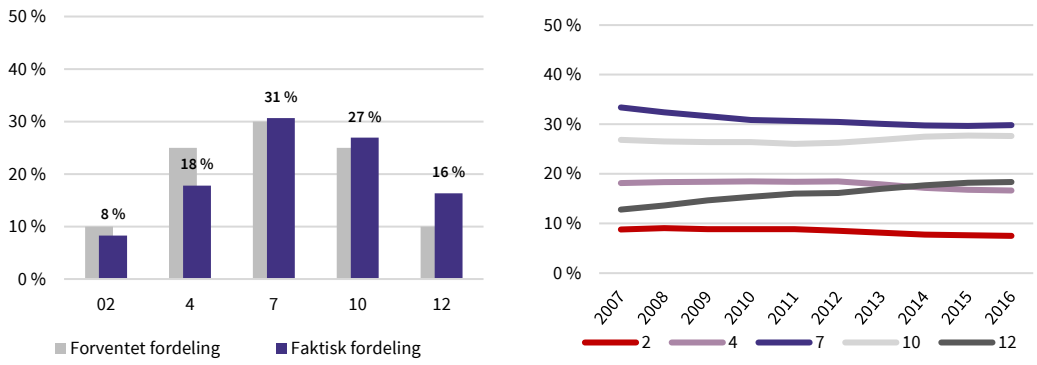
Gennemsnit: 7,0 12-andel: 12 % Fejlprocent: 6 % Største afv.: 02

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen*.
Note: Figuren til venstre viser den forventede fordeling af alle beståede karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle beståede karakterer på erhvervsakademiuddannelser. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin på erhvervsakademiuddannelser i perioden 2007-16. N = 504.690.

Der bliver generelt givet for stor en andel af 12-taller og for lille en andel af 4-taller på *professionsbacheloruddannelserne*, jf. figur 2.7. Kigger vi på mønstre på de forskellige hovedområder, er det særligt uddannelser inden for det designfaglige område samt medie, kommunikation og IT, som har en tendens til at tildele en meget stor andel af høje karakterer. Desuden ser det ud til, at der er en positiv sammenhæng mellem andelen af mundtlige eksamener og gennemsnitskarakterer.

FIGUR 2.7

Fordeling af og udvikling i karakterer på professionsbacheloruddannelser, 2007-16, %



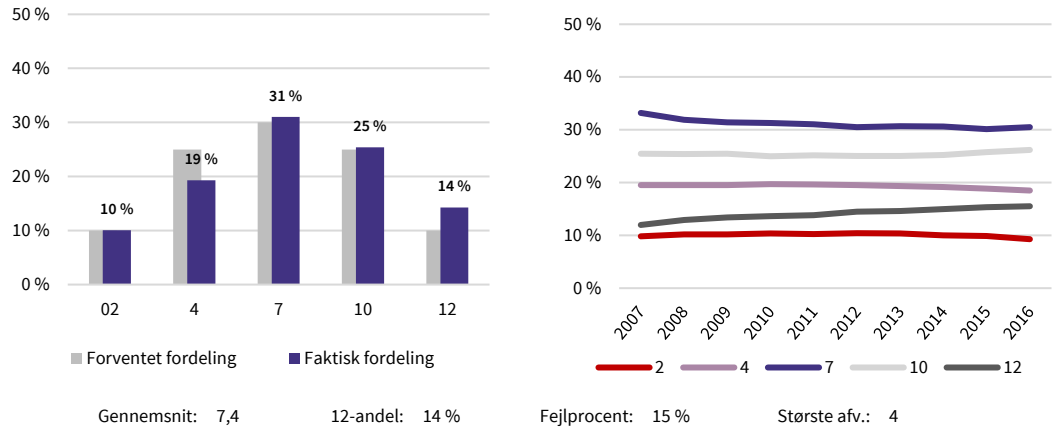
Gennemsnit: 7,7 12-andel: 16 % Fejlprocent: 25 % Største afv.: 4

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen.*
Note: Figuren til venstre viser den forventede fordeling af alle beståede karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle beståede karakterer på professionsbacheloruddannelser. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin på professionsbacheloruddannelser i perioden 2007-16. N = 1.830.829.

På *universitetsuddannelser* er der stor forskel i karaktergivningen på tværs af bachelor- og kandidatuddannelser. På *bacheloruddannelserne* ligger karaktergennemsnittet inden for den fastsatte grænse, om end andelen af 12-taller er for stor i forhold til det forventede niveau, jf. figur 2.8. På *kandidatuddannelserne* bliver der i særlig grad givet for mange 10- og 12-taller i forhold til det forventede niveau, og særligt 12-andelen er støt stigende over tid, jf. figur 2.9. Forskellen mellem bachelor- og kandidatuddannelser kan ikke forklares ud fra, at den afsluttende opgave fylder relativt mere på kandidatdelen end på bachelordelen, eftersom der ikke korrigeres for ECTS-vægte. På bacheloruddannelserne tildeler særligt RUC mange 12-taller, og på kandidatuddannelserne tildeler særligt DTU mange 12-taller.

FIGUR 2.8

Fordeling af og udvikling i karakterer på akademiske bacheloruddannelser, 2007-17, %

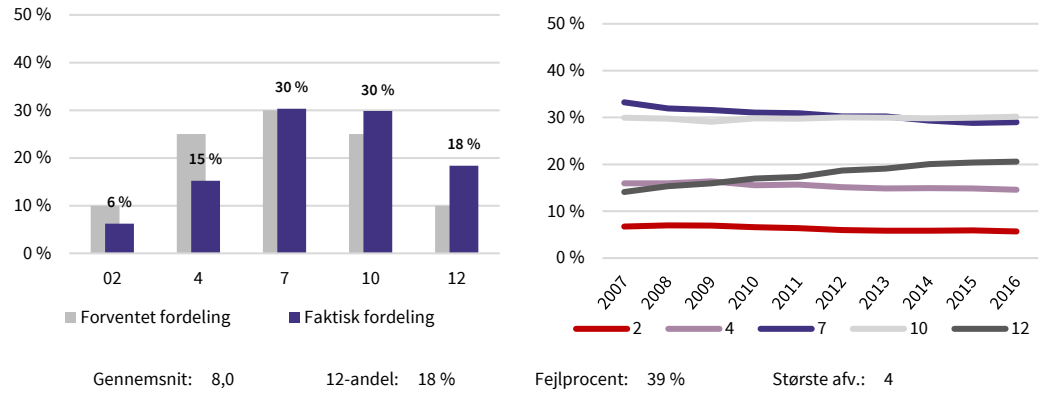


Gennemsnit: 7,4 12-andel: 14 % Fejlprocent: 15 % Største afv.: 4

Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen.*
Note: Figuren til venstre viser den forventede fordeling af alle bestående karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle bestående karakterer på akademiske bacheloruddannelser. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin på akademiske bacheloruddannelser i perioden 2007-16. N = 3.116.926.

FIGUR 2.9

Fordeling af og udvikling i karakterer på kandidatuddannelser, 2007-16, %



Kilde: *Kortlægning og analyse af karaktergivningen – baggrundsrapport til evaluering af 7-trins-skalaen*.
Note: Figuren til venstre viser den forventede fordeling af alle beståede karakterer (jf. Karakterkommissionens anbefalinger) samt den faktiske karakterfordeling for alle beståede karakterer på kandidatuddannelser. Figuren til højre viser den tidsmæssige udvikling i den relative andel af hvert bestået karaktertrin på kandidatuddannelser i perioden 2007-16. N = 1.555.320.

De *kunstneriske uddannelser* skiller sig ud, ved at karakterfordelingen er præget af meget færre lave karakterer og mange flere høje karakterer, end den forventede fordeling tilsiger. Dog ser der ikke ud til at have været nævneværdig karakterinflation siden 2011/12 på de kunstneriske uddannelser som helhed.

De *maritime uddannelser* følger den forventede karakterfordeling, samtidig med at der ikke ses en tendens til karakterinflation. Det kan dog bemærkes, at særligt karakteren 4 bliver givet relativt sjældent.

2.5 Perspektiver på stigning i andelen af karakteren 12

Fokusgruppeinterviews med bedømmere på alle områder samt ledere på de videregående uddannelser frembragte forskellige perspektiver på, hvorfor andelen af 12-taller er steget:

- 12-tallet blev i starten set som en exceptionel karakter.
- Nogle oplever pres med hensyn til at give højere karakterer.
- Eleverne/de studerende er mere målrettede og villige til at yde en ekstra indsats ved eksamen for at få en god karakter.
- Tvivl eller uenighed om karakteren kommer eleverne/den studerende til gode.
- Mundtlige prøver eller prøver med vejledning giver højere karakterer, mens skriftlige prøver ligger på det forventede karakterniveau.

12-tallet blev i starten set som en exceptionel karakter

Særligt bedømmerne på *grundskoleområdet* peger på, at stigningen i andelen af 12-taller kan hænge sammen med bedømmernes omstilling fra 13-tallet til 12-tallet. Bedømmerne har skullet

vænne sig til, at 12-tallet rummer andre og bredere faglige præstationer, hvor 10 % af eleverne forventes at få 12, sammenlignet med 13-tallet, hvor forventningen var, at blot 1 % af eleverne skulle have 13.

Nogle oplever pres med hensyn til at give højere karakterer

Særligt på *grundskoleområdet* oplever nogle bedømmere et pres med hensyn til at give højere karakterer. Bedømmerne oplever, at dette pres har forskellige årsager og kommer fra forskellige sider. Nogle beskriver, at det hænger sammen med, at karakterer ofte har forvaltningernes fokus og i større eller mindre grad anvendes til at sammenligne og rangere, hvilket fremgår af citaterne herunder:

> Når man ser på kommunens fokus, så er det nationale tests osv. Man ser på karakterer og faglige mål. Man ser jo ikke på inkluderende læringsfællesskaber og elever med særlige behov. Hvorfor ser man ikke på, om man formår at inkludere og rykke elever – også socialt?
>
> Bedømmer, grundskole

> Jeg ved ikke, om der ligger noget ubevidst i, at alle føler sig pressede i forhold til at skulle levere høje karaktergennemsnit og gode resultater. Fx sidste år var jeg ude som censor. Der sagde læreren: "Vi skal helst have nogle gode karakterer, ellers ryger vi på en obs-liste". Så tænkte jeg "det er jo frygteligt". Læreren var meget optaget af, at vi gav gode karakterer.
>
> Bedømmer, grundskole

Andre bedømmere kæder det sammen med karakterkravene til gymnasiet:

> Efter at der er kommet alle de her karakterkrav til gymnasiet, så kan man godt blive lidt presset til at skulle give dem lidt højere [karakterer]. Man kan godt have den tanke i baghovedet, at man føler, at de ikke kan bruge 4-tallet til noget.
>
> Bedømmer, grundskole

Eleverne/de studerende er mere målrettede og villige til at yde en ekstra indsats ved eksamen for at få en god karakter

En anden medvirkende forklaring på tendensen til højere karakterer er, ifølge bedømmerne, at eleverne/de studerende har et større fokus på karakterer og er mere målrettede og villige til at yde en indsats op til prøverne for at sikre sig en god karakter – og kun 10 og 12 anses som gode karakterer. To bedømmere fortæller her, hvordan dette kommer til udtryk:

> Den der holdning, at 7 ikke er godt nok, den hænger også sammen med, at der bliver givet mange flere 12-taller. Det forskyder hele skalaen. 7-tallet er blevet lidt devalueret. Det er ærgerligt, for det er da en skidegod karakter.
>
> Bedømmer, grundskole

> De, der bliver mest kede af det, er dem, der får 10. Der er så store ambitioner om, at man skal have 12. Det er en kulturel udvikling. Det skaber en stigende forventning om, at det er nemmere at få 12, og at det er det eneste acceptable.
>
> Bedømmer, videregående uddannelse

Blandt gymnasieeleverne er der desuden et øget fokus på karaktergennemsnit og dermed på de enkelte gymnasiekarakterer, da de har en oplevelse af, at det har stor betydning med hensyn til at komme ind på en videregående uddannelse. Det betyder ifølge bedømmerne på *de gymnasiale uddannelser*, at flere elever end tidligere er villige til at gøre en ekstra og meget målrettet strategisk

indsats for at forbedre deres karakterer. Bedømmerne oplever desuden karaktergivningen som et meget stort ansvar, fordi karaktererne betyder så meget for eleverne, og de store spring kan påvirke elevernes gennemsnit.

På *de videregående uddannelser* giver fokusset på karakterer, og for nogle studerende specifikt på 12-taller, ifølge bedømmerne sig udslag i, at de studerende er blevet mere fokuserede og bedre til at leve op til forventningerne. En bedømmer siger: "Det er interessant at se, hvor gode de er blevet til at kunne det, de bliver bedømt på". En anden bedømmer udtrykker det således:

> Vi snakker tit om, at vores studerende har knækket koden med hensyn til eksamen. De har lært, hvordan man lever op til kravene, uden at de egentlig laver noget som helst nyt. De opfylder bare kravene til perfektion. Dem har jeg givet en del 12-taller, som jeg har begrædt.
>
> Bedømmer, videregående uddannelse

Nogle bedømmere på de videregående uddannelser oplever på den baggrund også, at nogle studerende har fokus på at vælge projekter, som de oplever som sikre med hensyn til at få en god karakter, men som ikke udfordrer dem eller indebærer meget nytænkning. En bedømmer fortæller fx:

> Underviserne på de fagligt vanskelige områder får færre projekter. De studerende vælger altså de mere sikre projekter, hvor de får gode karakterer. De tør ikke tage de projekter, der udfordrer dem så meget, at de måske ikke kan få 12.
>
> Bedømmer, videregående uddannelse

Indførelse af læringsmål har på den måde gjort det tydeligere for eleverne/de studerende, hvad de bliver bedømt på, og gjort det muligt for eleverne/de studerende at yde en meget mere målrettet indsats til prøverne. Det relaterer sig ikke til karakterskalaen eller anvendelsen heraf, men i stedet til den læringskultur, der er på uddannelserne.

Tvivl eller uenighed om karakteren kommer eleverne/de studerende til gode

Derudover beskriver bedømmerne på *de gymnasiale uddannelser* og *de videregående uddannelser*, at de, hvis der er uenighed eller tvivl om, hvorvidt en elev/studerende skal have den ene eller den anden af to karakterer, der ligger ved siden af hinanden på skalaen, drøfter, hvad der taler for hhv. det ene og det andet, men også lader tvivlen komme eleven/den studerende til gode og ved fortsat tvivl vælger den højeste karakter. Det bakkes op af spørgeskemaundersøgelsen, hvor 53 % af bedømmerne på de videregående uddannelser og 43 % af bedømmerne på de gymnasiale uddannelser er enige i, at springene mellem 4, 7, og 10 jævnligt fører til en højere bedømmelse for ikke at give eleven/den studerende en for lav karakter (jf. figur 5.4). Når så stor en andel af bedømmerne jævnligt giver en højere bedømmelse, kan det være en medvirkede faktor med hensyn til den karakterinflation, som vi ser på særligt de videregående uddannelser.

Mundtlige prøver eller prøver med vejledning giver højere karakter, mens skriftlige prøver ligger på det forventede karakterniveau

Endelig peger bedømmerne – særligt på *de videregående uddannelser* – på, at prøveformen har betydning for karakterfordelingen. Generelt oplever bedømmerne, at det er lettere at få et 12-tal mundtligt end skriftligt. En af forklaringerne er, at eleverne/de studerende generelt opleves som stærkere mundtligt end skriftligt. En anden forklaring er, at fejl og uklare svar står tydeligere frem i en skriftlig prøve, hvor der ved mundtlige prøver er mulighed for at præcisere evt. misforståelser. Forskellen mellem mundtlige og skriftlige prøver illustreres i følgende citater:

> Noget af det, vi lægger vægt på [ved mundtlige prøver], er deres evne til analyse og refleksion. Nogle studerende kommer ind med en fantastisk analyse af deres egen opgave. Det hæver jo karakteren. Det er sjældent, at nogen kommer og taler sig selv ned. Det sker indimellem, men det er ikke tit. De kan blive stående eller gå op.
>
> Leder, professionshøjskole

> Ved skriftlige eksamener tæller vi fra bunden og op ved hjælp af et pointsystem. Der er hovedspørgsmål og underspørgsmål, som hver er tildelt en pointmængde. Man starter fra 00 og arbejder sig op i karaktersystemet. Når man er i tvivl om to karakterer, så ser man på helhedsindtrykket: Er der store tomme områder eller mindre fejl hist og pist?
>
> Bedømmer, videregående uddannelse

Afsluttende prøver med vejledning

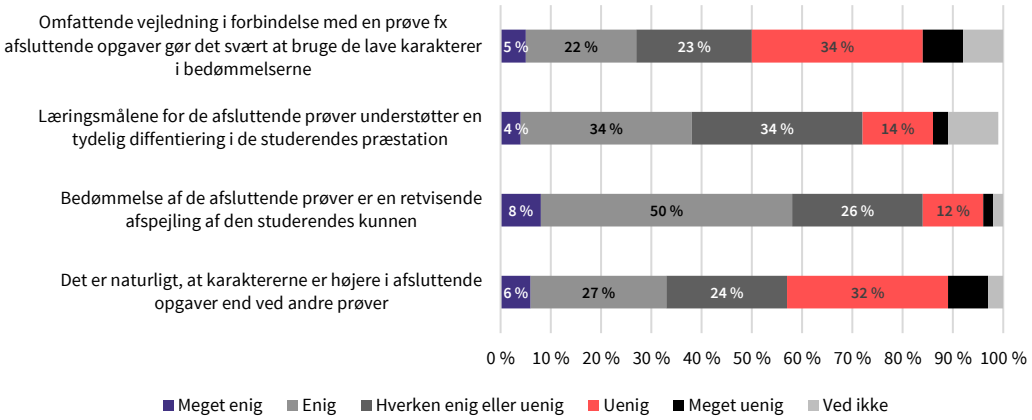
Andelen af 12-taller ved de afsluttende prøver på de videregående uddannelser var samlet set 27 % i perioden 2007-16 (se figur 3.5). Andelen har været stigende i perioden fra 21,8 % i 2007 til 29 % i 2016. På erhvervsakademiuddannelserne var andelen 17 % i hele perioden, mens den var 24 % på professionsbacheloruddannelserne og 30 % på universitetsuddannelserne.

Blandt bedømmerne på de videregående uddannelser er der meget delte meninger om, hvorvidt det er naturligt, at karakteren er højere ved afsluttende opgaver end ved andre prøver, jf. figur 3.9. 33 % af bedømmerne er meget enige eller enige i, at det er naturligt, at karaktererne er højere i forbindelse med afsluttende opgaver end ved andre prøver. Omvendt er 40 % uenige eller meget uenige. Samtidig er størstedelen (58 %) meget enige eller enige i, at bedømmelse af de afsluttende prøver er en retvisende afspejling af den studerendes kunnen.

Bedømmerne på universitetsuddannelserne (38 %) er i højere grad enige i udsagnet "Det er naturligt, at karaktererne er højere i forbindelse med afsluttende opgaver end ved andre prøver" end bedømmere på professionsbacheloruddannelserne (24 %) såvel som bedømmere på erhvervsakademiuddannelserne (25 %).

FIGUR 2.10

Hvor enig eller uenig er du i følgende udsagn om bedømmelse af afsluttende opgaver? Videregående uddannelser



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen i forbindelse med en afsluttende prøve i løbet af 2016, 2017 eller 2018. N = 2.556.

Fokusgruppeinterviewene med bedømmere og ledere på de videregående uddannelser frembragte et andet perspektiv, nemlig at det ikke overraskede, at karaktererne er højere ved de afsluttende prøver. Bedømmerne og lederne peger på flere årsager til, at karaktererne ligger højt i forbindelse med netop afsluttende projekter.

For det første beskriver de, at der i løbet af et uddannelsesforløb altid er frafald blandt de studerende, og en del af de frafaldne vil være studerende, der har det fagligt svært på uddannelsen. To bedømmere siger:

> Det er de modne, erfarne studerende, der kommer op i noget, der interesserer dem. Der har været et udskillelsesløb undervejs. Så jeg ser det ikke som noget mærkeligt.
>
> Bedømmer, videregående uddannelse

> Det giver mening, at karaktererne er højere i forbindelse med bachelorens afsluttende prøve. Sådan skal det se ud. Det ville være pinligt, hvis vores studerende ikke kunne præstere godt, når de har god tid og har gået på uddannelsen i tre år.
>
> Bedømmer, videregående uddannelse

For det andet mener lederne og bedømmerne, at det har betydning, at de studerende selv vælger, hvad de vil beskæftige sig med i forbindelse med den afsluttende opgave, og at de derfor har mulighed for at vælge et emne og fagområde, de interesserer sig særligt for og er dygtige inden for, og dermed også har større sandsynlighed for at klare sig godt i. Nogle bedømmere oplever også, at de studerende mere generelt vælger projekter og kurser, som de oplever som sikre, når det handler om at få gode karakterer, frem for at vælge projekter, der udfordrer, men hvor de måske ikke får 12.

For det tredje peger nogle bedømmere på, at det har betydning for niveauet, at de studerende i særlig grad får vejledning undervejs i arbejdet med de afsluttende projekter, hvilket fremgår af følgende citat:

> Det er meget tydeligt, at karaktererne for bacheloropgaven ligger højere end ved de løbende, skriftlige eksamener. Det har at gøre med formen. Det er en stor skriftlig opgave, hvor de studerende har individuel kontakt med vejleder. Og det kan være med til at løfte mellemgruppen op.
>
> Bedømmer, videregående uddannelse

I og med at skriftlige prøver følger den forventede karakterfordeling, er det især karaktergivningen ved de mundtlige og afsluttende prøver, der bidrager til, at karaktergivningen samlet set afviger fra den forventede fordeling.

3 Forståelse og anvendelse af skalaen

3.1 Baggrund og undersøgelsesspørgsmål

3.1.1 Karaktergivning skal være absolut

Karakterkommissionen havde en række forventninger til anvendelsen af karakterskalaen. For det første forventede man – og dette er også beskrevet i lovgivningen – at karaktergivningen skulle være absolut. Absolut karaktergivning defineres på følgende måde: "En absolut karaktergivning af en eksamenspræstation foretages alene ud fra en vurdering af, i hvor høj grad eksamenspræstationen lever op til de mål, der er fastsat for prøven. For at kunne foretage en absolut karaktergivning af en eksamenspræstation er det derfor en forudsætning, at der er opstillet præcise mål og vurderingskriterier for præstationer ved den pågældende prøve. Karaktergivningen foretages helt uafhængigt af kendskab til og sammenligning med andre eksamenspræstationer."[4]

Karaktergivning skal ske ud fra på forhånd fastlagte læringsmål
Et andet udgangspunkt var, at bedømmelserne skulle ske ud fra på forhånd fastlagte læringsmål, som også skulle være tilgængelige for elever/studerende. Dette blev betegnet som et skifte fra indholdsstyring til målstyring på uddannelserne.

På *grundskoleområdet og de gymnasiale uddannelser* fastlægges de faglige mål for de enkelte fag centralt i Undervisningsministeriet. Ministeriet fastlægger desuden konkrete vurderings-/bedømmelseskriterier for de enkelte prøver, og de skriftlige prøver afholdes som nationale prøver. Derved fastlægges et fælles fagligt niveau centralt i ministeriet. Som supplement til de faglige mål udarbejder Undervisningsministeriet en række andre redskaber til at støtte op om en korrekt og ensartet bedømmelse, herunder prøvevejledninger og karakterbeskrivelser for prøverne/fagene.

På *erhvervsuddannelserne* fastlægges de faglige mål centralt i læreplaner, uddannelsesbekendtgørelser og uddannelsesordninger. Bedømmelseskriterierne kan være fastlagt både centralt og decentralt. Eksempelvis beskrives bedømmelseskriterier for grundforløbs- og svendeprøver decentralt i et samarbejde mellem de faglige udvalg og skolerne. Uddannelsernes slutmål er centralt beskrevet i uddannelsesbekendtgørelsen for den enkelte erhvervsuddannelse. Målene for grundfagene fastlægges af Undervisningsministeriet, mens de uddannelsesspecifikke mål beskrives i et samarbejde mellem Undervisningsministeriet og uddannelsernes faglige udvalg.

På de videregående uddannelser fastlægges læringsmål for erhvervsakademi- og professionsbacheloruddannelserne og de maritime uddannelser i vid udstrækning centralt i Uddannelses- og Forskningsministeriet via uddannelsesbekendtgørelserne, mens de institutioner, som udbyder de

4 Betænkning om indførelse af en ny karakterskala til erstatning af 13-skalaen. Afgivet af Karakterkommissionen, november 2004.

kunstneriske uddannelser og universitetsuddannelserne, selv sammensætter strukturen i deres uddannelser og fastlægger læringsmålene. På alle uddannelser sker kvalitetsarbejdet dog decentralt, ligesom alle prøver tilrettelægges decentralt, dog med censur fra et nationalt censorkorps ved de fleste prøver.

Kvalitetssikring forventes for at sikre den relative fordeling

Karakterkommissionen beskrev også, at der skulle ske en kvalitetssikring af karaktergivningen ved at koordinere målene for faget, prøvens tilrettelæggelse og den population, der skal bestå prøven. Det skulle understøtte, at de bestående karakterer over tid og for store populationer nogenlunde følger en fast fordeling (karakterfordelingen i ECTS-skalaen), og at skalaen bruges på samme måde i hele uddannelsessystemet. Dette relative element i karakterskalaen fremgår dog ikke af den juridiske beskrivelse af karakterskalaen i karakterbekendtgørelsen.

I dette kapitel sætter vi fokus på følgende evalueringsspørgsmål: "Anvendes skalaen, som den var tænkt af Karakterkommissionen, og er kommissionens krav til en moderne karakterskala fortsat aktuelle?"

- Er international anvendelighed stadig centralt?
- Er målopfyldelse stadig centralt?
- Er en meget klar trindifferentiering stadig centralt?

Spørgsmålet om målopfyldelse behandles i dette kapitel, mens spørgsmålet om den internationale anvendelighed behandles i kapitel 7, og spørgsmålet om den klare trindifferentiering behandles i kapitel 5.

Kapitlet bygger på:

- Spørgeskemaundersøgelse blandt bedømmere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser.
- Spørgeskemaundersøgelse blandt ledere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser samt censorformænd på de videregående uddannelser
- Fokusgruppeinterviews og enkeltinterviews med bedømmere på grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser
- Enkeltinterviews med opgavekommissionsformænd og lærings- og fagkonsulenter på grundskoleområdet og de gymnasiale uddannelser.

3.2 Forståelsen og anvendelsen af skalaen

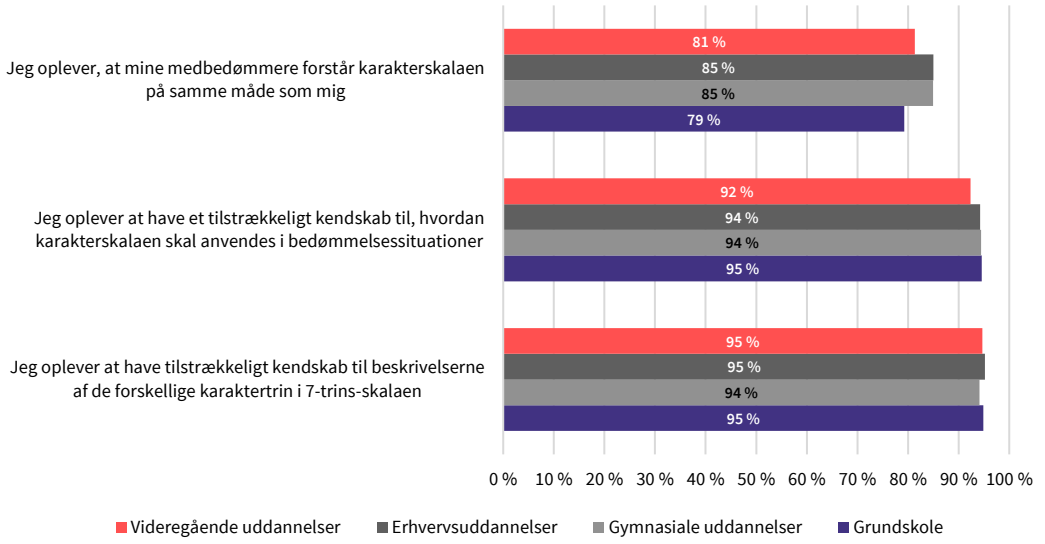
Helt overordnet er det værd at bemærke, at bedømmerne på tværs af grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser er meget enige med hensyn til deres vurdering af 7-trins-skalaen og deres forståelse og anvendelse af den. Fx viser spørgeskemaundersøgelsen (figur 3.1), at bedømmerne på tværs af områderne oplever, at:

- De har et tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin på 7-trins-skalaen (94-95 % af bedømmerne er meget enige eller enige heri)

- De har et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer (92-95 % af bedømmerne er meget enige eller enige heri)
- Deres medbedømmere forstår karakterskalaen på samme måde som dem (79-85 % af bedømmerne er meget enige eller enige heri).

FIGUR 3.1

Forståelse af skalaen blandt bedømmere, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

3.2.1 Bedømmerne anvender som udgangspunkt karakterskalaen absolut

Bedømmerne udtrykker i fokusgrupperne på tværs af *grundskoleområdet*, *de gymnasiale uddannelser* og *de videregående uddannelser*, at de er klar over, at de alene skal anvende skalaen absolut og ikke skal tilstræbe en fast fordeling af karaktererne. Dette fremgår fx af følgende citater:

> Nogle tænker måske, at man ikke kan give flere 12-taller, når man har givet det fjerde 12-tal. Men det kan man jo ikke tillade sig. At undgå at give 12 til dig, bare fordi du var den femte i rækken. Det er ikke retfærdigt.
>
> Bedømmer, grundskolen

> Jeg opfatter karakterskalaen som absolut. Men jeg kan da godt huske, da karakterskalaen kom, så var det meldingen fra Undervisningsministeriet, at man forventede, at 10 % skulle have 12. Jo jo, hvis eleverne kan gøre det. Jeg har altid oplevet det som en forventning til skalaens kunnen snarere end en ordre til mig om, at nu skal jeg give 10 % af eleverne 12. Jeg bruger absolutte karakterer. De kan alle sammen få -3 og alle sammen få 12.
>
> Bedømmer, gymnasial uddannelse

På *grundskoleområdet* peger nogle bedømmere dog på, dels at de møder andre bedømmere, som stadig har den opfattelse, at man skal tilstræbe en bestemt fordeling af karakterer, dels at karakterskalaens relative element indirekte har betydning for deres karaktergivning, idet de oplever, at en del af bedømmelsesværktøjerne udformes ud fra ønsket om at ramme den intenderede fordeling. En bedømmer siger herom:

> Det er jo den [den intenderede fordeling], man laver vejledningerne og omsætningstabellerne ud fra. I matematik får vi jo den der fordelingstabel, efter at de har rettet nogle hundrede stykker.
>
> Bedømmer, grundskolen

Lederne og censorformændene er i spørgeskemaundersøgelsen blevet bedt om at vælge tre styrker ved 7-trins-skalaen ud af ni mulige. I tabel 4.1 er der gengivet de tre styrker, som er valgt af flest ledere. "At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden" er i toptre for alle områder. 46 % af lederne på grundskoleområdet har valgt det som en styrke. Det samme gælder 57 % af lederne på erhvervsuddannelserne, 61 % af lederne på de gymnasiale uddannelser, 62 % af lederne på de videregående uddannelser og 86 % af censorformændene. Dette tyder på, at Karakterkommissionens krav om, at bedømmelse skal være absolut frem for relativ, stadig er aktuelt.

3.2.2 Bedømmernes perspektiver på trinbeskrivelsen

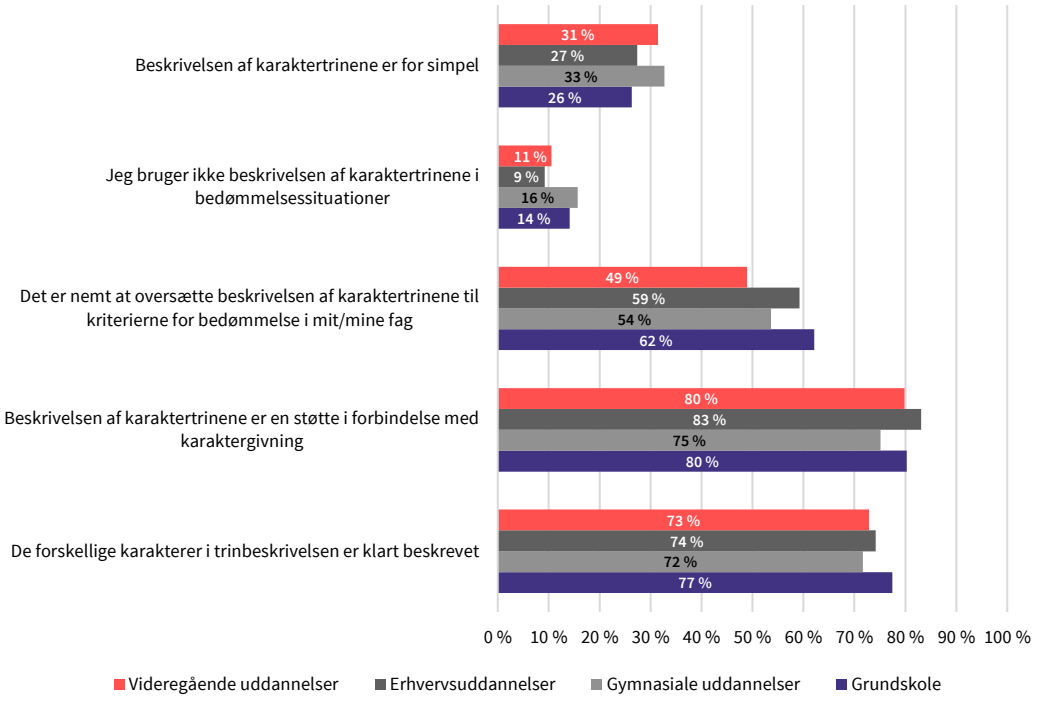
Bedømmerne er på tværs af alle uddannelsesområder overvejende enige i, at de forskellige karakterer er klart beskrevet i trinbeskrivelsen (72-77 %), og at beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning (75-83 %), jf. figur 3.2.

Omkring halvdelen eller flere af bedømmerne på tværs af alle uddannelsesområder (fra 49 % (videregående uddannelser) til 62 % (grundskole)) er desuden enige i, at det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i deres fag.

Omvendt er kun en mindre del af bedømmerne på tværs af uddannelsesområderne enige i, at de ikke bruger beskrivelsen af karaktertrinene i bedømmelsessituationer (9-16 %), mens 26-33 % er enige i, at beskrivelsen af karaktertrinene er for simpel, jf. figur 3.2.

FIGUR 3.2

Bedømmernes perspektiver på beskrivelsen af karaktertrinene, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

Trinbeskrivelse anvendes særligt ved tvivl eller uenighed

Et perspektiv i interviewene med bedømmerne er, at trinbeskrivelsen anvendes meget aktivt. Trinbeskrivelsen anvendes særligt ved usikkerhed eller uenighed om, hvilken karakter en given elev/studerende skal have. En bedømmer fortæller fx, at han tager trinbeskrivelsen frem og læser beskrivelserne op, således at han og censor for hvert læringsmål kan vurdere, om der fx er tale om en del mangler og dermed karakteren 7 eller adskillige væsentlige mangler og dermed karakteren 4. To andre bedømmere fortæller følgende om deres brug af trinbeskrivelsen:

> Jeg gør det, at jeg sætter mig med opgaven og samtidig sidder og kigger på opgaveformuleringen, de har fået, for at se, om de opfylder de krav, der er. Så kigger jeg på den ramme, der er i vores studieordning. Hvis jeg er i tvivl, så går jeg ind og kigger på, om de opfylder de ting, de skal på det her niveau. Og til sidst kigger jeg så på selve 7-trins-skalaens formuleringer. Og giver karakterer derudfra.
>
> Bedømmer, videregående uddannelse

> Denne her [trinbeskrivelsen] har jeg liggende. Den bruger jeg tit. Vi har ikke brugt den i dag, men når vi er mere uenige, så synes jeg i virkeligheden, det her tit er klargørende. Den trækker jeg relativt meget på. Den har jeg altid med, når jeg giver karakterer.
>
> Bedømmer, videregående uddannelse

Delte meninger om trinbeskrivelsens fokus på fejl og mangler

På tværs af alle områderne har bedømmerne forskellige holdninger til, om trinbeskrivelserne har for meget fokus på mangler, jf. figur 3.3:

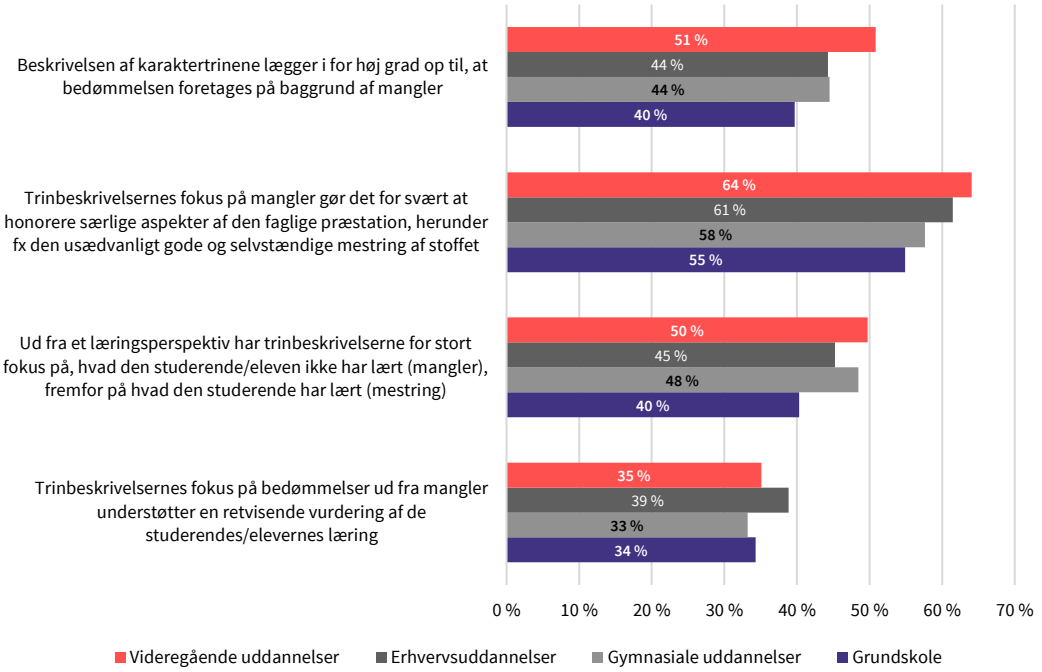
- Fra 55 % (grundskole) til 64 % (videregående uddannelser) af bedømmerne er enige i, at trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation, herunder fx den usædvanlig gode og selvstændige mestring af stoffet. Fra 20 % (gymnasiale uddannelser) til 25 % (grundskole) er uenige eller meget uenige.
- Fra 40 % (grundskole) til 50 % (videregående uddannelser) af bedømmerne er enige i, at trinbeskrivelserne ud fra et læringsperspektiv har for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), frem for på, hvad den studerende har lært (mestring). Fra 21 % (grundskole) til 25 % (gymnasiale uddannelser) er uenige eller meget uenige.
- Fra 39 % (grundskole) til 50 % (videregående uddannelser) af bedømmerne er enige i, at beskrivelsen af karaktertrinene i for høj grad lægger op til, at bedømmelsen foretages på baggrund af mangler. Fra 24 % (erhvervsuddannelserne) til 31 % (grundskole) er uenige eller meget uenige.
- Fra 40 % (gymnasiale uddannelser) til 51 % (videregående) af bedømmerne er enige i, at trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring. Fra 24 % (erhvervsuddannelser) til 33 % (videregående uddannelser) er uenige eller meget uenige.

Spørgeskemaundersøgelsen og fokusgrupperne med bedømmere på de videregående uddannelser tegner et billede af, at undervisere fra de naturvidenskabelige og tekniske uddannelser i mindre grad end undervisere fra de humanistiske og samfundsfaglige uddannelser mener, at fokus på mangler er uhensigtsmæssigt. Der er tilsyneladende ikke forskelle at spore på tværs af fag på grundskoleområdet, de gymnasiale uddannelser eller erhvervsuddannelserne.

FIGUR 3.3

Bedømmernes perspektiver på trinbeskrivelsens fokus på fejl og mangler, %

Hvor enig eller uenig er du i følgende udsagn om beskrivelsen af de forskellige trin?



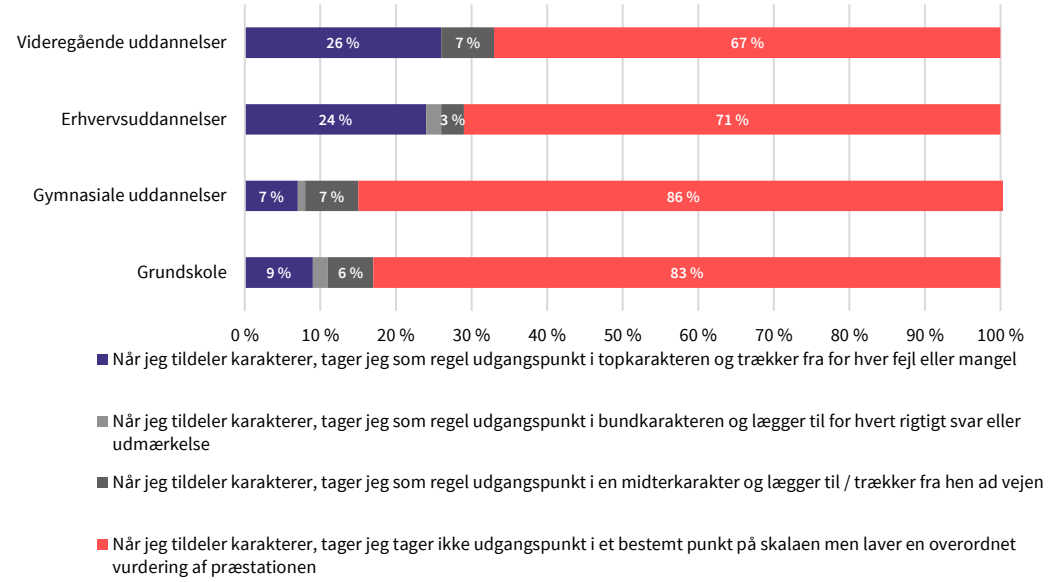
Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

Vi har i spørgeskemaundersøgelsen spurgt bedømmerne, hvordan de anvender skalaen, når de tildeler karakterer, jf. figur 3.4. Figuren viser, at langt størstedelen af bedømmerne (fra 67 % (videregående uddannelser) til 86 % (gymnasiale uddannelser)) ikke tager udgangspunkt i noget bestemt sted på skalaen, når de tildeler karakter ved en mundtlig prøve. Kun 6-26 % tager udgangspunkt i topkarakteren og trækker fra for fejl og mangler. Billedet er det samme, når vi kigger på karaktertildeling ved skriftlige prøver.

FIGUR 3.4

Bedømmernes anvendelse af skalaen ved mundtlige prøver, %

Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en mundtlig prøve?



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N =2.946.

Samme billede går igen i fokusgruppeinterviewene med bedømmerne. Særligt bedømmerne på *grundskoleområdet* og *de gymnasiale uddannelser* beskriver nemlig generelt, at de ikke starter et bestemt sted på skalaen, men i højere grad foretager en helhedsvurdering af elevens præstation. En bedømmer fortæller:

> Jeg oplever det ikke så meget som en jagt på fejl og mangler, men mere en kritisk vurdering af den enkelte elevs præstation. Hvordan man bedømmer, kommer helt an på, hvem der kommer ind ad døren. Man sidder generelt med et åbent sind som eksaminator og censor.
>
> Bedømmer, grundskole

I tråd med resultaterne fra spørgeskemaundersøgelsen, som viser, at en større andel af bedømmerne på de videregående uddannelser tager udgangspunkt i karakteren 12 og trækker fra for fejl og mangler, giver fokusgruppeinterviewene med bedømmerne på *de videregående uddannelser* et indtryk af, at fejl og mangler fylder mere i bedømmernes karaktergivning end på de andre områder. Dette beskriver bedømmerne som drevet af et større fokus på eksamensklager. To bedømmere fortæller herom:

> Hvis jeg nu starter med at eksaminere opad, og vi ender med nogle spørgsmål omkring en 7'er i sværhedsgrad, men jeg når ikke at teste 10 og 12, fordi det er tydeligt, at det kan den studerende ikke svare på, så bliver eksaminationen på 7'er-niveauet. Men så går den studerende ud og siger "jeg kunne svare på det hele, jeg skulle have 12". Hvis jeg ikke kan sige "der var rent faktisk nogle ting, du ikke kunne svare på" [...] Så kan den gå hele vejen og blive omstødt.
>
> Bedømmer, videregående uddannelse

> Jeg gør dem [de studerende] opmærksom på det. Det er en del af mit arbejde. De skal vide, hvad de kan klage over. De kan jo ikke klage over, at de bliver spurgt om ting, de ikke ved noget om. Derfor gør jeg også meget ud af at stille spørgsmål, de ikke kan svare på. Ellers forstår de det ikke, hvis de har svaret på alt.
>
> Bedømmer, videregående uddannelse

Fra et læringsperspektiv opleves det ikke som meningsfuldt at bedømme mangler

På tværs af bedømmerne fra grundskolen, de gymnasiale uddannelser og de videregående uddannelser er det et udbredt perspektiv i fokusgruppeinterviewene, at det ud fra et læringsperspektiv ikke er meningsfuldt at skulle bedømme eleverne/de studerende ud fra mangler. Dels fordi mangler i en læringssituation, ifølge bedømmerne, ikke er entydigt negativt, men også kan rumme et udviklingspotentiale – at man lærer ved at fejle. Dels fordi bedømmerne beskriver, at det, når det drejer sig om at bygge mennesker op, er centralt at bygge på det, de kan, og ikke alene på det, de ikke kan. To bedømmere fortæller herom:

> Jeg vil hellere danne mig et fuldstændigt billede i stedet for at lave fejlfinding. Selvom skalaen godt kan lægge op til det, og det er formuleret i systemet, men jeg bryder mig ikke om den tilgang. Det er en pædagogisk overvejelse, for hvis man bare finder fejl og trækker point fra, så tør eleverne ikke åbne munden, fordi de bliver bange for at sige noget forkert. Jeg vil hellere se på mennesket bagved og støtte dem.
>
> Bedømmer, gymnasial uddannelse

> Mange af mine kollegaer kritiserer, at det er en mangelskala. Det er pædagogisk meget uheldigt, at vi altid fokuserer på deres fejl og mangler. Vi burde fokusere på det, de er lykkedes med.
>
> Bedømmer, videregående uddannelse

Bedømmerne lægger også vægt på, at fokus på mangler opbygger en kultur blandt eleverne/de studerende, hvor deres hovedfokus er at undgå mangler:

> Det helt store problem er det med at finde fejl. Vi skal altid finde fejl. Det er håbløst. Med 13-skalaen kunne vi belønne det exceptionelle. Vi har fremelsket en nulfejlskultur. Det er dårligt for vores samfund. Man skal tage chancer. Før handlede det om selvstændighed, det er væk nu. De skal bare gøre, hvad der bliver sagt.
>
> Bedømmer, videregående uddannelse

Samlet set tegner evalueringen dermed ikke et entydigt billede af, om et fokus på mangler i forbindelse med bedømmelsen er hensigtsmæssigt, og ej heller af, at dette faktisk er bedømmernes fokus.

3.3 Målopfyldelse og bedømmelsesværktøjer

Et udgangspunkt for anvendelsen af skalaen er, at bedømmelserne skulle ske ud fra på forhånd fastlagte faglige mål, som også skulle være tilgængelige for elever/studerende.

Lederne i grundskolen, på de gymnasiale uddannelser, på erhvervsuddannelserne og på de videregående uddannelser samt censorformændene ved de videregående uddannelser er i spørgeskemaundersøgelsen blevet bedt om at vælge tre styrker ved 7-trins-skalaen ud af ni mulige. I tabel 3.1 er der gengivet de tre styrker, som er valgt af flest ledere. Med undtagelse af grundskoleområdet er "At bedømmelser sker ud fra eksplicitte læringsmål" blandt toptre på alle områder. 50 % af lederne i grundskolen og på erhvervsuddannelserne har valgt det som en styrke. Det samme gælder 52 %

af lederne på de gymnasiale uddannelser, 70 % af lederne på de videregående uddannelser og 68 % af censorformændene. Blandt lederne på de videregående uddannelser og censorformændene er det den mest valgte styrke.

TABEL 3.1

Ledernes perspektiv på styrker ved karakterskalaen, toptre fordelt på områder

	1	2	3
Grundskole	"At skalaen er anvendelig internationalt" (81 %)	"At de forskellige karakterer i trinbeskrivelsen er klart beskrevet" (53 %)	"At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden (relativ karaktergivning) " (47 %)
Erhvervsuddannelser	"At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden (relativ karaktergivning)" (57 %)	"At skalaen er anvendelig internationalt" (54 %)	"At bedømmelser sker ud fra eksplicitte læringsmål" (50 %)
Gymnasiale uddannelser	"At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden (relativ karaktergivning)" (60 %)	"At skalaen er anvendelig internationalt" (55 %)	"At bedømmelser sker ud fra eksplicitte læringsmål" (52 %)
Videregående uddannelser	"At bedømmelser sker ud fra eksplicitte læringsmål" (70 %)	"At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden (relativ karaktergivning)" (63 %)	"At skalaen er anvendelig internationalt" (39 %)
Censorformænd	"At bedømmelser sker ud fra eksplicitte læringsmål" (68 %)	"At bedømmelserne er baseret på målopfyldelse (absolut karaktergivning), frem for hvordan de studerende/eleverne klarer sig i forhold til hinanden (relativ karaktergivning)" (64 %)	"At de forskellige karakterer i trinbeskrivelsen er klart beskrevet" (39 %)

Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evaluering af 7-trins-skalaen*.
Note: Tallet i parentes indikerer andelen af ledere, som har angivet en given styrke som første-, anden- eller tredjeprioritet. Grundskole: N = 79, gymnasiale uddannelser: N = 122, erhvervsuddannelser: N = 73, videregående uddannelser: N = 370.

3.3.1 Bedømmernes perspektiv på bedømmelsesværktøjerne på grundskoleområdet, de gymnasiale uddannelser og erhvervsuddannelserne

På grundskoleområdet og de gymnasiale uddannelser fastlægges de faglige mål for de enkelte fag centralt i Undervisningsministeriet. Undervisningsministeriet udarbejder desuden konkrete vurderings-/bedømmelseskriterier for de enkelte prøver, og de skriftlige prøver afholdes som nationale prøver. Derved fastlægges et fælles fagligt niveau centralt i ministeriet. Som supplement til de faglige mål udarbejder Undervisningsministeriet en række andre redskaber til at støtte op om en korrekt og ensartet bedømmelse, herunder prøvevejledninger og karakterbeskrivelser for prøverne/fagene.

På *erhvervsuddannelserne* anvendes en kombination af centrale og lokale bedømmelsesværktøjer. I grundfag er rammerne for faget og prøven beskrevet i den centralt udarbejdede læreplan med faglige mål, kernestof, eksaminationsgrundlag og bedømmelsesgrundlag, men opgaven til prøven stilles af skolen på grundlag af den prøvebeskrivelse, som skolen inden for rammerne af den centralt udarbejdede læreplan har udfoldet i den lokale undervisningsplan. De afsluttende prøver udformes typisk i et samarbejde mellem skolerne og de faglige udvalg. Uddannelsernes slutmål er fastlagt i uddannelsesbekendtgørelserne, og den konkrete tilrettelæggelse af prøven og udformningen af censorvejledning sker i et samarbejde mellem skolerne og de faglige udvalg.

De forskellige bedømmelsesværktøjer fremgår af tabel 3.2.

TABEL 3.2

Oversigt over bedømmelsesværktøjer (ud over karakterbekendtgørelsen)

Niveau	Bedømmelsesværktøjer
Grundskolen	• Fælles Mål • Bekendtgørelse om folkeskolens prøver, herunder vurderingskriterier fordelt på fag • Prøvevejledninger med blandt andet karakterbeskrivelser, der har til formål at præcisere og uddybe de prøvekrav, der stilles i prøvebekendtgørelsen, og at tydeliggøre den sammenhæng, der er mellem prøvebekendtgørelsen, folkeskolens formål, fagområdet og Fælles Mål • De centralt stillede skriftlige opgaver
De gymnasiale uddannelser	• Læreplaner med faglige mål, kernestof og bedømmelseskriterier • Vejledninger til læreplanerne, herunder karakterbeskrivelser, som beskriver, hvad der karakteriserer præstationer til udvalgte karakterer • Bekendtgørelse om prøver og eksamen på de almene og studieforberedende ungdoms- og voksenuddannelser • Vejledning til de centralt stillede skriftlige opgaver
Erhvervsuddannelserne	• Læreplaner for grundfag med faglige mål, kernestof og bedømmelseskriterier • Vejledninger til læreplanerne for grundfag • Fagmøder for grundfagslærere og faglærere • Uddannelsesordninger for de uddannelsesspecifikke fag med faglige mål/niveauer • Niveaubeskrivelse af uddannelsesspecifikke fag – fastlagt i hovedbekendtgørelsen • Lokale undervisningsplaner for fag/undervisningsforløb, herunder prøvebeskrivelser • National standard for grundforløbsprøver – gældende for de enkelte uddannelser • Vejledning til centralt udarbejdede afsluttende prøver for de enkelte uddannelser – udarbejdet af faglige udvalg og skoler • Bekendtgørelse om prøver og eksamen på de erhvervsrettede uddannelser • Særlig censur i udvalgte grundfag • Eksterne censorer fra erhvervslivet ved afsluttende prøver • Årlige udviklingsredegørelser (faglige udvalgs status over uddannelsen, herunder beståelsesprocenter ved afsluttende prøver)

Spørgeskemaundersøgelsen blandt bedømmere på *grundskoleområdet, de gymnasiale uddannelser og erhvervsuddannelserne* viser overordnet, at bedømmerne i varierende grad anvender de centrale bedømmelsesværktøjer, som er til rådighed, ved mundtlige og skriftlige prøver:

- 81 % af bedømmerne på *grundskoleområdet* bruger i meget høj grad eller i høj grad de bedømmelsesværktøjer, der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.), ved skriftlige prøver, og 70 % bruger dem i meget høj grad eller i høj grad ved mundtlige prøver.
- 67 % af bedømmerne på *de gymnasiale uddannelser* bruger i meget høj grad eller i høj grad de bedømmelsesværktøjer, der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.), ved skriftlige prøver, og 60 % bruger dem i meget høj grad eller i høj grad ved mundtlige prøver.
- 64 % af bedømmerne på *erhvervsuddannelser* bruger i meget høj grad eller i høj grad de bedømmelsesværktøjer, der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.), ved skriftlige prøver, og 64 % bruger dem i meget høj grad eller i høj grad ved mundtlige prøver.

Der er også variation med hensyn til bedømmernes vurdering af de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, jf. figur 3.6. 82 % af bedømmerne på grundskoleområdet er meget enige eller enige i, at de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, giver dem et solidt udgangspunkt for bedømmelsen. Det samme gælder 57 % og 60 % af bedømmerne på hhv. de gymnasiale uddannelser og erhvervsuddannelserne.

79 % af bedømmerne på grundskoleområdet er meget enige eller enige i, at de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, er med til at sikre en ensartet bedømmelsespraksis. Det gælder 61 % af bedømmerne på de gymnasiale uddannelser og 54 % af bedømmerne på erhvervsuddannelserne.

60 % af bedømmerne på grundskoleområdet er desuden meget enige eller enige i, at de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, sikrer sammenhæng mellem Fælles Mål/faglige mål/målbeskrivelser og deres bedømmelse. Det samme gælder 57 % og 51 % af bedømmerne på hhv. de gymnasiale uddannelser og erhvervsuddannelserne.

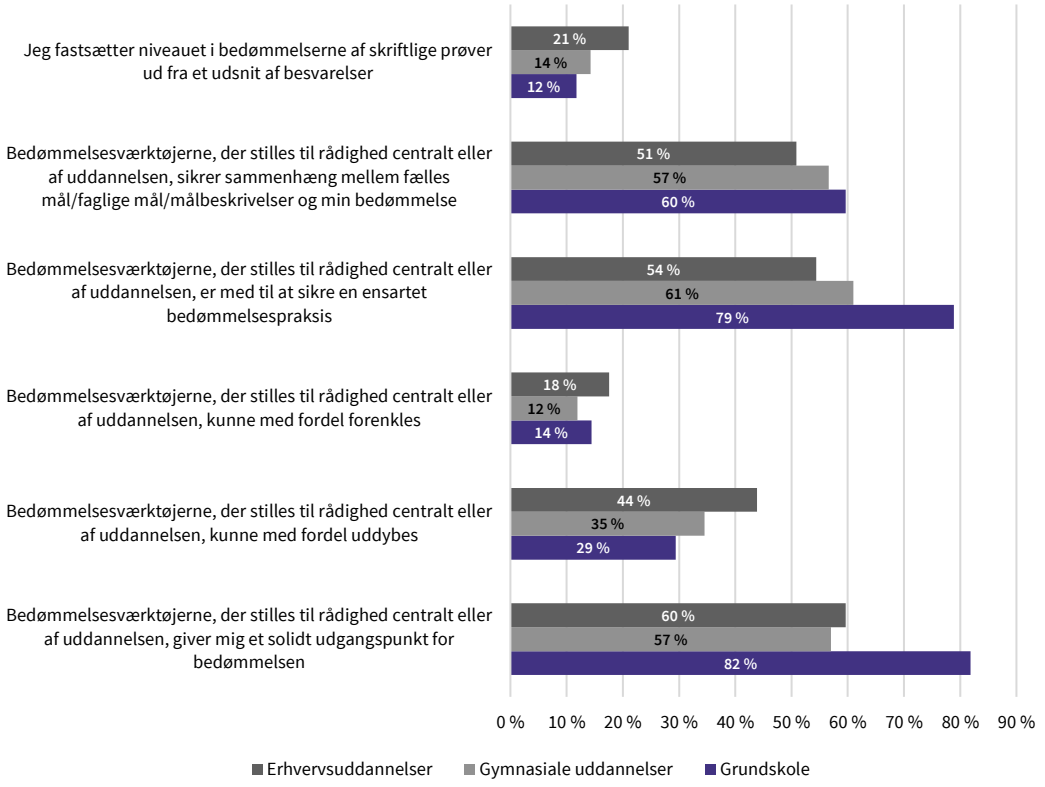
Der er ikke et udbredt ønske om, at bedømmelsesværktøjerne skal forenkles (12-18 % af bedømmerne er enige eller meget enige), men der tegner sig et billede af, at en gruppe bedømmere (29-44 %) gerne ser, at de bedømmelsesværktøjer, der stilles til rådighed centralt eller af uddannelsen, uddybes. Fx efterspørger nogle bedømmere på de gymnasiale uddannelser i fokusgrupperne, at de fagspecifikke karakterbeskrivelser udvides til at omfatte hvert trin på skalaen frem for kun top-, midt- og bundpræstationen.

Et lignende billede tegner sig for bedømmelsesværktøjerne til de skriftlige prøver, jf. bilagsrapporten *Bedømmernes perspektiver – Bilagsrapport til evaluering af 7-trins-skalaen*.

FIGUR 3.5

Bedømmernes vurdering af bedømmelsesværktøjer, fordelt på område, %

Hvor enig eller uenig er du i følgende udsagn om bedømmelsesværktøjer ved skriftlige prøver?



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en skriftlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018. Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 374, gymnasiale uddannelser: N = 721, erhvervsuddannelser: N = 57.

Bedømmelsesværktøjerne anvendes særligt ved tvivl eller uenighed

Bedømmerne på *grundskoleområdet* og *de gymnasiale uddannelser* beskriver, at bedømmelsesværktøjerne særligt anvendes, hvis der opstår tvivl eller uenighed med hensyn til bedømmelsen af en elev/studerende, og at bedømmelsesværktøjerne derfor er med til at understøtte en ensartethed i bedømmelsen. En bedømmer udtrykker det således:

> Den generelle trinbeskrivelse er god at have liggende, hvis man er lidt i tvivl. Hvis censor har spillet noget andet ud end en selv, så kan man finde gode argumenter i den. Så jeg bruger den meget.
>
> Bedømmer, gymnasial uddannelse

Bedømmerne peger desuden på, at bedømmelsesværktøjerne er vigtige, når det handler om at skabe ensartethed i bedømmelse og tolkning af målene, frem for at risikere, at bedømmernes individuelle kæpheste inden for faget kan blive styrende.

3.3.2 Bedømmelsesværktøjer på de videregående uddannelser

På de videregående uddannelser har bedømmerne ud over trinbeskrivelsen primært læringsmålene og evt. andre lokalt udarbejdede bedømmelsesredskaber til rådighed som værktøjer ved bedømmelserne.

Bedømmerne er ikke blevet spurgt direkte om, hvorvidt de anser målopfyldelse som en central del af bedømmelsen. Bedømmerne på de videregående uddannelser er dog blevet spurgt om, hvorvidt de generelt mener, at de faglige mål er formuleret tydeligt og anvendeligt med hensyn til brug i bedømmelsessituationen. Det er 44 % af bedømmerne enige eller meget enige i, at de er. Andelen, der er enige i udsagnet, er statistisk signifikant forskellig mellem bedømmere fra professionsbacheloruddannelserne (40 %) og bedømmere fra erhvervsakademiuddannelserne (49 %). Der er med andre ord ikke noget, der indikerer, at der her er forskel på, om målene er formuleret centralt eller decentralt.

Spørgeskemaundersøgelsen blandt bedømmere på de videregående uddannelser tegner desuden et noget broget billede af sammenhængen mellem bedømmelse og læringsmål, jf. figur 3.7. Figuren viser, at:

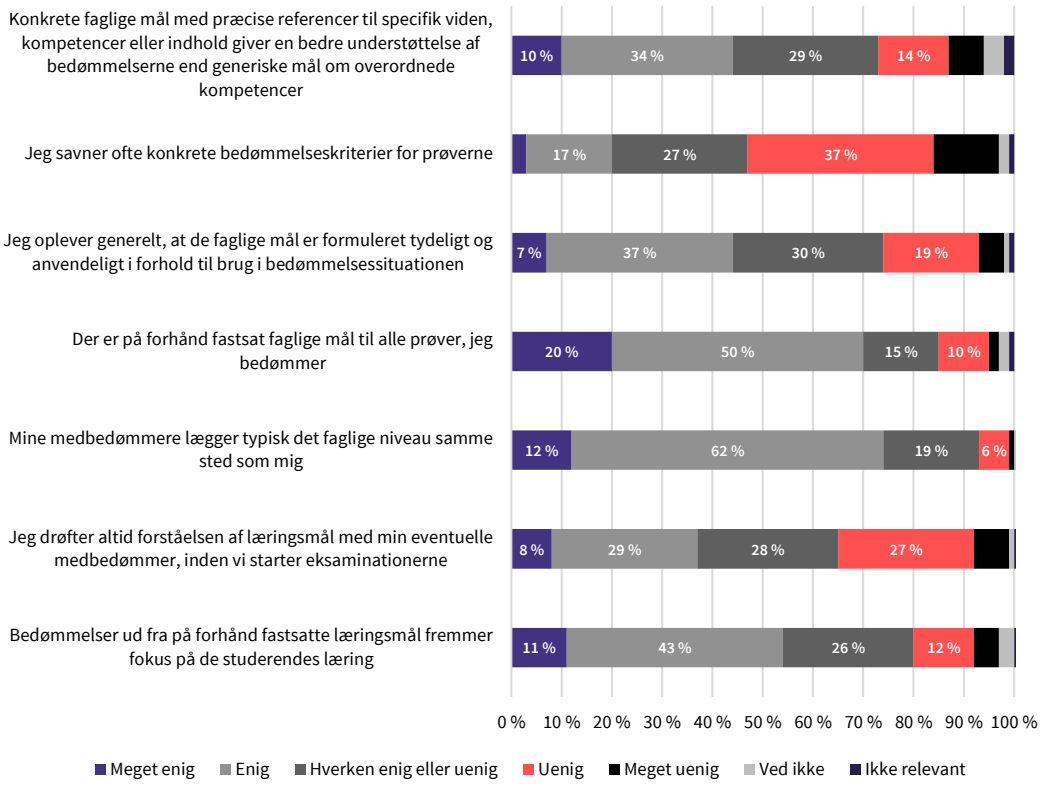
- 74 % er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig". Omvendt er 7 % uenige.
- 70 % er enige i udsagnet "Der er på forhånd fastlagt faglige mål til alle prøver, jeg bedømmer". Omvendt er 12 % uenige.
- 54 % er enige i udsagnet "Bedømmelser ud fra på forhånd fastlagte læringsmål fremmer fokus på de studerendes læring". Omvendt er 17 % uenige.
- 44 % er enige i udsagnet "Jeg oplever generelt, at de faglige mål er formuleret tydeligt og anvendeligt med hensyn til brug i bedømmelsessituationen". Omvendt er 21 % uenige.
- 44 % er enige i udsagnet "Konkrete faglige mål med præcise referencer til specifik viden, kompetencer eller indhold giver en bedre understøttelse af bedømmelserne end generiske mål om overordnede kompetencer". Omvendt er 21 % uenige.
- 37 % er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne". Omvendt er 34 % uenige.

Ser vi nærmere på udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne", er der statistisk signifikante forskelle mellem bedømmere fra forskellige sektorer. 46 % af bedømmerne på erhvervsakademiuddannelserne er enige i udsagnet, mens dette gælder 36 % af bedømmerne på professionsbacheloruddannelserne og 34 % af bedømmerne på universitetsuddannelserne.

FIGUR 3.6

Bedømmernes vurdering af sammenhængen mellem læringsmål og bedømmelse, %

Hvor enig eller uenig er du i følgende spørgsmål om sammenhængen mellem læringsmål og bedømmelse?



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: N = 2.946.

Spørger man censorformændene, er 89 % meget enige eller enige i, at der er opstillet læringsmål for alle faglige prøver på de videregående uddannelser. 74 % er meget enige eller enige i, at de faglige læringsmål er formuleret, så de understøtter bedømmelser ud fra målopfyldelse, mens 20 % er meget enige eller enige i, at der er stor forskel på kvaliteten af de faglige læringsmål inden for respondentens censorområde.

I fokusgrupperne er et gennemgående tema, at det ikke altid er let at anvende læringsmålene i en bedømmelsessituation, hvilket illustreres i de følgende citater:

> Bedømmelsen skulle jo gerne ske på baggrund af målbeskrivelserne. Dilemmaet er, at målbeskrivelserne ikke altid er så præcise. Man vil gerne have en vis åbenhed i delmålene og samtidig en konkret beskrivelse, som gør det muligt at bedømme præcist. Vores studieordning er relativt ny, så vi redigerer stadig løbende. Målene bruges i bedømmelsesprocessen, ved at man kigger på de overordnede mål og de delmål, vi selv har formuleret.
>
> Bedømmer, videregående uddannelse

> Problemet med læringsbeskrivelser er, at de er så diffuse. Det kan være svært at vide, hvor man skal skyde på dem [de studerende] for at afprøve målene.
>
> Bedømmer, videregående uddannelse

> Jeg oplevede i starten, at læringsmålene var for abstrakte. Både i forhold til karaktergivning og i forhold til, hvor meget vi kræver af de studerende. Jo mere abstrakte, des mere kan man læse i en uendelighed. Det var problemet de første år, fordi læringsmålene var for generelle. Nu er de så blevet ultraspecifikke.
>
> Bedømmer, videregående uddannelse

> Der var nogle børnesygdomme [ved læringsmålene]. Vi lavede nogle mere runde formuleringer, og nu er det blevet vores bedste ven i forhold til eksamensklager. I forhold til skalaen gør det det faktisk ekstremt nemt at arbejde med. Hvis man har lavet gode beskrivelser.
>
> Bedømmer, videregående uddannelse

Som citaterne illustrerer, bliver læringsmålene mange steder revideret ofte, hvilket kan gøre det svært i en bedømmelsessituation, hvis man hvert år skal bedømme efter nye mål, især hvis bedømmer og censor ikke drøfter målene inden udprøvningen. Desuden giver fokusgrupperne og de individuelle interviews et indtryk af, at der er forskellige hensyn på spil. I en bedømmelsessituation skal målene helst være så konkrete som muligt, men for ikke at revidere dem hele tiden eller låse undervisningen for meget skal målene helst være på et vist abstraktionsniveau.

Bedømmerne oplever ikke behov for supplerende bedømmelsesværktøjer

Bedømmerne på de videregående uddannelser oplever generelt ikke, at der er behov for supplerende bedømmelsesværktøjer. Blot 24 % svarer ja med hensyn til dette i forbindelse med afsluttende opgaver (specialer, bachelorprojekter, afsluttende projekter mv.), 20 % svarer ja, når det gælder mundtlige prøver, 18 % svarer ja i forbindelse med skriftlige prøver, og 17 % svarer ja med hensyn til projektopgaver.

3.3.3 Understøttelsen af bedømmerne

Evalueringen viser, at understøttelsen af bedømmerne primært er individuel, og at der ikke nødvendigvis er en fast systematik praksis med hensyn til det. Det er således især en-til-en-oplæring eller sparring fra kollegaer, overværelse af kollegas prøveafholdelse og afholdelse af de første prøver sammen med erfaren censor/eksaminator, som lederne tilbyder nye bedømmere, mens kurser eller seminarer er mindre udbredte, jf. figur 3.8. Billedet er det samme, hvis vi ser på, hvad lederne tilbyder erfarne bedømmere. Her er det også primært sparring mellem bedømmere i teams/grupper/par og overværelse af kollegaers prøveafholdelse, som lederne initierer (se bilagsrapporten *Ledelsernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen*). Sidemandsoplæring har mange fordele, men også en indbygget risiko for, at misforståelser og kultur med hensyn til, hvordan skalaen skal anvendes, og lignende videreføres fra én bedømmer til en anden.

Indtrykket fra fokusgrupperne med ledere på de videregående uddannelser er i tråd med resultaterne fra spørgeskemaundersøgelsen. En leder fortæller fx om uddannelsens tilgang til oplæring:

> Vi bruger meget en-til-en i vores understøttelse af bedømmerne. Vores erfaring er, at det skal være meget konkret, før det bliver meningsfyldt for bedømmerne, i hvert fald for de nye bedømmere. Det skal være en konkret opgave, vi gennemgår sammen og så siger "hvad trækker op, hvad trækker ned, hvad lægger vi vægt på i præcis den her besvarelse?". Når vi gør det mere generelt, samler større grupper, bliver det meget diffust, hvad der ligger til grund for niveaudelingen.
>
> Leder, professionsbacheloruddannelse

Samtidig fremhæver lederne, at bedømmerne tager bedømmelse meget alvorligt:

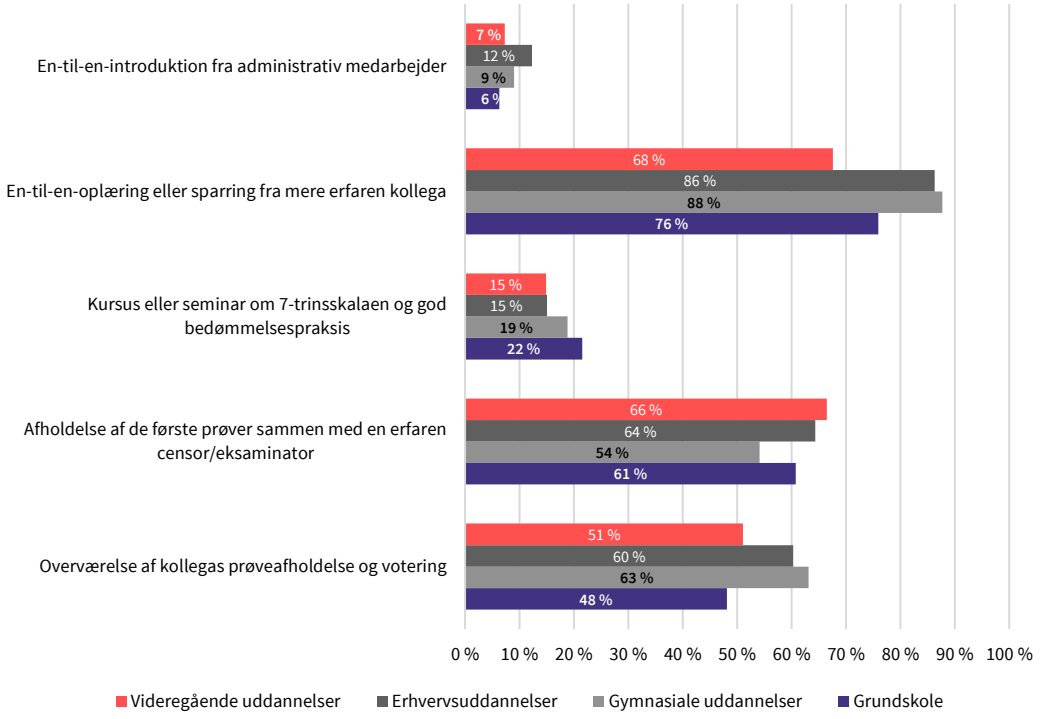
> Det er noget af det, de nye spørger mest om. "Hvordan sørger jeg for ikke at knolde eksamen op?" De er meget optagede af at læse op på regler, spørge de mere erfarne. Det er noget, der er meget opmærksomhed på. Jeg behøver ikke løbe hen ad gangen og spørge de nye "har du nu husket at sætte dig ind i eksamen?". Det er vigtigt og også sårbart, så derfor vil underviserne gerne være klædt på.
>
> Leder, kunstnerisk uddannelse

Indtrykket fra fokusgrupperne med ledere på de videregående uddannelser er dog også, at det er meget forskelligt, hvor systematisk praksis med hensyn til at klæde bedømmere på er. På nogle uddannelsesinstitutioner er det en del af adjunktforløbet eller et obligatorisk kursus for alle nyansatte. Fx fortæller en leder, at man har et obligatorisk kursus i eksamensafholdelse for alle nyansatte undervisere, hvilket blandt andet er foranlediget af, at institutionen har mange undervisere fra udlandet.

FIGUR 3.7

Ledernes initiativer med hensyn til at oplære nye bedømmere, %

Hvilke af følgende aktiviteter til at oplære nye bedømmere i karaktergivning, brug og forståelse af karakterskalaen benytter I på din skole/institution/uddannelse?



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af ledere, som har angivet, at de benytter en given aktivitet. Grundskole: N = 79, gymnasiale uddannelser: N = 122, erhvervsuddannelser: N = 73, videregående uddannelser N = 370.

Hvad angår censorerne på grundskoleområdet, viser interviews med Undervisningsministeriets læringskonsulenter, at der umiddelbart er mere systematik med hensyn til at klæde de beskikkede censorer på sammenlignet med de lokale censorer. Nogle kommuner prioriterer at sende udvalgte bedømmere på kurser i prøveformerne. Andre steder er der den praksis, at de beskikkede censorer på skolerne støtter og vejleder de lokale bedømmere.

Censorformændene på det videregående uddannelsesområde sikrer på forskellig vis, at censorerne bedømmer ensartet på tværs af institutioner og beslægtede uddannelser. Her er det ligeledes meget individualiseret i form af vejledende materiale om god censorskik mv., der sendes ud til censorerne, eller drøftelse af det faglige niveau i bedømmelserne. Deciderede kurser for nye censorer afholdes af 21 % af censorformændene, mens 14 % afholder kurser for erfarne censorer. Uddannelses- og Forskningsministeriet arbejder dog på et fælles censorkursus for alle censorer på det videregående uddannelsesområde.

Bedømmerne oplever ikke behov for mere understøttelse

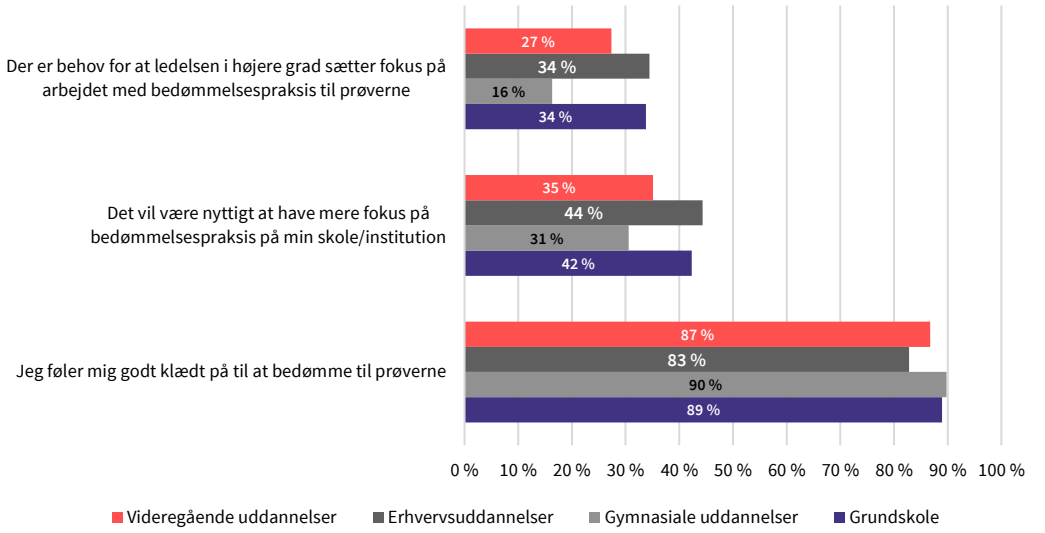
Spørgeskemaundersøgelsen blandt bedømmere peger dog på tværs af uddannelsesområderne ikke på noget større behov for mere understøttelse, jf. figur 3.9. Størstedelen af bedømmerne føler sig godt klædt på til at bedømme prøverne. Dog er 31-44 % af bedømmerne enige i, at det vil være nyttigt at have mere fokus på bedømmelsespraksis på deres skole/institution, og 16-34 % er enige

i, at der er behov for, at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne. Fokusgrupperne giver indtryk af, at bedømmerne særligt oplever, at der kan være behov for en større systematik med hensyn til at klæde nye bedømmere på til karaktergivning, mens de erfarne bedømmere oplever, at netop erfaringen klæder dem godt på. En risiko ved denne tilgang kan dog som beskrevet være, at der ikke bliver rettet op på misforståelser og problematisk kultur i forbindelse med karaktergivningen.

FIGUR 3.8

Bedømmernes perspektiver på oplæring, %

Hvor enig eller uenig er du i følgende udsagn?



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

3.4 Kvalitetssikringen af karaktergivningen (efter karaktergivningen)

Selvom 7-trins-skalaen er en absolut skala med karaktergivning ud fra målopfyldelse, var det Karakterkommissionens forventning, at de bestående karakterer over tid og for store populationer ville sprede sig nogenlunde normalfordelt. Det relative element i karakterskalaen betyder, at karaktergivningen over tid og for store populationer bør følge karakterfordelingen fra ECTS-skalaen. Det gælder ned på de enkelte fag på grundskoleområdet og de gymnasiale uddannelser (fx danskfaget på de gymnasiale uddannelser på tværs af uddannelsesretninger), men er ikke nærmere specificeret på de videregående uddannelser og erhvervsuddannelserne.

På *grundskoleområdet og de gymnasiale uddannelser* er kvalitetssikringen todelt. Dels er der intern kvalitetssikring på den enkelte skole/institution, herunder med hensyn til, hvordan skolen/institutionen og skolens ledelse arbejder med at sikre kvalitet og ensartethed i karaktergivningen, dels er

der central kvalitetssikring af karaktergivningen, herunder hvordan Undervisningsministeriets opgavekommissioner og fagkonsulenter arbejder med kvalitetssikringen, når de producerer og kvalitetssikrer skriftlige prøveopgaver og foretager forcensur i forbindelse med prøveafholdelsen.

På *erhvervsuddannelserne* er der ekstern censur ved alle prøver. Undervisningsministeriet beskikker særlige censorer i udvalgte grundfag med henblik på stikprøvevis at kvalitetssikre prøveafholdelsen på erhvervsskolerne. De særlige censorer afrapporterer til fagkonsulenterne. Undervisningsministeriet afholder desuden et årligt censormøde, hvor resultatet af den særlige censur drøftes. De afsluttende prøver gennemføres ligeledes med ekstern censur, oftest med censorer fra erhvervslivet. Prøverne gennemføres typisk på baggrund af detaljerede prøvebeskrivelser og censorvejledninger med tydelige bedømmelseskriterier.

På det videregående uddannelsesområde foregår kvalitetssikringen primært lokalt på de enkelte institutioner, og det er meget forskelligt, hvor faste rammer der lovgivningsmæssigt er for, hvordan denne kvalitetssikring skal varetages. Derudover har censorformændene en rolle, idet de som led i kvalitetssikringen af uddannelserne blandt andet skal afgive en årlig beretning og medvirke til en løbende dialog om udviklingen i uddannelsen, faget eller fagområdet ved mindst hvert andet år at afholde censormøder med censorerne i censorkorpset og kontaktmøder for uddannelsesinstitutionerne og censorerne, jf. eksamensbekendtgørelserne på det videregående uddannelsesområde.

3.4.1 Den centrale kvalitetssikring på grundskoleområdet og de gymnasiale uddannelser

Kvalitetssikring af karaktergivningen på *grundskoleområdet og de gymnasiale uddannelser* kan foregå både før og efter karakterafgivningen. Kvalitetssikringen kan dreje sig dels om at klæde bedømmerne på til en ensartet bedømmelse, når de afgiver karakterer, dels om at følge op på karakterfordeling og prøver efter prøve og karaktergivning.

På *grundskoleområdet* skelnes der mellem statsligt beskikkede censorer og lokale censorer. Alle skriftlige censorer er statsligt beskikkede, mens ca. 400 af de mundtlige beskikkede censorer er statsligt beskikkede. De lokale censorer er udpeget af skolelederen og indgår typisk som en del af de kommunale censorringe.

På *gymnasieområdet* er alle censorer statsligt beskikkede.

Kvalitetssikringsaktørerne på grundskoleområdet og de gymnasiale uddannelser beskriver, at der fra Undervisningsministeriets side, herunder læringskonsulenterne i grundskolen og fagkonsulenterne i de gymnasiale uddannelser, gøres en lang række tiltag, som er henvendt direkte til de beskikkede skriftlige censorer og i mindre omfang til de beskikkede mundtlige censorer, for at sikre en ensartet bedømmelse. Der er ingen tiltag fra centralt hold, som er henvendt direkte til de lokale censorer på grundskoleområdet med hensyn til at sikre en ensartet bedømmelse.

Evalueringen har ikke haft fokus på den centrale kvalitetssikring på erhvervsuddannelserne.

Bedømmelsesværktøjer til understøttelse af karaktergivningen ved prøver

På *grundskoleområdet* og *de gymnasiale uddannelser* udarbejder Undervisningsministeriet en række bedømmelsesværktøjer til de skriftlige og mundtlige prøver til understøttelse af karaktergivningen. Blandt de centrale bedømmelsesværktøjer er prøvevejledningerne i grundskolens enkelte fag og vejledningerne til de gymnasiale læreplaner. Disse indeholder vejledende karakterbeskrivelser og vurderingskriterier. Derudover bliver der udarbejdet rettevejledninger og omsætningstabeller til de skriftlige prøver, hvor det er relevant.

På *erhvervsuddannelserne* anvendes en kombination af centrale og lokale bedømmelsesværktøjer. I grundfag er rammerne for faget og prøven beskrevet i den centralt udarbejdede læreplan med faglige mål, kernestof, eksaminationsgrundlag og bedømmelsesgrundlag, men opgaven til prøven stilles af skolen på grundlag af den prøvebeskrivelse, som skolen inden for rammerne af den centralt udarbejdede læreplan har udfoldet i den lokale undervisningsplan. De afsluttende prøver udformes typisk i et samarbejde mellem skolerne og de faglige udvalg. Uddannelsernes slutmål er fastlagt i uddannelsesbekendtgørelserne, og den konkrete tilrettelæggelse af prøven og udformningen af censorvejledning sker i et samarbejde mellem skolerne og de faglige udvalg.

Aktiviteter for at sikre en ensartet bedømmelse

På *grundskoleområdet* iværksætter Undervisningsministeriet følgende aktiviteter for beskikkede censorer for at sikre ensartet bedømmelse:

- Censormøder for alle beskikkede censorer i forbindelse med mundtlige og skriftlige prøver med fokus på vurderingskriterier, bedømmelse af konkrete besvarelser mv.
- Ekstern forcensur af skriftlige prøver i matematik, dansk læsning og retskrivning, engelsk, fysik/kemi, biologi og geografi.
- Evalueringsmøde for opgavekommissioner og læringskonsulenter i forbindelse med alle skriftlige prøver før endelig karaktergivning med opfølgende webinar for beskikkede censorer
- Rettegrupper og SkoleKom-konferencer for beskikkede censorer.

På *de gymnasiale uddannelser* gennemføres der ligeledes en række initiativer for lærere, herunder censorer, for at sikre en ensartet bedømmelse:

- Praksis vedrørende karaktergivning er en central del af pædagogikumforløbet, herunder kurset i fagdidaktik, for nye gymnasielærere.
- Fagkonsulenterne afholder regionalmøder/dialogmøder for lærerne i deres fag, og i den forbindelse drøftes forhold med hensyn til karaktergivning ofte. Derudover vurderer fagkonsulenterne løbende, om der er behov for at afholde censorkurser for de skriftlige censorer.
- På kurserne Fagudvikling i Praksis (FIP), som fagkonsulenterne årligt afholder for lærerne i hvert enkelt af gymnasiets fag, tages aktuelle emner op, herunder hvis det vurderes relevant emner som evaluering og karaktergivning.
- Fagkonsulenterne afholder et møde med de skriftlige censorer i hvert enkelt fag, forud for at censorerne fastsætter karaktererne. Her skabes der konsensus blandt censorerne om, hvordan besvarelser af de konkret stillede opgaver skal bedømmes. Udgangspunktet for dette møde er, at fagkonsulenten har indhentet oplysninger fra censorerne om udfaldet af årets besvarelser og censorernes kommentarer til de stillede opgaver.
- Censorernes tilbagemeldinger til fagkonsulenterne danner også grundlag for den evaluering af opgavesættene, som gennemføres hvert år. Denne evaluering er dels en hjælp til opgavekommissionerne i forbindelse med udarbejdelse af kommende opgavesæt, dels en hjælp til lærerne, der kan se eksempler på anvendelse af vurderingskriterierne.

Der er fokus på kvalitetssikring af prøvekarakterer frem for årskarakterer

Kvalitetssikringsaktørerne på *grundskoleområdet* fortæller, at der som udgangspunkt ikke gæres noget fra centralt hold for at sikre ensartet fastsættelse af årskarakterer. Årskaraktererne skal gives ud fra karakterskalaen og Fælles Mål, men der er ikke udarbejdet yderligere bedømmelsesværktøjer eller taget initiativer til på anden vis at kvalitetssikre afgivningen af årskarakterer. I matematik

forsøger Undervisningsministeriet sig dog med et lokalt kursus til to kommuner om fastsættelse af årskarakterer.

Kvalitetssikringsaktørerne påpeger, at der kunne være mere systematisk fokus på afgivning af årskarakterer, og de oplever, at nogle lærere ikke er tilstrækkeligt klædt på til at give årskarakterer. To kvalitetssikringsaktører siger:

> Der bliver ikke gjort noget fra centralt hold i forhold til årskarakterer. Det er sårbart. Det er ikke altid, at faglæreren har styr på, hvad der skal gives karakter efter. Vi ser tit, at der bliver givet karakter efter adfærd, fremtoning og deltagelse i undervisningen i stedet for opfyldelse af fagets mål. Der er begyndt at blive opmærksomhed på dette på skolerne.
> Kvalitetssikringsaktør, grundskoleområdet

> Vi ser, at der ikke altid bliver givet karakterer efter, hvad eleverne kan rent fagligt, men pædagogiske karakterer for eksempelvis at række hånden op.
> Kvalitetssikringsaktør, grundskoleområdet

Spørgeskemaundersøgelsen viser da også, at elevens deltagelse og opførsel indgår som et element hos nogle bedømmere, selvom standpunkts- og årskarakterer skal udtrykke elevens faglige standpunkt på et givent tidspunkt. Fx angiver 59 % af bedømmerne på grundskoleområdet og 49 % af bedømmerne på de gymnasiale uddannelser, at vurdering af elevernes deltagelse i meget høj grad eller i høj grad indgår. Vurdering af elevernes opførsel indgår i meget høj grad eller i høj grad for 7 % af bedømmerne på grundskoleområdet og 3 % af bedømmerne på de gymnasiale uddannelser.

Opfølgning på karakterfordelinger og prøver

På *grundskoleområdet* og *de gymnasiale uddannelser* indberetter skolerne prøvekaraktererne til Styrelsen for It og Læring. Undervisningsministeriet, herunder læringskonsulenter, fagkonsulenter, opgaveudvalg og opgavekommissioner, samler op på karakterfordelingen til prøverne og på censorernes erfaringer med opgavesættene og med elevernes forståelse og besvarelse heraf. Opsamlingerne på censorernes erfaringer og karakterfordelingerne kan give anledning til at justere opgavernes indhold, sværhedsgrad og formulering. Samtidig kan man justere vejledninger, hvis eksempelvis elementer i prøven og faget viser sig vanskelige at bedømme. Opsamlingen på karakterfordelingerne kan ifølge kvalitetssikringsaktørerne være et vigtigt redskab til at vide, i forbindelse med hvilke typer opgaver der er særlig brug for opmærksomhed.

På *erhvervsuddannelserne* er der ekstern censur ved alle prøver. Undervisningsministeriet beskikker særlige censorer i udvalgte grundfag med henblik på stikprøvevis at kvalitetssikre prøveafholdelsen på erhvervsskolerne. De særlige censorer afrapporterer til ministeriets fagkonsulenter. Undervisningsministeriet afholder desuden et årligt censormøde, hvor resultatet af den særlige censur drøftes.

3.4.2 Den decentrale kvalitetssikring

Lederne er i spørgeskemaundersøgelsen blevet spurgt om, hvorvidt de ser forskellige opgaver som en del af deres arbejde med karaktergivning, jf. figur 3.10. Det fremgår, at:

- Fra 64 % (videregående uddannelser) til 84 % (gymnasiale uddannelser) af lederne i meget høj grad eller i høj grad ser det at understøtte, at bedømmerne bruger karakterskalaen på en korrekt og ensartet måde, som en del af deres arbejde med karaktergivningen

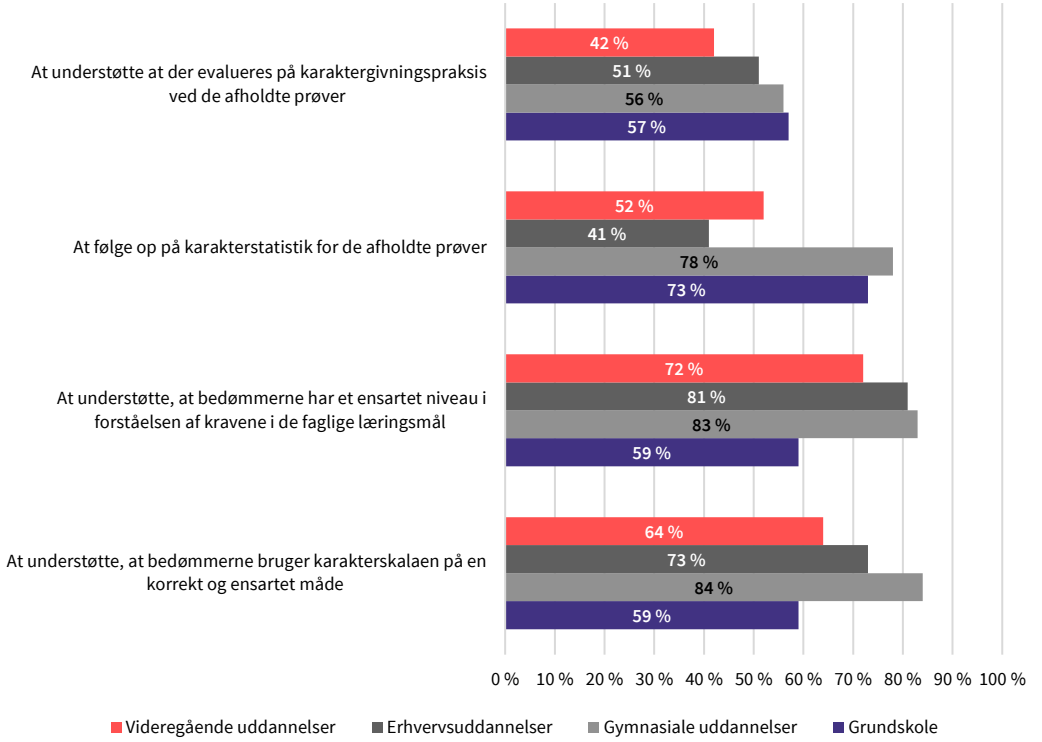
- Fra 59 % (grundskole) til 83 % (gymnasiale uddannelser) af lederne i meget høj grad eller i høj grad ser det at understøtte, at bedømmerne har et ensartet niveau i forståelsen af kravene i de faglige læringsmål, som en del af deres arbejde med karaktergivningen
- Fra 41 % (erhvervsuddannelserne) til 78 % (gymnasiale uddannelser) af lederne i meget høj grad eller i høj grad ser det at følge op på karakterstatistik for de afholdte prøver som en del af deres arbejde med karaktergivningen
- Fra 42 % (videregående uddannelser) til 57 % (grundskole) af lederne i meget høj grad eller i høj grad ser det at understøtte, at der evalueres på karaktergivningspraksis ved de afholdte prøver, som en del af deres arbejde med karaktergivningen.

Figuren viser også, at lederne på de videregående uddannelser generelt adskiller sig, ved at der er en mindre andel af lederne, der i meget høj grad eller i høj grad ser en given opgave som en del af deres arbejde med karaktergivning.

FIGUR 3.9

Ledelsernes perspektiver på egen rolle i forbindelse med karaktergivning, %

I hvilken grad ser du følgende som en del af dit arbejde med karaktergivningen på din skole/institution/uddannelse?



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen*.
Note: Figuren visen andelen af ledere, som har svaret "I meget høj grad" eller "I høj grad". Grundskole: N = 79, gymnasiale uddannelser: N = 122, erhvervsuddannelser: N = 73, videregående uddannelser: N = 370.

Lederne er desuden blevet spurgt om, hvordan de typisk reagerer på problemer med karaktergivningen. På tværs af områder er den typiske reaktion fra lederne, at de drøfter det med underviserne/de interne bedømmere (se *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evaluering af 7-trins-skalaen*).

På det videregående uddannelsesområde angiver 8 % af lederne, at de faglige læringsmål bliver justeret. De lederne på videregående uddannelser, der deltog i fokusgruppeinterviewene, gav desuden udtryk for, at justering af læringsmål/læringsudbyttebeskrivelser som oftest er drevet af andre hensyn end karaktergivningen og udviklingen heri. Det kan fx være udviklingen på arbejdsmarkedet og ændrede kompetencebehov.

Kvalitetssikring på de videregående uddannelser

På de videregående foretages der ikke kvalitetssikring af karaktergivningen fra centralt hold. Det varierer, hvor stor en andel af lederne der ser forskellige kvalitetssikringsopgaver som en del af deres arbejde. Figur 3.11 viser, at:

- 74 % af lederne i meget høj grad eller i høj grad ser det at følge op på evt. tilbagemeldinger fra censorsystemet angående karaktergivningen som en del af deres arbejde med karaktergivningen
- 52 % af lederne i meget høj grad eller i høj grad ser det at følge op på karakterstatistik for de afholdte prøver som en del af deres arbejde med karaktergivningen
- 42 % af lederne i meget høj grad eller i høj grad ser det at understøtte, at der evalueres på karaktergivningspraksis ved de afholdte prøver, som en del af deres arbejde med karaktergivningen.

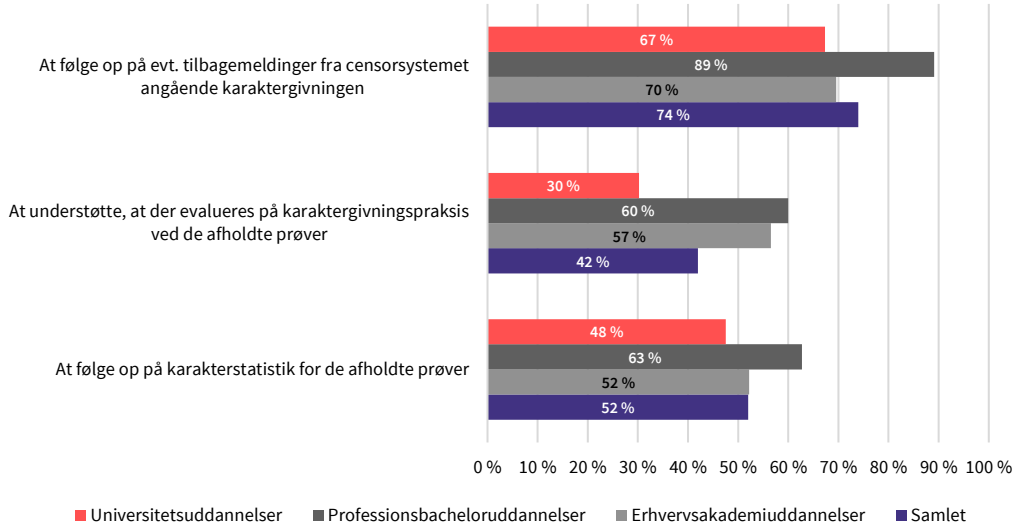
Der er interessante, statistisk signifikante forskelle mellem ledere af hhv. erhvervsakademi-, professionsbachelor- og universitetsuddannelser, jf. figur 4.10:

- 30 % af lederne af universitetsuddannelser ser i meget høj grad eller i høj grad det at understøtte, at der evalueres på karaktergivningspraksis ved de afholdte prøver, som en del af deres arbejde med karaktergivningen, hvilket er signifikant forskelligt fra andelen af ledere af hhv. erhvervsakademi- (56 %) og professionsbacheloruddannelser (60 %)
- 47 % af lederne af universitetsuddannelser ser i meget høj grad eller i høj grad det at følge op på karakterstatistik for de afholdte prøver som en del af deres arbejde med karaktergivningen, hvilket er signifikant forskelligt fra andelen af ledere af professionsbacheloruddannelser (63 %), men ikke fra andelen af ledere af erhvervsakademiuddannelser (52 %)
- 89 % af lederne af professionsbacheloruddannelser ser i meget høj grad eller i høj grad det at følge op på evt. tilbagemeldinger fra censorsystemet angående karaktergivningen som en del af deres arbejde med karaktergivningen, hvilket er signifikant forskelligt fra andelen af ledere af hhv. universitets- (68 %) og erhvervsakademiuddannelser (69 %).

FIGUR 3.10

Ledelserne på de videregående uddannelsers perspektiver på egen rolle i forbindelse med karaktergivning, samlet og fordelt på sektor, %

I hvilken grad ser du følgende som en del af dit arbejde med karaktergivningen på din skole / institution / uddannelse?



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen* samt *Tabelrapport, de videregående uddannelser*.
Note: Figuren visen andelen af ledere, som har svaret "I meget høj grad" eller "I høj grad". Erhvervsakademiuddannelser: N = 23, professionsbacheloruddannelser: N = 110, universitetsuddannelser: N = 202.

Når det kommer til ledernes opfølgning på udfordringer med karaktergivningen, viser spørgeskemaundersøgelsen, at problemer med karaktergivningen typisk følges op af drøftelser med de interne bedømmere og på professionsbachelor- og erhvervsakademiuddannelser også af drøftelser i uddannelsesnævn/ledernetværk. Problemerne følges kun sjældnere eller aldrig op med justeringer af de faglige læringsmål eller justeringer af eksameners sværhedsgrad. Trods stigningen i det generelle karakterniveau svarer 72 % af lederne på de videregående uddannelser, at de inden for de seneste to år aldrig har taget initiativ til at mindske antallet af høje karakterer. 83 % svarer dog også, at de ikke har taget initiativ til at øge det.

Lederne på videregående uddannelser er også blevet spurgt om deres brug af data i kvalitetssikringsarbejdet, dvs. hvilke oplysninger de benytter i arbejdet med at understøtte og kvalitetssikre karaktergivningen på uddannelsen, jf. figur 3.12:

- 64 % af lederne svarer, at de efter hver eksamenstermin eller årligt benytter statistik for karakterfordelingen ved hver prøve i en given eksamenstermin i deres kvalitetssikringsarbejde.
- 60 % af lederne svarer, at de efter hver eksamenstermin eller årligt benytter statistik for den samlede karakterfordeling på uddannelsen i en given eksamenstermin i deres kvalitetssikringsarbejde.
- 51 % af lederne svarer, at de efter hver eksamenstermin eller årligt benytter udviklingen i den samlede karaktergivning for uddannelsen over tid i deres kvalitetssikringsarbejde.

- 49 % af lederne svarer, at de efter hver eksamenstermin eller årligt benytter bedømmernes vurdering af karaktergivningen på uddannelsen i deres kvalitetssikringsarbejde.
- 17 % af lederne svarer, at de efter hver eksamenstermin eller årligt benytter statistik for karaktergivningen ved andre institutioner og/eller beslægtede uddannelser i deres kvalitetssikringsarbejde.

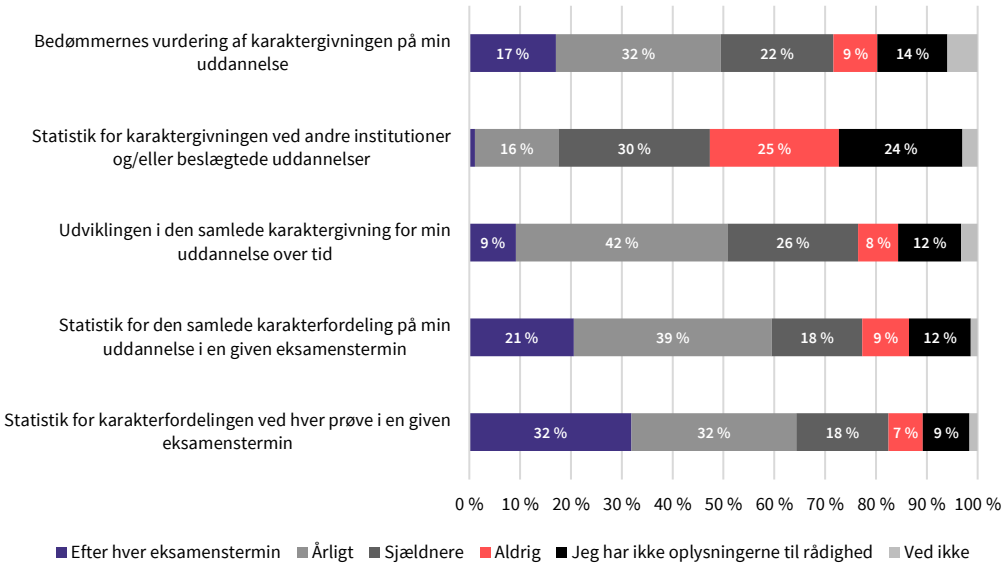
Der er igen markante forskelle mellem svar fra ledere af hhv. erhvervsakademi-, professionsbachelor- og universitetsuddannelser. I flere tilfælde er andelen af ledere af erhvervsakademi- og professionsbacheloruddannelser, der efter hver eksamenstermin eller årligt benytter oplysningerne i deres arbejde, signifikant større end den tilsvarende andel af ledere af universitetsuddannelser. Fx svarer 48 % af lederne af universitetsuddannelser, at de efter hver eksamenstermin eller årligt benytter statistik for den samlede karakterfordeling på uddannelsen i en given eksamenstermin i deres kvalitetssikringsarbejde, hvilket er signifikant forskelligt fra andelen af ledere af hhv. erhvervsakademi- (74 %) og professionsbacheloruddannelser (71 %).

Ca. 20 % af lederne har ikke oplysninger om den samlede karaktergivning på uddannelsen eller benytter dem aldrig i deres kvalitetssikring. Det gælder for 49 % af lederne, når det kommer til oplysninger om karaktergivning på andre institutioner og/eller beslægtede uddannelser.

Disse oplysninger er helt centrale med hensyn til den måde, som Karakterkommissionen forestillede sig, at karaktergivningen skulle kvalitetssikres på. Det er derfor et tegn på, at kvalitetssikringen ikke alle steder foregår, som den var tænkt.

FIGUR 3.11

Ledernes brug af oplysninger i arbejdet med understøttelse og kvalitetssikring af karaktergivningen, videregående uddannelser, %



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen*.
Note: N = 370.

Der tegner sig et noget andet billede i fokusgruppeinterviewene med ledere på de videregående uddannelser. Det overordnede indtryk fra de gennemførte fokusgruppeinterviews er nemlig, at ledelserne generelt er opmærksomme over for karaktergivning og karakterniveauer på uddannelserne. De interviewede ledere fremhæver en række elementer i deres kvalitetssikring, herunder:

- Karakterstatistik
- Tilbagemeldinger fra censorer, enten direkte eller via de årlige censorrapporter
- Behandling af eksamensklager.

Alle interviewede ledere angiver, at de i et eller andet omfang betragter karakterstatistik som en del af deres kvalitetssikring. Om karakterstatistikken er på overordnet niveau (fx gennemsnit) eller vedrører udvalgte fag (typisk karakterer i forbindelse med afsluttende prøver) eller karakterer fra alle eksamener og fx opdelt på eksterne og fastansatte undervisere, varierer meget. Eksempelvis fremhæver en leder, at man holder et årligt møde, hvor karakterstatistik for alle afholdte prøver drøftes. Lederne giver også udtryk for, at de går i dialog med underviserne, hvis der er store udsving i karaktergivning fra ét år til et andet. Lederne fremhæver desuden, at de er særligt opmærksomme over for karaktergivning, hvis der er indført en ny prøveform. To ledere fortæller:

> Jeg overvåger karaktergivningen og ser på, om hele karakterskalaen bliver brugt, eller karakterne klumper sig sammen. Og jeg ser på, om der er udsving fra år til år. Hvis det svinger meget, så kan det være tegn på, at der ikke er den nødvendige ensartethed i karaktergivningen. Og jeg godkender læringsmål, og i den forbindelse ser jeg særligt på, om de gør det muligt at skelne mellem gode og dårlige præstationer. Så kigger jeg på vores nøgletal og ser på, om der er forskel på forskellige grupper af studerende, fx i forhold til adgangsgivende eksamen, for at se på, om vi fx skal gøre noget ved vores optag.
>
> Leder, videregående uddannelse

> Vi kigger især på ændringer gennem årene. Er der udsving fra år til år? Men vi kigger også på, hvor på karakterskalaen karaktergivningen ligger, og så sammenligner vi med andre uddannelser. Og så behandler vi censorrapporten i studienævnet.
>
> Leder, videregående uddannelse

Lederne fremhæver også akkreditering og brug af eksterne censorer som en kvalitetssikringsmekanisme. Med hensyn til det eksterne censorsystem handler det blandt andet om, at ekstern censur i sig selv er en kvalitetssikring af karaktergivningen. Hvad angår akkreditering, handler det om, at man mange steder i forbindelse med akkrediteringen har formuleret politikker for kvalitetssikring, hvor karakterstatistik og behandling heraf beskrives. Det skal dog bemærkes, at det ikke er et krav, at statistik over karakterer skal være en del af institutionernes kvalitetssikring.

Størstedelen af censorformændene har fokus på karaktergivning

Selvom opfølgning på karaktergivning ikke specifikt er beskrevet som en opgave, viser spørgeskemaundersøgelsen, at censorformændene betragter en række opgaver relateret til karaktergivning som en del af deres arbejde:

- 91 % ser det at understøtte, at censorerne forstår og bruger karakterskalaen korrekt, som en del af deres arbejde.
- 82 % ser det at følge op på, om ankesagsbehandlingen og sager om bedømmelse tager udgangspunkt i de opstillede læringsmål, som en del af deres arbejde.

- 78 % ser det at følge op på, om der er et ensartet niveau i de faglige krav i bedømmelserne på tværs af censorkorpsets uddannelser og institutioner, som en del af deres arbejde.
- 76 % ser det at følge op på, om bedømmelserne tager udgangspunkt i de opstillede læringsmål, som en del af deres arbejde.
- 73 % ser det at følge op på udviklingen i karaktergivningen inden for censorkorpsets uddannelser over tid som en del af deres arbejde.
- 73 % ser det at understøtte, at institutionerne inddrager karaktergivningen i kvalitetssikringen og udviklingen af deres uddannelser, som en del af deres arbejde.
- 71 % ser det at følge op på, om der er opstillet læringsmål for alle prøver, som en del af deres arbejde.

Størstedelen af censorformændene (71 %) følger også op på udviklingen i karaktergivningen inden for censorområdets uddannelser. Dog angiver kun 42 % af censorformændene, at de har adgang til karakterstatistik for alle bedømmelser på korpsets uddannelser. Hvad angår samarbejde med institutionerne, oplever 52 % af censorformændene i høj grad eller i nogen grad, at institutionerne samarbejder om at følge op på udviklingen i karaktergivningen inden for deres censorområde. 19 % svarer, at de slet ikke er af den opfattelse, at institutionerne samarbejder herom.

4 Karakterskalaens udformning

4.1 Baggrund og undersøgelsesspørgsmål

Med overgangen fra 13-skalaen til 7-trins-skalaen gik man fra en dansk skala med ti karaktertrin til en skala med syv trin. Talværdierne i 7-trins-skalaens karaktertrin blev fastlagt med udgangspunkt den internationale ECTS-skala (se "7-trins-skalaen og ECTS-skalaen – appendiks B" for yderligere beskrivelse af sammenhængen). Det betød, at der var større spring mellem de tre midterste bestå-ede karakterer 4, 7 og 10 end mellem de øvrige karakterer. Samtidig var det en prioritet for Karak-terkommissionen at understøtte, at der var en klar trindifferentiering og dermed en større sikker-hed i karaktergivningen. Det skulle understøttes både gennem beskrivelsen af karaktertrinene og springene imellem dem. Karakterkommissionens begrundelse for reduktionen af antallet af trin var, at kommissionen i forbindelse med interviews og samtaler ofte blev mødt med det synspunkt, at det kan være svært at afgøre, om en given præstation skal vurderes til fx 7 eller 8 eller fx 8 eller 9.

> Antallet af trin og springene mellem trinene i karakterskalaen skal være afpasset således, at valget mellem en karakter og nabokarakteren er udtryk for en sikker differentiering. Samtidig skal skalaen være så finkornet, at den kan benyttes i forbindelse med den formative vurdering.
>
> Betænkning om indførelse af en ny karakterskala til erstatning af 13-skalaen, 2004

Af de syv trin er to ikke-bestået:

- 00 – den utilstrækkelige præstation – gives for den utilstrækkelige præstation, der ikke demon-strerer en acceptabel grad af opfyldelse af fagets mål
- -3 – den ringe præstation – gives for den helt uacceptable præstation.

Karakterkommissionen valgte at indføre to ikke-beståede karakterer placeret symmetrisk omkring den laveste beståede karakter. De to karakterer skulle give mulighed for at sende et signal til den studerende om, hvor langt præstationen var fra at være bestået. I appendiks B uddybes det, hvor-for man endte med karakteren -3.

I dette kapitel sætter vi fokus på evalueringsspørgsmål 3: "Hvad betyder karakterspringene og ka-rakteren -3 for karaktergivningen?"

Evalueringen har primært fokuseret på den summative brug af karakterskalaen, selvom den forma-tive brug også berøres. Med summativ menes, at karakteren er et udtryk for eksaminandens faglige standpunkt ved afslutningen af et undervisningsforløb. Med formativ menes, at skalaen anvendes som et hjælpemiddel til at støtte den studerende i vedkommendes læreproces forud for den sum-mative vurdering, idet man her skal hjælpe den studerende til at forbedre sine præstationer.

Kapitlet bygger på:

- Deskriptive og regressionsbaserede analyser af data for alle tilgængelige afgivne karakterer i perioden 2007-16 (2007-15 for de gymnasiale uddannelser).

  Erhvervsuddannelserne er ikke repræsenteret i de deskriptive og regressionsbaserede analyser, da der kun forelå data for en toårig periode, og det dermed ikke var muligt at analysere udviklingen i karaktergivningen.
- Spørgeskemaundersøgelse blandt bedømmere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser.
- Spørgeskemaundersøgelse blandt ledere på grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser samt censorformændene på de videregående uddannelser.
- Fokusgruppeinterviews og enkeltinterviews med bedømmere på grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser.

4.2 Perspektiver på springene på karakterskala

4.2.1 De brede karaktertrin gør det ikke lettere at blive enige

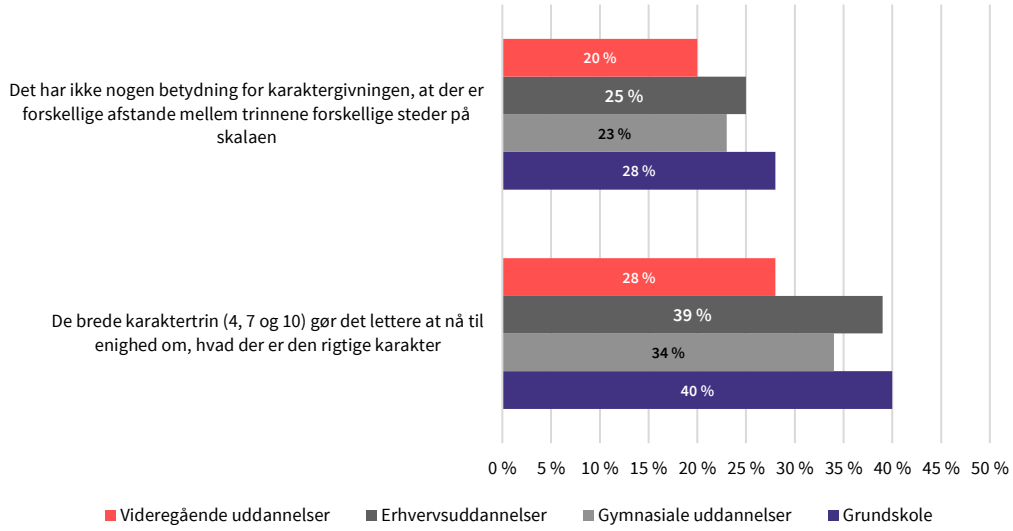
Evalueringen viser, at de store spring for en del af bedømmerne ikke understøtter en større sikkerhed i karaktergivningen. På tværs af grundskoleområdet, de gymnasiale uddannelser, erhvervsuddannelserne og de videregående uddannelser angiver 28-40 % af bedømmerne, at de er meget enige eller enige i, at de brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter, jf. figur 4.1. Særligt bedømmerne på grundskoleområdet oplever dette. Det er samtidig de færreste bedømmere (fra 20 % (videregående uddannelser) til 28 % (grundskole)), der vurderer, at det ikke har betydning for karaktergivningen, at der er forskellige afstande mellem trinene forskellige steder på skalaen.

I fokusgrupperne var der på den ene side en gruppe bedømmere, der oplever, at karaktergivningen bliver lettere, og enigheden mellem censor og eksaminator større, når karaktererne er brede, og der skal differentieres mindre. På den anden side oplever andre bedømmere, at de brede karakterer vanskeliggør karaktergivningen.

FIGUR 4.1

Andelen af bedømmere, som er enige eller meget enige i følgende udsagn om springene i karakterskalaen, %

Hvor enig eller uenig er du i følgende udsagn om springene i karakterskalaen



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

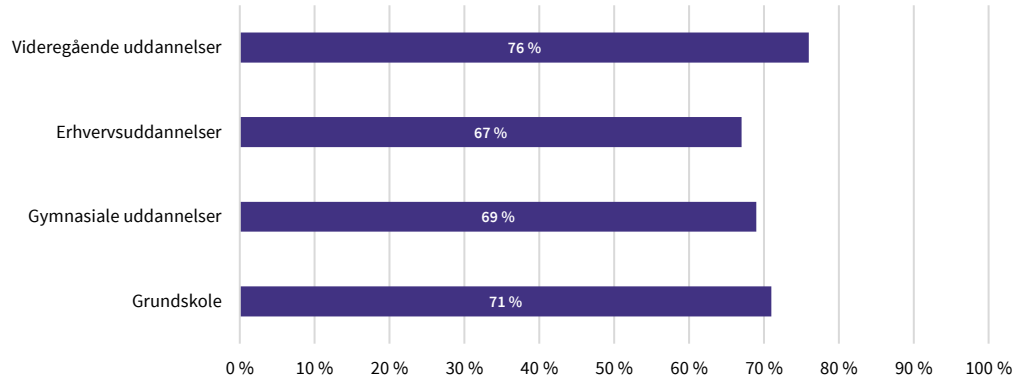
De store spring mellem 4, 7 og 10 udfordrer

Spørgeskemaundersøgelsen blandt bedømmerne viser, at flertallet oplever, at der er områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, jf. figur 4.2.

Der er dog forskelle, alt efter hvilket fag bedømmerne på de gymnasiale uddannelser underviser i. 79 % af bedømmerne i dansk angiver, at der er områder på skalaen, hvor det er svært at vurdere, hvilken karakter eleven skal have. Det samme gælder 76 % af bedømmerne i engelsk og 49 % af bedømmerne i matematik. En lignende, men ikke så udtalt tendens ses på grundskoleområdet. På de videregående uddannelser er der ikke nævneværdige forskelle på tværs af hovedområderne.

FIGUR 4.2

Andel af bedømmere, som oplever, at der er områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, %

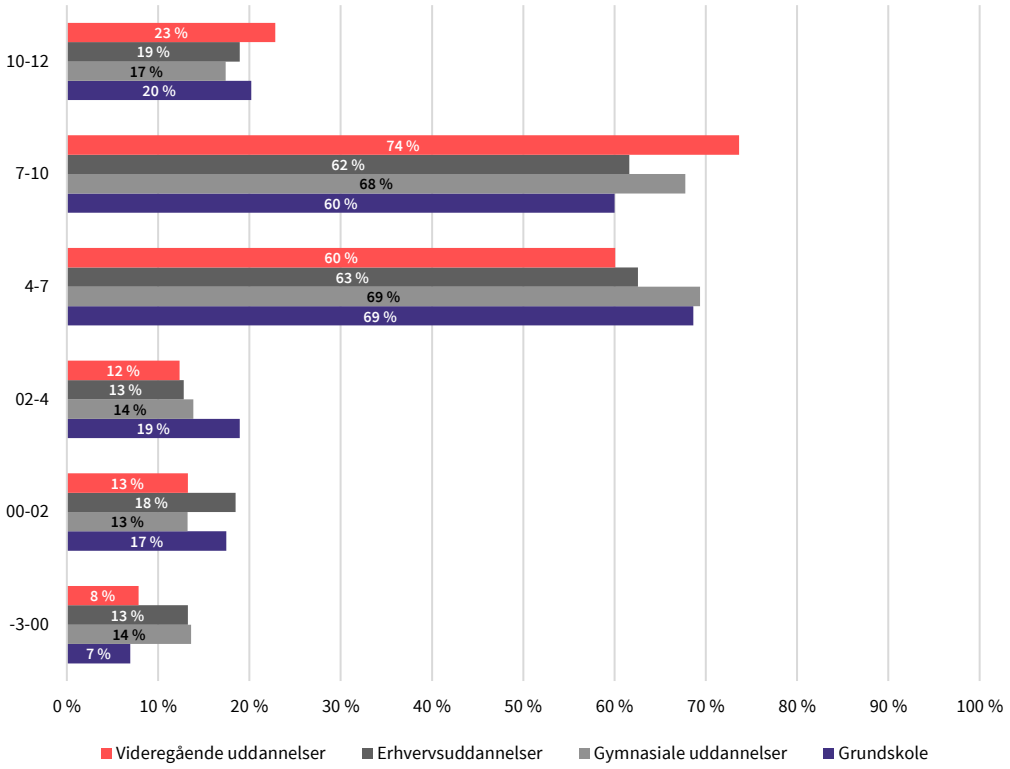


Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

Bedømmerne oplever primært, at det er svært at vurdere, hvilken karakter eleven/den studerende skal have, når de ligger mellem 4 og 7 eller mellem 7 og 10, jf. figur 4.3. Over 60 % af de bedømmere, der oplever, at der er steder på skalaen, hvor det er svært at afgøre, hvilken karakter eleven/den studerende skal have, vurderer, at det er mellem 4 og 7 og mellem 7 og 10, at det er svært at vurdere.

FIGUR 4.3

Bedømmernes vurdering af, hvor på skalaen det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*

I fokusgruppeinterviewene er der bedømmere, som beskriver, at de store spring på midten af skalaen mellem 4, 7 og 10 gør det lettere for dem at give karakterer. Andre bedømmere oplever omvendt, at spændet på midten af skalaen kan give særlig anledning til diskussion mellem censor og eksaminator, og at det blandt andet kan handle om, at de store spring gør bedømmelsen sværere.

De bedømmere, der oplever udfordringer med midterkaraktererne, beskriver, at det skyldes, at trinene er så brede, at det kan være vanskeligt at afgøre, om der er tale om den ene eller den anden karakter, og om en elev/studerendes præstation eksempelvis er til et stort 7-tal eller et lille 10-tal, fordi karaktererne kan rumme meget forskelligartede præstationer.

> Det [springene] gør selvfølgelig, at der er større diskussion og overvejelser. Også fordi det er så stort et spænd, og vi måske har siddet med en opgave inden, der også har fået et 7-tal. Så giver det anledning til en længere votering, at der er de her store spring.
>
> Bedømmer, gymnasial uddannelse

Derudover peger nogle bedømmere på, at trinbeskrivelsens formuleringer ikke bidrager til særlig afklaring. Et perspektiv i interviewene med bedømmerne på *grundskoleområdet* og *de gymnasiale uddannelser* er, at ordlyden i den generelle trinbeskrivelse efterlader et for stort fortolkningsrum. Der er bedømmere, der oplever, at de bærende ord i beskrivelsen af de forskellige karakterer, fx ”jævnt”, ”fortrinligt” og ”utilstrækkeligt”, er for overordnede og derfor ikke tilstrækkeligt præcise

som bedømmelsesgrundlag. En særlig kritik rettes af ordene "fremragende", der betegner præstationen ved karakteren 12, og "fortrinligt", der betegner præstationen ved 10-tallet, fordi det kan være vanskeligt at vurdere, hvilket af de to tillægsord der reelt betegner den bedste præstation.

Kun 10 og 12 betragtes som gode karakterer – 7-tallet er utilfredsstillende for eleverne/de studerende

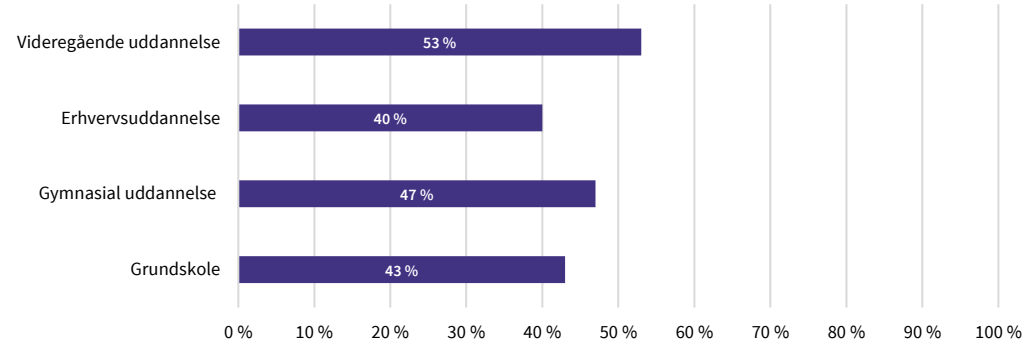
Bedømmerne fortæller i fokusgruppeinterviewene og de kvalitative interviews, at eleverne/de studerende har en tendens til kun at betragte 10 og 12 som gode karakterer, mens 7-tallet ikke oplæves som en god karakter. Bedømmerne peger på, at en del af forklaringen kan være, at mange af elevernes forældre selv bedst kender 13-skalaen og derfor betragter et 7-tal som en under middelkarakter. Nogle bedømmere, særligt de, der også har bedømt med 13-skalaen, beskriver, at de også selv har skullet lære at betragte både 4 og 7 som gode karakterer.

Blandt gymnasieeleverne er der desuden et øget fokus på karaktergennemsnit og dermed på de enkelte karakterer, da de har en oplevelse af, at det har stor betydning med hensyn til at komme ind på en videregående uddannelse. Det betyder ifølge bedømmerne på de gymnasiale uddannelser, at en del elever er villige til at gøre en ekstra og meget målrettet strategisk indsats for at forbedre deres karakterer. Lærerne oplever desuden karaktergivningen som et meget stort ansvar, fordi karaktererne betyder så meget for eleverne, og de store spring kan gøre en stor forskel for elevernes gennemsnit. Det er dog ikke lærerne på de gymnasiale uddannelser, men derimod underviserne på de videregående uddannelser, der er mest enige i, at de store spring mellem 4, 7 og 10 jævnligt fører til en højere bedømmelse for ikke at give eleven/den studerende en for lav karakter, jf. figur 4.4. 53 % af bedømmerne på de videregående uddannelser er enige i, at de store spring mellem 4, 7 og 10 jævnligt fører til en højere bedømmelse for ikke at give eleven/den studerende en for lav karakter.

FIGUR 4.4

Andelen af bedømmere, som er enige eller meget enige i følgende udsagn om springene på karakterskalaen, %

Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give eleven/den studerende en for lav karakter



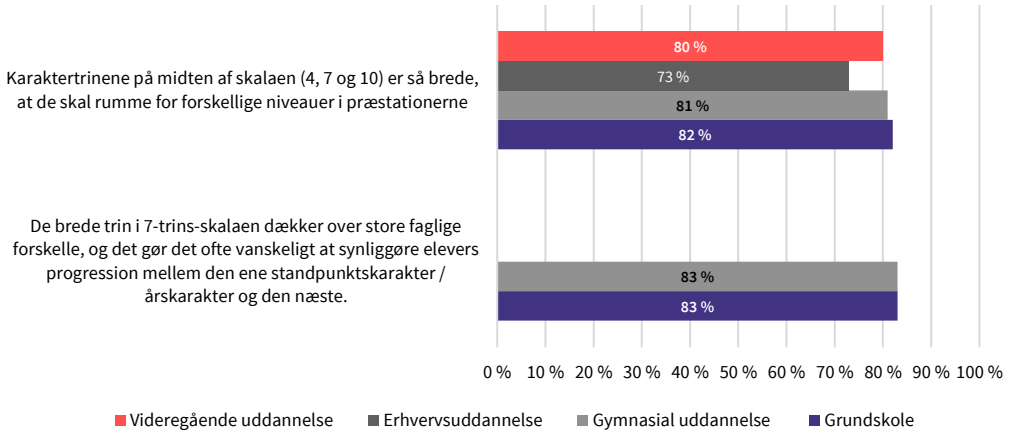
Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser N = 314, videregående uddannelser N = 2.946.

Svært at anvende skalaen til løbende karaktergivning

Evalueringen peger på, at bedømmerne på grundskoleområdet og de gymnasiale uddannelser oplever, at skalaen er svær at anvende til løbende karaktergivning. Langt størstedelen af bedømmerne (73-82 %) er enige eller meget enige i, at karaktererne på midten af skalaen (4,7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne, jf. figur 4.5. 83 % af bedømmerne på *grundskoleområdet* og *de gymnasiale uddannelser* angiver desuden, at de er enige eller meget enige i udsagnet "De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevers progression mellem den ene standpunktskarakter/årskarakter og den næste" (se figur 4.5).

FIGUR 4.5

Andelen af bedømmere, som er enige eller meget enige i følgende udsagn om springene på karakterskalaen, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

I fokusgruppeinterviewene nævner nogle bedømmere, at de savner i højere grad at kunne nuancere og differentiere i forbindelse med karaktergivningen på midten af skalaen. For bedømmerne i grundskolen og de gymnasiale uddannelser drejer det sig om at kunne bedømme forskellige elever mere retvisende og nuanceret og om, at de store spring på midten af skalaen har stor betydning for elevernes karaktergennemsnit, men også om, at de store spring kan have betydning for elevernes motivation og muligheden for at synliggøre elevernes progression. Det kan fx være svært for en elev at få samme karakter som en anden på trods af tydelige niveauforskelle:

> I de mundtlige eksamener er der ofte brug for et 9-tal, hvis eleven ligger lige imellem 7 og 10. Det er så ærgerligt, hvis man ikke kan anerkende, at de har undergået en udvikling, fordi de har bevæget sig fra et lille 7-tal til et stort 7-tal. Ja, der er virkelig stor forskel på de dårligste og de bedste 7-taller. Jeg kan godt finde på lige at give en elev på et stort 7-tal et 10-tal.
>
> Bedømmer, grundskole

Bedømmerne undrer sig desuden over, at man netop på midten af skalaen, hvor hovedparten af eleverne/de studerende ligger, har valgt så store spring og dermed en begrænset mulighed for differentiering, mens springene er mindre andre steder på skalaen, hvor færre elever/studerende befinder sig. På den baggrund efterlyser disse bedømmere flere karakterer på midten af skalaen.

Lederne identificerer desuden i spørgeskemaundersøgelsen følgende svaghed som den væsentligste (jf. tabel 4.1): "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne". Den eneste undtagelse er lederne på de gymnasiale uddannelser, der har "De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevers progression mellem den ene standpunktskarakter/årskarakter og den næste" som den væsentligste svaghed, dog efterfulgt af "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne".

TABEL 4.1

Ledernes perspektiv på svagheder ved karakterskalaen, toptre fordelt på områder

	1	2	3
Grundskole	"Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne" (70 %)	"Skalaen rummer ikke i tilstrækkelig grad mulighed for at honorere den ekstraordinært gode præstation" (55 %)	"De brede trin på 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevers progression mellem den ene standpunktskarakter / årskarakter og den næste" (53 %)
Erhvervsuddannelser	"Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne" (84 %)	"Skalaen rummer ikke i tilstrækkelig grad mulighed for at honorere den ekstraordinært gode præstation" (55 %)	"De brede trin på 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevers progression mellem den ene standpunktskarakter/årskarakter og den næste" (45 %)
Gymnasiale uddannelser	"De brede trin på 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevers progression mellem den ene standpunktskarakter/årskarakter og den næste" (59%)	"Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne" (56 %)	"At bedømmelserne i for høj grad skal tage udgangspunkt i mangler i præstationen" (49 %)
Videregående uddannelser	"Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne" (72 %)	"Skalaen rummer ikke i tilstrækkelig grad mulighed for at honorere den ekstraordinært gode præstation" (54 %)	"At bedømmelserne i for høj grad skal tage udgangspunkt i mangler i præstationen" (42 %)
Censorformænd	"Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne" (74 %)	"Skalaen rummer ikke i tilstrækkelig grad mulighed for at honorere den ekstraordinært gode præstation" (48 %)	"Der er for få trin på skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse" (37 %)

Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evaluering af 7-trins-skalaen*.
Note: Tallet i parentes indikerer andelen af ledere, som har angivet en given styrke som første-, anden- eller tredjeprioritet. Grundskole: N = 79, gymnasiale uddannelser: N = 122, erhvervsuddannelser: N = 73, videregående uddannelser: N = 370.

Omvendt vælger kun 9-12 % af lederne "De brede karaktertrin (4, 7 og 10) understøtter en entydig tildeling af karakterer" som en styrke, hvilket gør det til den mindst valgte styrke blandt lederne, jf. *Ledernes perspektiv på 7-trins-skalaen – Bilagsrapport til Evaluering af 7-trins-skalaen*.

De store spring på midten af skalaen har stor betydning for karaktergennemsnittet

Generelt beskriver bedømmerne, at tvivlen, hvis der er tvivl om, hvilken midterkarakter eleven/den studerende skal have, kommer eleven/den studerende til gode. Dette sker ud fra en bevidsthed om, at karakteren kan have stor betydning for eleven/den studerende. Bedømmerne peger blandt andet på, at det kan have store konsekvenser for elevernes/de studerendes karaktergennemsnit.

> Et 7-tal kræver to 12-taller for at blive opvejet. Det virker ikke rimeligt. Der straffes for hårdt, når man går ned fra 10 til 7. Springet på tre enheder er for stort.
>
> Bedømmer, videregående uddannelse

Interviewene peger således på, at bedømmerne er opmærksomme på den skævvridning, som de ujævne spring i skalaen kan medføre. Fx giver et 7-tal og et 12-tal tilsammen et lavere gennemsnit end to 10-taller. Ligeledes giver et -3-tal og et 02-tal et lavere gennemsnit end to 00-taller, selvom de to 00-taller begge er dumpede karakterer.

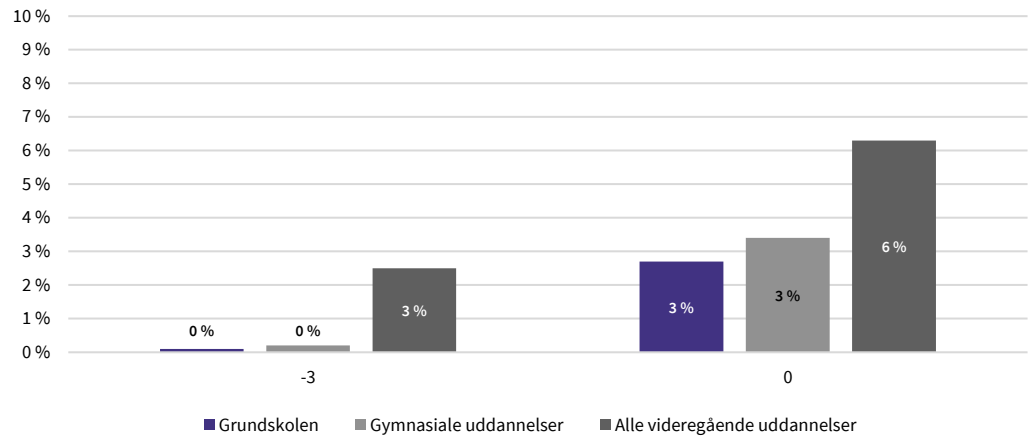
4.3 Brug af og perspektiver på -3 og 00

4.3.1 -3 er en undtagelseskarakter med stor symbolsk betydning

På tværs af *grundskoleområdet, de gymnasiale uddannelser* og *de videregående uddannelser* fortæller bedømmerne i fokusgruppeinterviewene, at de meget sjældent anvender karakteren -3. Denne pointe understøttes af de deskriptive og regressionsbaserede analyser af karakterdata, som viser, at under 1 % af de karakterer, som er blevet givet siden 2007 i grundskolen og på de gymnasiale uddannelser, har været -3. På de videregående uddannelser gør det sig gældende for 3 % af karaktererne, jf. figur 4.6.

FIGUR 4.6

Andelen af ikke-beståede karakterer på de enkelte uddannelsestyper, 2007-15/16, %



Kilde: EVA, 2018.
Note: Figuren viser den faktiske karakterfordeling for ikke-beståede karakterer i perioden 2007-16 (for de gymnasiale uddannelser kun til og med 2015).

Adspurgt om, hvad forskellen på -3 og 00 er, og hvornår -3 kan komme i spil, svarer bedømmerne, at man ved 00 som elev/studerende har gjort et forsøg, men ikke er lykkedes, mens -3 kommer i

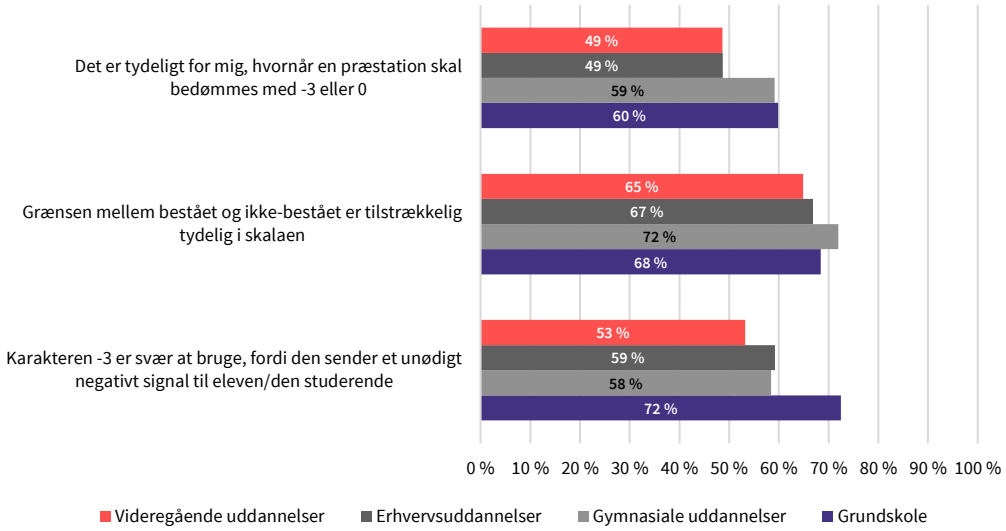
spil ved en fuldstændig uacceptabel eller decideret fraværende præstation. En bedømmer fra grundskolen eksemplificerer, hvordan fraværet af forsøg kan se ud:

> Jeg har prøvet at give -3 til en mundtlig eksamen. Der sad en elev, der ikke ville fremføre sin præstation, hun ville ikke svare på nogen spørgsmål og heller ikke trække et spørgsmål. Hun svarede ikke på noget.
>
> Bedømmer, grundskolen

I spørgeskemaundersøgelsen svarede ca. 50-60 % af bedømmerne, at det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 eller 00, jf. figur 4.7. Omvendt er 14-30 % uenige eller meget uenige i, at det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 og 00. Særligt bedømmerne på de videregående uddannelser (30 %) er uenige heri.

FIGUR 4.7

Andelen af bedømmere, som er enige eller meget enige i følgende udsagn om syvtrinsskalaen, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: Figuren visen andelen af bedømmere, som er enige eller meget enige i de angivne udsagn. Grundskole: N = 665, gymnasiale uddannelser: N = 1.187, erhvervsuddannelser: N = 314, videregående uddannelser: N = 2.946.

I fokusgruppeinterviewene fortæller bedømmerne, at de oplever, at -3 har stor symbolsk betydning. Bedømmerne er generelt kritiske i forhold til symbolikken i at have en minuskarakter. Derudover begrunder nogle bedømmere deres kritik af det at have en minuskarakter med den betydning, som en sådan karakter har for elevernes/de studerendes gennemsnit. Af den grund anvender nogle bedømmere principielt ikke karakteren -3. Bedømmerne siger blandt andet om -3, at "der er noget helt upædagogisk ved -3 på alle måder", "det er utilstedeligt. Det gør ikke noget godt for de studerendes læring og burde smides væk", "det virker som at gnide salt i såret" og "det er lidt udtryk for, at man er dummere, end da man kom".

Bedømmerne på *de gymnasiale og videregående uddannelser* har også det perspektiv, at anvendelsen af -3 skal sende et signal til eleven/den studerende om, at det ikke er inden for rækkevidde at bestå faget, og at eleven/den studerende måske endda bør genoverveje sit studievalg:

> Jeg synes, at -3 skal signalere, at det her er så langt fra bestået, at det er urealistisk. 00 skal signalere, at du skal arbejde på noget, og så er der 02 på den anden side.
>
> Bedømmer, gymnasial uddannelse

> Jeg har brugt -3 tre gange i min karriere som underviser [ni år]. Det er brugt til at signalere, at du er ikke bare dumpet, men at det, du leverede, var seriøst dårligt. Jeg var engang censor med en underviser, der sagde, at for ham var 00 et signal om, at præstationen var skidt, men det kan ske for os alle, hvor -3 betyder, at den studerende skal finde noget andet at lave.
>
> Bedømmer, videregående uddannelse

Som det fremgår af figur 4.7, er der forskel på, i hvor høj grad bedømmerne vurderer, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende. 72 % af bedømmerne på grundskolen er enige/meget enige i dette, hvilket er en signifikant højere andel end på de øvrige uddannelser. 53 % af bedømmerne på de videregående uddannelser, 59 % af bedømmerne på erhvervsuddannelserne og 58 % af bedømmerne på de gymnasiale uddannelser er enige eller meget enige i dette udsagn.

Kun 15-23 % af lederne vælger "At den negative karakter -3 sender et tydeligt signal om, at præstationen ikke er tilstrækkelig" som en styrke, hvilket gør det til den fjerdemindst valgte styrke (se bilagsrapporten *Ledernes perspektiv – Bilagsrapport til Evalueringen af 7-trins-skalaen*). Lederne havde mulighed for at vælge tre styrker ud af i alt ni mulige.

Karakteren -3 er sværere at anvende ved mundtlige bedømmelser

Bedømmerne på *de gymnasiale og videregående uddannelser* fortæller, at -3 i højere grad anvendes ved skriftlige bedømmelser end ved mundtlige. Dette skyldes, dels at man ved en mundtlig bedømmelse for det meste kan få hevet et forsøg på faglighed ud af eleven/den studerende, mens der ved skriftlig bedømmelse kun er det at bedømme, man har på skrift. Dels at det opleves som sværere at give -3 til en person, man sidder over for ansigt til ansigt, frem for til en anonym elev/studerende på skrift.

Dumpegrænsen er tydelig, men vigtig at diskutere ekstra grundigt

I spørgeskemaundersøgelsen svarer 65-72 % af bedømmerne på tværs af områderne, at de er enige eller meget enige i udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkeligt tydelig i skalaen", jf. figur 4.7. Dette understøttes af de kvalitative fokusgruppeinterviews. Her fortæller bedømmerne, at de som regel ikke er i tvivl om, hvorvidt en elev/studerende skal dumpe eller ej.

Bedømmerne beskriver, at dumpekarakterer er noget af det, man diskuterer allermest i forbindelse med karaktergivning. Dette skyldes den betydning, en dumpekarakter har for eleven/den studerende. Bedømmerne gør derfor meget ud af at have argumenterne på plads med hensyn til, hvorfor eleven/den studerende dumper. Derfor mener bedømmerne, at det er vigtigt at bruge tid på at være ekstra grundig i gennemgangen af, hvilke argumenter der taler for hhv. dumpet og bestået, og sikre sig, at der ikke er noget, der gør, at eleven ikke skal dumpe. Det kan blandt andet ske, ved at bedømmerne undervejs i den mundtlige prøve stiller spørgsmål, som eleven/den studerende ikke kan svare på, og dermed får belyst konkrete fejl og mangler, eller at de bruger ekstra lang tid på voteringen. En bedømmer fortæller blandt andet:

> For mit vedkommende hænger det også sammen med, at det er en alvorlig besked at give til en elev, meget mere alvorlig end de andre karakterer. Derfor synes jeg også, at det er vigtigt, at man lige gør sig ekstra umage og er ekstra sikker på, at det er den rigtige karakter, når man giver et 00.
>
> Bedømmer, gymnasial uddannelse

5 Belønning og identifikation af særligt dygtige studerende

5.1 Baggrund og undersøgelsesspørgsmål

7-trins-skalaen har ikke en undtagelseskarakter til den ekstraordinære præstation, som den tidligere skalas 13-tal. 13-tallet blev på den gamle skala givet for den usædvanlig selvstændige og udmærkede præstation, mens 12-tallet på 7-trins-skalaen gives for den fremragende præstation, der demonstrerer udtømmende opfyldelse af fagets mål med ingen eller få uvæsentlige mangler. Det forventes, at omkring 10 % af alle beståede karakterer er 12.

Det er et bevidst valg, at 7-trins-skalaen ikke har nogen undtagelseskarakter som det gamle 13-tal for at gøre skalaen mere internationalt anvendelig. Der er kun få karakterskalaer internationalt, som har undtagelseskarakterer. En Eurydice-rundspørge blandt de europæiske lande foretaget af Styrelsen for Forskning og Uddannelse peger på, at topkarakteren i de udenlandske skalaer ofte bruges til 13-30 % af de beståede prøver, og tal fra USA peger på, at andelen kan være endnu højere (se evt. "7-trins-skalaen og ECTS-skalaen" appendiks B). 13-tallet var derfor svært at forklare for personer i udlandet.

Samtidig pegede Karakterkommissionen på, at der var stor forskel mellem sektorer, fag og hovedområder med hensyn til, hvor ofte bedømmerne gav 13.

Det er dog blevet påpeget, at ændringen gør det sværere for aftagerne at identificere de dygtigste danske studerende, og at det er uhensigtsmæssigt, at skalaen ikke giver mulighed for at belønne den ekstraordinære præstation.

Uddannelses- og Forskningsministeriet har implementeret en talentbekendtgørelse, som blandt andet giver institutionerne mulighed for at belønne særligt dygtige studerende. Men mulighederne for belønning er ikke indbygget i karaktergivningen.

I dette kapitel sætter vi fokus på evalueringsspørgsmål 4: Er der tilstrækkelige muligheder for at identificere særligt dygtige studerende og belønne den ekstraordinære præstation?

Kapitlet bygger på:

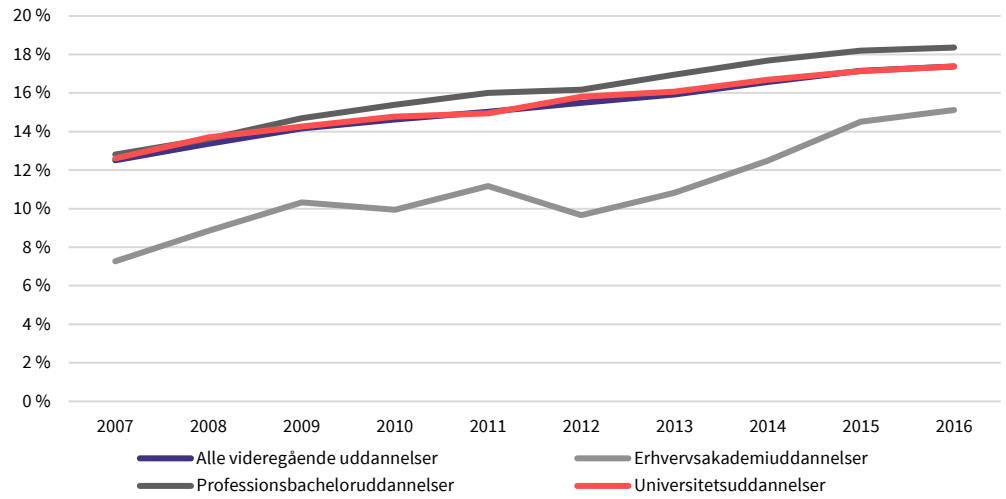
- Deskriptive og regressionsbaserede analyser af data for alle tilgængelige afgivne karakterer i perioden 2007-16 på grundskoleområdet, de gymnasiale uddannelser og videregående uddannelser (2007-15 for gymnasiale uddannelser)
- Spørgeskemaundersøgelse blandt ledere og bedømmere på de videregående uddannelser
- Fokusgruppeinterviews med bedømmere og ledere på de videregående uddannelser
- Interviews med arbejdsgiverorganisationer (Dansk Industri, Dansk Erhverv (medlemsrepræsentant), Dansk Byggeri og Finanssektorens Arbejdsgiverforening).

5.2 Udbredelse af 12-tallet på de videregående uddannelser

Siden indførelsen af 7-trins-skalaen har været en stigning i anvendelsen af 12-tallet på de videregående uddannelser samlet set og på tværs af de tre store sektorer, jf. figur 5.1. Udviklingen i 12-tallet har været stort set identisk på professionsbachelor- og universitetsuddannelserne, hvor der har været en stigning fra knap 13 % 12-taller i 2007 til omtrent 18 % 12-taller i 2016. Andelen af uddelte 12-taller på erhvervsakademiuddannelserne er lavere, men til gengæld har stigningen været mere markant samlet set. Man er således gået fra at tildele et 12-tal ved omtrent 7 % af alle eksamener i 2007 til at tildele 12-taller ved 15 % af alle eksamener i 2016.

FIGUR 5.1

Udvikling i 12-andelen ved alle videregående uddannelser og fordelt på tre største sektorer, 2007-16, %

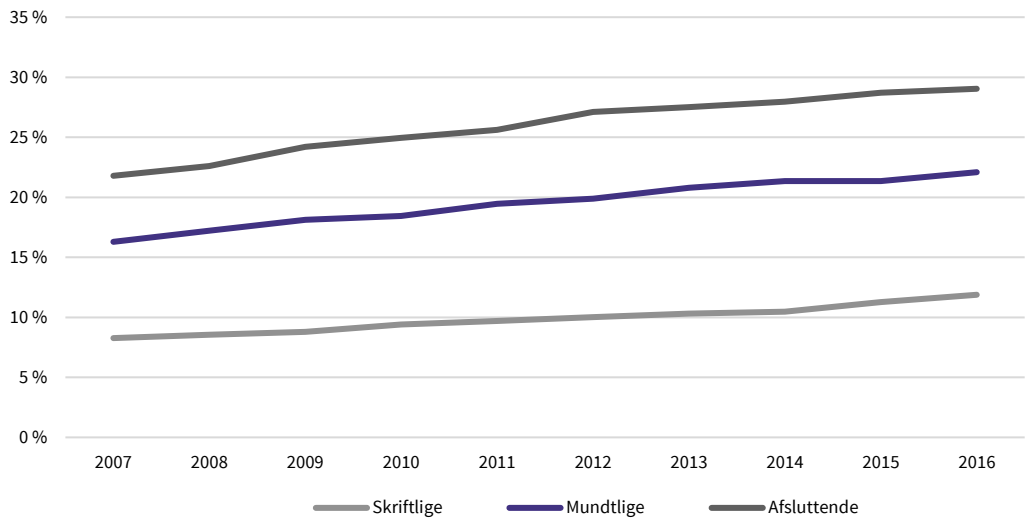


Kilde: EVA på grundlag af data indsamlet af Uddannelses- og Forskningsministeriet.
Note: Figuren viser den tidsmæssige udvikling i 12-andelen ud af alle beståede karakterer ved alle videregående uddannelser, erhvervsakademi-, professionsbachelor- og universitetsuddannelser i perioden 2007-16. Alle: N = 7.248.401, erhvervsakademiuddannelser: N = 504.690, professionsbacheloruddannelser: N = 1.830.829, universitetsuddannelser: N = 4.672.246.

Figur 5.2 viser udviklingen i 12-andelen fordelt på de tre primære prøveformer på alle videregående uddannelser. Vi kan se, at andelen af 12-taller har været stigende for alle prøveformer. 12-andelen for skriftlige prøver er steget fra ca. 8 % i 2007 til lidt over 12 % i 2016, mens den for mundtlige prøver er steget fra omtrent 16 % i 2007 til over 22 % i 2016. 12-andelen for afsluttende prøver har gennem hele perioden ligget på det højeste niveau af samtlige prøveformer, og siden 2007 er 12-andelen steget fra ca. 22 % til mere end 29 % i 2016.

FIGUR 5.2

Udvikling i 12-andelen ved alle videregående uddannelser fordelt på prøveformer, 2007-16, %



Kilde: EVA på grundlag af data indsamlet af Uddannelses- og Forskningsministeriet.
Note: Figuren viser den tidsmæssige udvikling i 12-andelen ud af alle beståede karakterer ved alle videregående uddannelser fordelt på de tre eksamensformer; skriftlige, mundtlige og afsluttende. Skriftlige: N = 2.514.476, mundtlige: N 1.875.661, afsluttende: N = 522.888.

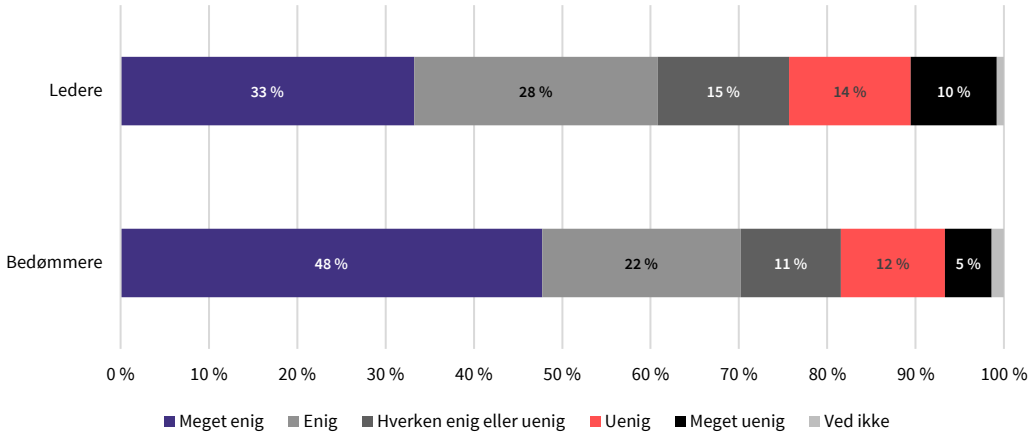
I kortlægningen af karakterfordelingen undersøgte vi, hvilken betydning udviklingen i populationen af de studerende over tid har haft for inflationen i 12-andelen. Analysen viser, at inflationen havde været højere, hvis populationen af studerende i de efterfølgende år havde været den samme som i 2007. Dette skyldes, at man over årene har udvidet optaget på de videregående uddannelser, så der i dag optages langt flere, end der gjorde i 2007. Stigningen i optaget har ført til, at flere studerende fra eksempelvis ikke-akademiske hjem bliver optaget på en videregående uddannelse, hvilket altså lægger en dæmper på inflationen. For samtlige år siden 2007 estimeres inflationen til at være højere, hvis vi tager højde for udviklingen i sammensætningen af de studerende. Det er dermed muligt at konkludere, at vi ikke med observerbare kontrolvariable kan forklare, hvorfor inflationen har fundet sted. Det kan på den anden side ikke afvises, at uobserverede faktorer, såsom øget effektivitet, studieintensitet etc., vil kunne forklare en del af den inflationsudvikling, som vi har set siden indførelsen af 7-trins-skalaen.

5.3 Perspektiver på muligheder for identifikation og belønning af særligt dygtige studerende

Størstedelen af lederne (61 %) og bedømmerne (70 %) på det videregående uddannelsesområde er enige eller meget enige i, at karakterskalaen bør have mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet, jf. figur 5.3. Og både lederne på de videregående uddannelser og censorformændene peger på, at det, at skalaen ikke i tilstrækkelig grad rummer mulighed for at honorere den ekstraordinært gode præstation, er en svaghed ved skalaen. Blandt lederne er der 54 %, der anser det som en svaghed (se tabel 4.1). Det samme gælder 49 % af censorformændene.

FIGUR 5.3

Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet på 13-skalaen, %



Kilde: *Bedømmernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen* samt *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen.*
Note: bedømmere: N = 2.946, ledere: N = 370.

I fokusgruppeinterviewene med ledere og bedømmere er der forskellige perspektiver på det gamle 13-tal. Nogle af bedømmerne og lederne savner at kunne belønne den exceptionelle præstation, hvor den studerende gør noget ud over det sædvanlige. De beskriver 12-tallet som en meget bred karakter, der kan dække over meget forskellige præstationer. De savner at kunne differentiere og nuancere bedømmelsen af de studerende, der toppræsterer, sådan at man kan skelne mellem de studerende, der får 12, fordi de ikke havde nogle væsentlige fejl og mangler og lever op til målbeskrivelsen, og de studerende der får 12, fordi de leverede en hel særegen præstation. En bedømmer og en leder udtrykker det således:

> Jeg kunne godt tænke mig en ekstra topkarakter. Jeg sidder hvert eneste år og kunne tænke mig at give 13. Det er på et helt andet niveau, og der er det ærgerligt, at man ikke kan give dem det med ud i verden.
>
> Bedømmer, videregående uddannelse

> Der lå en signalværdi i 13-tallet, fordi det var ekstraordinært. En udmærkelse ville ikke have samme signalværdi, som 13-tallet havde. Det handler jo også, om de signaler vi sender til de studerende. Alle får 12, så hvordan kan man udmærke sig nu? Hvad skal man egentlig stræbe efter?
>
> Leder, videregående uddannelse

Denne leder berører et perspektiv, som også var udbredt i fokusgruppeinterviews med både bedømmere og ledere, nemlig at en exceptionel karakter kan være med til at motivere de studerende til ikke blot at målopfylde, men at yde en ekstra indsats.

For andre bedømmere og ledere handler det ikke om, at de savner en egentlig karakter for den exceptionelle præstation, såsom 13-tallet, men at de på anden vis gerne ville kunne markere, når den studerende leverer en exceptionel præstation. En bedømmer fortæller fx:

> Man kunne godt bruge plus. Det skal ikke tælle med. Vi skal være sammenlignelige med det internationale karaktersystem.
>
> Bedømmer, videregående uddannelse

Endelig er der bedømmere, der ikke savner 13-tallet eller på anden måde at kunne differentiere 12-tallet blandt andet af hensyn til international sammenlignelighed. En bedømmer siger:

> Hvis Danmark var sin egen verden, var det fint at få 13-tallet, men fra hvad jeg kunne forstå, stod flere af vores elever med den gamle skala dårligere i et internationalt miljø. Vores kandidater skal kunne stå stærkt i en international konkurrence. Det er jo meningen, at der skal være flere, der får 12, fordi det både dækker 11 og 13.
>
> Bedømmer, videregående uddannelse

Bedømmerne er med andre ord opmærksomme på en af de ulemper en exceptionel karakter som 13-tallet var på den gamle skala, da den stillede de studerende, som ville til udlandet og studere, i en dårligere konkurrencesituation.

Sværere for arbejdsgivere at identificere de dygtige

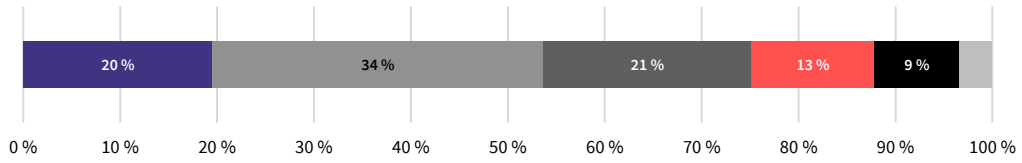
Interviews med arbejdsgiverorganisationerne viser, at selvom ansættelse ikke blot er baseret på karakterer, så spiller karakter en rolle – særligt ved ansættelse af nyuddannede – og det særligt i virksomheder, hvor de har mange ansøgere. Arbejdsgiverorganisationer ønsker ikke umiddelbart 13-tallet tilbage, men med stigningen i andelen, som får 12, oplever de, at deres medlemsvirksomheder har stadigt sværere ved at skelne mellem de studerende og identificere de ekstraordinært dygtige. De peger derfor på, at det er centralt for dem, at inflation i karaktererne modarbejdes, så 12-tallet reelt kun tildeles i omegnen af 10 % af de bestående præstationer, som Karakterkommissionen beskrev det. Det gælder også ved de afsluttende opgaver.

Også blandt lederne på universitetsuddannelserne er der en mindre andel, som ikke mener, at skalaen i tilstrækkelig grad understøtter, at de kan identificere de dygtigste studerende i forbindelse med forskerrekruttering, jf. figur 5.4. Som det fremgår, er 21 % af lederne på universitetsuddannelserne meget uenige eller uenige i, at den eksisterende karakterskala i tilstrækkelig grad understøtter identifikation af de dygtigste studerende i forbindelse med instituttets forskerrekruttering.

FIGUR 5.4

Skalaen understøtter i tilstrækkelig grad, at vi på mit institut kan identificere de dygtigste studerende i forbindelse med forskerrekruttering (ledere på universitetsuddannelserne)

Hvor enig eller uenig er du i følgende udsagn? -Skalaen understøtter i tilstrækkelig grad, at vi på mit institut kan identificere de dygtigste studerende i forbindelse med forskerrekruttering



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen*.
Note: N = 202.

5.4 Øvrige muligheder for identifikation af særligt dygtige studerende

Bekendtgørelse om talentinitiativer på de videregående uddannelser på Uddannelses- og Forskningsministeriets område (talentbekendtgørelsen) giver en række muligheder for at belønne særligt dygtige studerende med udmærkelse på eksamensbeviset, anerkendelse af ekstracurriculære aktiviteter på eksamensbeviset eller ekstra ECTS-point. Generelt er det de interviewede lederes holdning, at mulighederne i talentbekendtgørelsen er gode, men at de ikke har set eksempler på dem i brug.

Fokusgruppeinterviewene med bedømmere indeholder også kun få eksempler på udnyttelse af mulighederne.

Arbejdsgiverorganisationerne efterspørger øget brug af de eksisterende muligheder. Tilbagemeldingen fra deres medlemsvirksomheder er, at de ikke oplever, at mulighederne bliver udnyttet i dag, men at der er en generel interesse blandt virksomheder for en øget brug af mulighederne, dog uden at der går inflation i brugen.

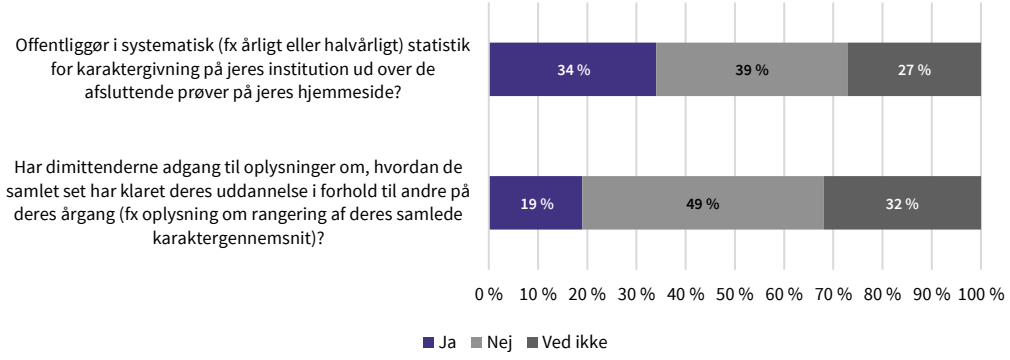
Ud over mulighederne i talentbekendtgørelsen kan uddannelsesinstitutionerne benytte andre metoder til identifikation af særligt dygtige studerende, herunder fx offentliggørelse af statistik over karaktergivning på et givent fag/en given uddannelse.

Figur 5.5. viser, at 19 % af lederne på de videregående uddannelser angiver, at dimittenderne har adgang til oplysninger om, hvordan de samlet set har klaret deres uddannelse i forhold til andre på deres årgang (fx oplysning om rangering af deres samlede karaktergennemsnit). Samtidig svarer 34 % af lederne, at der på en hjemmeside systematisk (fx årligt eller halvårligt) offentliggøres statistik for karaktergivning på institutionen ud over de afsluttende prøver.

Interviews med arbejdsgiverorganisationerne gav ikke noget klart svar på, om deres medlemsvirksomheder ville bruge statistik om karaktergivning i deres rekrutteringsarbejde.

FIGUR 5.5

Brug af øvrige muligheder for identifikation af særligt dygtige studerende, %



Kilde: *Ledernes perspektiver på 7-trins-skalaen – Bilagsrapport til Evalueringen af 7-trins-skalaen*.
Note: N = 370.

6 Skalaens internationale anvendelighed

6.1 Baggrund – 7-trins-skalaens understøttelse af international anvendelighed

International anvendelighed var et Karakterkommissionens fem centrale krav til en moderne karakterskala[5]. Danmark implementerede i 2001 store dele af anbefalingerne fra det europæiske ECTS-system, som havde til formål at skabe transparens med hensyn til uddannelserne.

Den danske karakterskala blev dog fremhævet som en fortsat hindring for de studerendes internationale mobilitet. Især undtagelseskarakteren 13 i 13-skalaen blev i den sammenhæng fremhævet som en udfordring. Da udenlandske uddannelsesinstitutioner ofte ikke forstod, at 13-tallet var en undtagelseskarakter, risikerede danske studerende, der søgte optagelse på gode institutioner i udlandet, at blive mødt med en forventning om, at de skulle have lutter 13-taller for at blive optaget.

Det affødte et politisk ønske om at understøtte, at danske studerende blev stillet mere rimeligt, når de søgte optagelse på udenlandske uddannelser. Det var derfor en del af Karakterkommissionens kommissorium, at kommissionen skulle afdække og inddrage "hvilke karaktersystemer, der benyttes i internationale sammenhænge, herunder ECTS-skalaen". ECTS-skalaen indgik den gang i ECTS-systemet som et redskab til at skabe transparens netop i forbindelse med karaktergivning.

Karakterkommissionen vurderede, at 13-skalaen ikke var kompatibel med ECTS-skalaen eller de nordamerikanske karakterskalaer[6]. Kommissionen byggede 7-trins-skalaen op efter den europæiske ECTS-skala (se appendiks B for yderligere beskrivelse). Det skulle understøtte, at karaktererne i 7-trins-skalaen kunne oversættes en-til-en til karaktererne i ECTS-skalaen og derigennem gøre dem internationalt anvendelige.

En forudsætning for, at de danske karakterer kan oversættes en-til-en til ECTS-skalaen, er, at de nogenlunde følger den forventede fordeling (10 % – 12, 25 % – 10, 30 % – 7, 25 % – 4, 10 % – 2). Oversættelse via ECTS-skalaen kan derfor være misvisende.

I 2009 blev ECTS-skalaen dog skrevet ud af anbefalingerne i ECTS-systemet og erstattet af nye anbefalinger til at understøtte gennemsigtighed og formidling af karakterer på tværs af de nationale karakterskalaer. Anbefalingerne er nu i stedet, at landene skal offentliggøre karakterfordelinger for

[5] Se Karakterkommissionen, 2004, kapitel 6, og boks i denne evaluerings indledning.

[6] Karakterkommissionen, 2004, inklusive henvisning til den meste udbredte karakterskala i USA med karaktererne A, B, C, D (bestået) og F (ikke-bestået) evt. med brug af "+" og "–" for særligt gode eller dårlige præstationer inden for karaktererne.

de bestående karakterer samt beskrivelser af deres karakterskalaer, og hvordan de bruges. Oversættelse og formidling af karakterer anbefales altså at ske direkte mellem skalaerne ved hjælp af karakterfordelinger uden et mellemled.

Den første evaluering af 7-trins-skalaen konkluderede i 2013, at karaktererne efter 7-trins-skalaen var internationalt anvendelige og ikke hindrede international mobilitet. Evalueringen pegede også på, at ECTS-skalaen blev brugt som oversættelsesredskab. Den belyste dog ikke betydningen af de nye anbefalinger i ECTS-systemet.

Kapitlet afdækker derfor i lyset af udviklingen spørgsmålet "Er 7-trins-skalaen internationalt anvendelig?".

Kapitlet belyser i den forbindelse, hvorvidt det stadig er relevant at understøtte, at karakterskalaen skal være international relevant, hvordan 7-trins-skalaens internationale relevans opleves i praksis samt betydningen af de nye anbefalinger i ECTS-systemet med hensyn til understøttelsen af de danske karakterers anvendelighed i udlandet.

Kapitlet fokuserer på de videregående uddannelser og bygger på:

- Enkeltinterviews med otte studerende, der har taget dele af en uddannelse eller en hel uddannelse i udlandet
- To fokusgruppeinterviews med internationale medarbejdere fra de videregående uddannelsesinstitutioner
- Interview med professor Anthony Vickers ved University of Essex, som var tovholder på udviklingen af det europæiske projekt Egracons til at understøtte oversættelsen af karakterer på tværs af karakterskalaer
- Rundspørge om karaktergivning i andre lande via det europæiske Eurydice-samarbejde
- ECTS Users' Guide fra 1998, 2004, 2009 og 2015
- Desktopresearch-systemer til oversættelse af karakter i forbindelse med optag udarbejdet af Styrelsen for Forskning og Uddannelse.

ECTS-skalaen

ECTS-skalaen blev udviklet som en del af ECTS-systemets (European Credit Transfer and Accumulation Systems) anbefalinger til at understøtte Bologna-processens målsætning om international sammenlignelighed af uddannelser. Skalaen er et værktøj til at oversætte karakterer internationalt på tværs af forskellige karakterskalaer. Den indgik i ECTS-systemets anbefalinger frem til 2009.

Skalaen er et værktøj til at oversætte karakterer fra en national karakterskala til en anden. Det består af fem bestående karakterer og to ikke-bestående karakterer. De bestående karakterer skal oversættes efter en fast fordeling, så de bedste 10 % af præstationerne får A, de næstbedste 25 % får B, de næste 30 % får C, de næste 25 % får 25 og de 10 % svageste præstationer får E.

Der findes ikke en karakterfordeling for de ikke-bestående karakterer.

Oversættelse ved hjælp af ECTS-skalaen kræver, at de nationale karakterer giver tilstrækkeligt detaljeret information om karakterfordelingen til, at karaktererne kan indplaceres i fordelingen.

Tabellen viser, hvordan de danske karakterer, jf. karakterbekendtgørelserne, oversættes til ECTS-skalaen på de danske eksamensbeviser:

-3	00	02	4	7	10	12
Fx	F	E	D	C	B	A

Opkoblingen til ECTS-skalaen går igen i de officielle engelske beskrivelser af 7-trins-skalaen.

6.2 Er spørgsmålet om international anvendelighed stadig relevant?

Tre faktorer var drivende for Karakterkommissionens fokus på, at den danske karakterskala skulle være internationalt anvendelig:

- Flere studerende, der tager hele eller dele af uddannelser i andre lande (international mobilitet)
- Store internationale forskelle i karaktersystemerne
- Et stort internationalt fokus på gennemsigtighed i uddannelsessystemerne, herunder anbefalingen af ECTS-skalaen i ECTS-systemet.

Afsnittet viser, at den internationale mobilitet er steget siden 7-trins-skalaens implementering, og at der er ikke sket en konvergens af de internationale karakterskalaer. Den danske karakterskala skiller sig blandt andet ud ved at være en talskala i stedet for en bogstavskala, ved at have talværdier, der ikke er bundet op på et bagvedliggende antal point i en given eksamen, samt ved at være en ikke-lineær skala med spring mellem talværdierne. Dette tyder på, at det stadig er relevant at understøtte, at de danske karakterer er anvendelige internationalt.

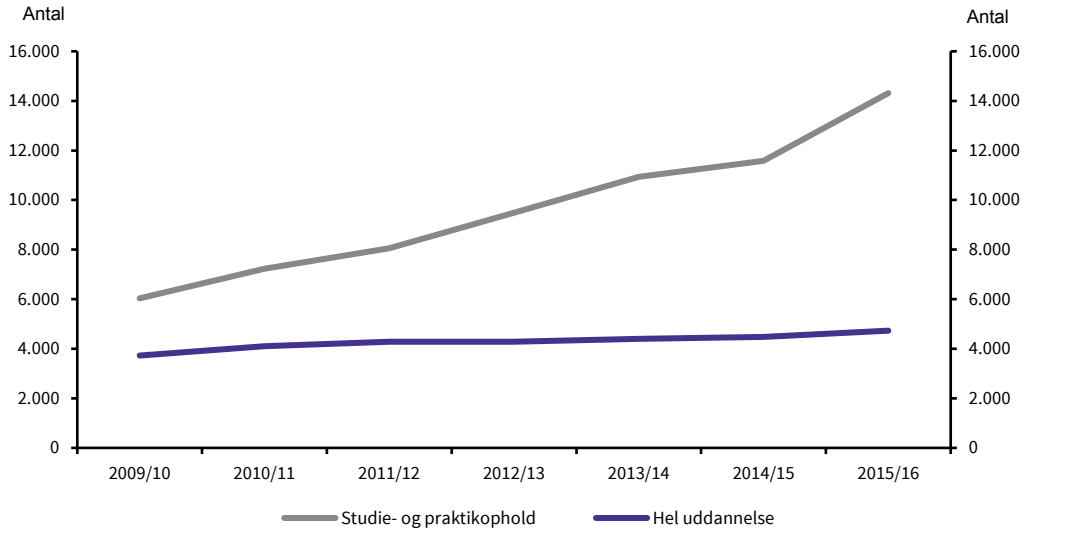
6.2.1 International mobilitet og karakterformidling

Hvert år rejser et stort antal danske studerende ud for at tage hele eller dele af deres uddannelse i udlandet. Tilsvarende modtager Danmark hvert år studerende fra udlandet enten på studieophold eller som optagne på hele uddannelser i Danmark.

Antallet af danske studerende, der tager uddannelse i udlandet, er steget de seneste år. Især antallet af praktik- og studieophold i udlandet som del af en dansk uddannelse er steget, jf. figur 6.1.

FIGUR 6.1

Udviklingen i antallet af danske studerende, der studerer i udlandet, 2009/10-15/16

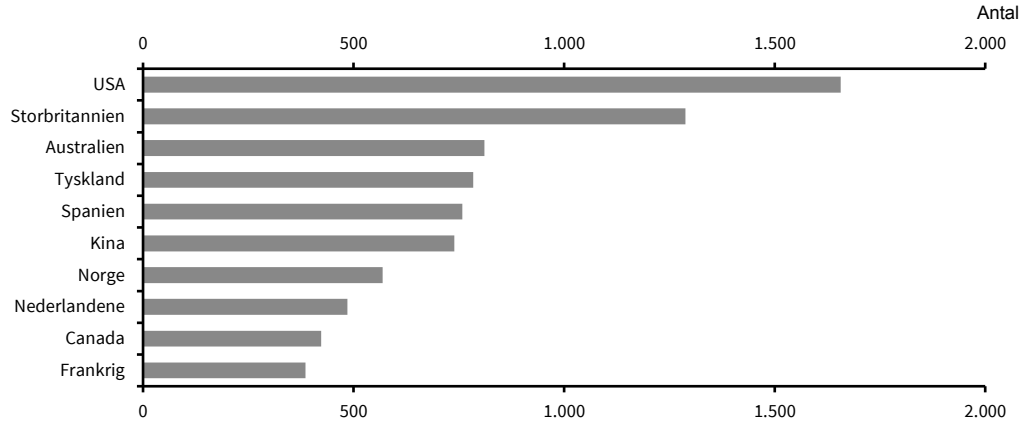


Kilde: Uddannelses- og Forskningsministeriet på baggrund af data indsamlet af Danmarks Statistik (studie- og praktikophold) og SU-data (hel uddannelse).
Note: Studerende på en hel uddannelse i udlandet er opgjort på kalenderår. Data for 2015/16 dækker således kalenderåret 2016.

Danske studerende i udlandet modtager oftest SU til at studere i USA, Storbritannien eller Australien. Det er også de lande, som de oftest søger udlandsstipendier til selvarrangerede ophold i, jf. figur 6.2.

FIGUR 6.2

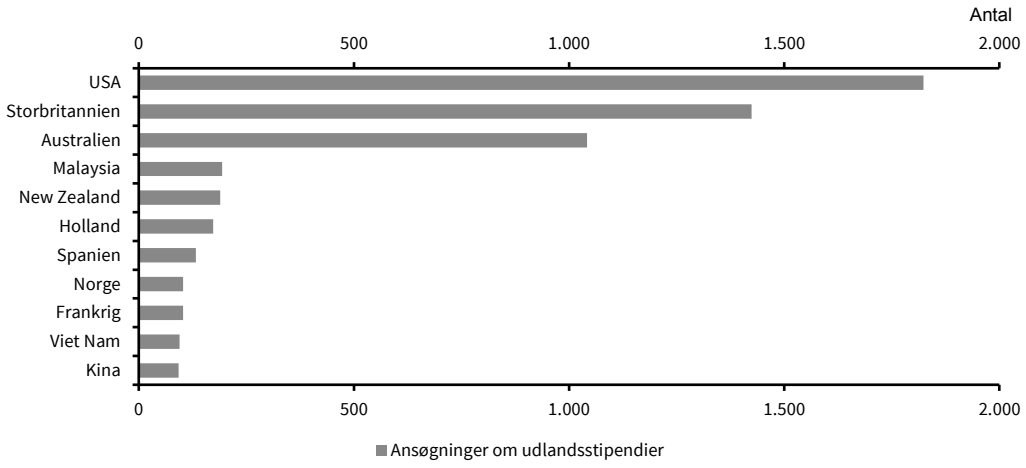
Antal rejsende studerende med SU til uddannelse i udlandet i 2015/16



Kilde: Styrelsen for Forskning og Uddannelse på grundlag af data fra Danmarks Statistik.

FIGUR 6.3

De ti lande, som flest studerende søgte udlandsstipendium til hel uddannelse eller studieophold i, 2015/16



Kilde: Styrelsen for Institutioner og Uddannelsesstøtte.
Note: Opgørelsen viser antal unikke studerende, som har søgt stipendium til uddannelse i et givet land.

Mange studerende studerer dog også i europæiske lande som Tyskland, Spanien og Norge. Spørgsmålet om international anvendelighed opstår, fordi karakter- og bedømmelsessystemer er et af de områder, hvor der er særligt stor diversitet på tværs af lande og i flere lande også på tværs af institutioner.[7] Forskellene er illustreret med en beskrivelse af det britiske, norske, svenske, tyske

7 Se Karran, 2005, for en sammenligning af de europæiske karaktersystemer i 2003.

og franske karaktersystem i tabel 6.1. De fem systemer adskiller sig med hensyn til både karakterskalaernes opbygning med antallet af trin, trinenes betegnelser med bogstaver, tal eller andet, og om de bruger eksplicitte procentgrænser for, hvor meget af en prøve der skal være bestået for at opnå en given karakter, etc. Anvendelsen af skalaerne er også forskellig både med hensyn til brugen af skalaerne, fx brugen af topkarakteren, og hvilken præstation der kræves for at bestå, og med hensyn til rammerne i systemet, det være sig vedrørende institutionernes autonomi i forbindelse med brugen af karakterskala eller transparens i karaktergivningen via statistik.

TABEL 6.1

Karaktersystemer i europæiske lande

	Norge	Sverige	Tyskland	Storbritannien	Frankrig
Skala	A-B-C-D-E-F	Mest anvendte skala: bestået med ære/bestået/ikke bestået	1-2-3-4-5	Fire beståede karakterer (1. klasse, 2. klasse øverst, 2. klasse nederst og 3. klasse). En ikke-bestået karakter	Mest udbredte: 0-20, (op til 20 point for præstationen med mulighed for halve point).10-20 point er bestået.
Anvendelse	Alle videregående uddannelser	Institutionerne vælger skala. P.t. fem i brug inkl. ECTS-skalaen (dog absolut)	Både videregående uddannelser og ungdomsuddannelser	Kun videregående uddannelser	Institutionerne vælger skala. 0-20 er mest udbredt
Topkarakterens andel af beståede karakterer	13,5 %	Ca. 30 % (ved den mest anvendte skala, men varierer afhængigt af skala og fagligt område)	Ca. 19 %	26 %	-

Kilde: Survey via Eurydice foretaget af Styrelsen for Forskning og Uddannelse og Karran, 2005.
Note: Generelt for de europæiske karakterskalaer gælder det, at de typisk bruges absolut, evt. knyttet til eksplicitte læringsmål. Både tal og bogstavskalaer forekommer, de fleste har mindst fem beståede trin, og de fleste har en eksplicit fastsættelse af procentandelen af rigtige eller antallet af point, der skal opnås for at opnå en given karakter – satserne kan variere mellem fag og institutioner. Skalaerne har typisk lineære spring mellem karaktererne.

Med undtagelse af princippet om, at karakterer skal gives absolut (ud fra graden af opfyldelse af læringsmål/kriterier), er der ikke en bestemt måde at give karakterer på, herunder heller ikke en bestemt karakterskala, som går igen og er typisk set i internationalt perspektiv.

Selvom en del lande bruger bogstavskalaer, som kan minde om den mest anvendte version af den amerikanske skala, anvendes skalaerne ofte forskelligt, fx med forskellige krav til, hvor stor en andel af de mulige point der skal være opnået for at bestå eller for at opnå topkarakter. Direkte sammenligning af karaktererne kan derfor være misvisende[8].

8 Karran, 2004, 2005; Søndergaard, 2008.

Fokusgrupperne med de internationale medarbejdere og interviewene med danske studerende, som har søgt uddannelse i udlandet, peger på, at spørgsmålet om de danske karakterers internationale anvendelighed oftest handler om, hvorvidt de danske karakterer kan formidles retvisende i udlandet. Især tre kontekster fremhæves som relevante:

- Danske studerende, som søger om optagelse på en hel uddannelse i udlandet eller søger studieophold i udlandet uden en udvekslingsaftale, har brug for at formidle deres karakterer for at komme i betragtning til optagelse.
- Danske uddannelsesinstitutioner har brug for at kunne forklare den danske karakterskala, når de skal indgå udvekslingsaftaler med institutioner i udlandet, herunder når de skal indgå aftaler om, hvilke krav de skal stille til deres egne studerende, hvis de skal kunne benytte sig af aftalerne.
- Danske institutioner og udenlandske studerende skal kunne formidle de danske karakterer til parter i udlandet, når de udenlandske studerende kommer tilbage til deres hjemland. Det drejer sig især om formidling til hjemmeinstitutionen i forbindelse med udveksling og dokumentation for deres aktivitet i forbindelse med modtagelse af udlandsstipendier.

Især på erhvervsakademierne og professionshøjskolerne og blandt nogle studerende fremhæves desuden en problematik om, at det kan være svært at bruge karakterer fra udlandet til at opnå uddannelse i Danmark. Udfordringen her er dog i mindre grad den danske skala og i højere grad, at oversættelse af karakterer ses som komplekst og ressourcekrævende. Et gennemgående element er, at formidling er centralt, når de danske karakterer skal bruges i udlandet.

> Der er brug for den der dialog hele tiden.
> International medarbejder, universitet

> Vi har sådan nogle standardforklaringer [af karaktererne], vi kan sende af sted [...] Man er nødt til at kunne snakke sammen om det, ellers går det galt.
> International medarbejder, universitet

Samlet set er der dog helt generelt en skepsis over for at søge at gøre karakterer og karakterskalaer internationalt kompatible. Både studerende og medarbejdere påpeger, at de internationale forskelle i uddannelsessystemerne er så markante, at det aldrig vil være muligt at skabe fuldstændig kompatibilitet.

> Man ville aldrig nogensinde få noget, som er absolut kompatibelt, fordi uddannelserne, læringskulturen og karaktergivningen generelt er forskellig landene imellem.
> International medarbejder, universitet

De internationale medarbejdere påpeger, at en ny skala vil skabe forvirring og medføre, at formidlingsarbejdet skal genstartes fra bunden. En ny skala ses derfor i denne kontekst som et tilbageskridt for den internationale anvendelighed af de danske karakterer. Flere af de internationale medarbejdere nævner, at hvis det blev besluttet at ændre skalaen, ville det være hensigtsmæssigt, at den danske skala lagde sig op af en udenlandsk skala, som danske institutioner i høj grad samarbejder med. Her gives der dog forskellige eksempler som Sverige, Tyskland, USA (hvor der anvendes flere skalaer) eller Storbritannien, alt efter hvem den pågældende institution samarbejder mest med.

> Det ville være et større internationalt problem, hvis vi allerede nu indførte en ny skala. Fordi det er skide besværligt. Vi sidder og roder rundt i det med dem, der har indført nye skalaer og pointsystemer. Det vigtigste er ikke at lave om igen nu.
>
> International medarbejder, kunstnerisk institution

Formidling og samtale ses altså som et vilkår ved internationalt samarbejde og mobilitet i uddannelse, uanset karakterskala.

6.3 7-trins-skalaens internationale anvendelighed

Både studerende og internationale medarbejdere beskriver 7-trins-skalaen overordnet set som internationalt anvendelig. Skalaen opleves som nemmere at formidle til udlandet end 13-skalaen og opleves ikke som en hindring for international mobilitet.

> Det her er meget bedre end den gamle skala. [...] At gå tilbage er ikke en mulighed. Vi skal bare et skridt længere frem.
>
> International medarbejder, erhvervsakademi

> Skalaen er mere klar at forklare over for udlandet end 13-skalaen. Jeg synes ikke, at problemet så meget er 7-trins-skalaen. Det er mere ECTS-skalaen.
>
> International medarbejder, kunstnerisk institution

Både studerende og internationale medarbejdere oplever, at 7-trins-skalaen kan formidles i forbindelse med international mobilitet. Skalaens opbinding på ECTS-skalaen fungerer som en god ramme for formidling af danske karakterer ved samarbejdsaftaler med europæiske partnere. Det er imidlertid samtidig opbindingen på ECTS-skalaen, der beskrives at volde størst problemer i forbindelse med formidlingen af danske karakterer uden for Europa. Dette får flere studerende og internationale medarbejdere til at efterspørge værktøjer fra ministeriet som en bedre engelsk guide til 7-trins-skalaen eller omregningstabeller ligesom dem, der eksisterer for udenlandske eksamensbeviser ved ungdomsuddannelser (Eksamenshåndbogen), der kan muliggøre en mere retvisende og ensartet oversættelse eller formidling af karakterer. De understreger dog, at det er centralt, at der ikke stilles krav om ekstra administrative procedurer eller brug af faste, mekaniske metoder til omregning.

6.3.1 Formidling af 7-trins-skalaen i udlandet

De danske studerendes studieophold arrangeres typisk via uddannelsesaftaler indgået mellem deres danske institution og værtsinstitutionen i udlandet. I den sammenhæng er det den danske institution, der varetager formidlingen af den danske karakterskala til udlandet, og den danske institution, der sikrer, at de danske studerende lever op til evt. karakterkrav i udlandet. Desuden har den danske institution en opgave i at understøtte, at den udenlandske institution forstår de karakterer, som dens studerende får i Danmark.

De internationale medarbejdere fra uddannelsesinstitutionerne fortæller i fokusgruppeinterviews, at der ved indgåelse af samarbejdsaftaler følger et selvstændigt formidlingsarbejde med omkring karaktererne efter 7-trins-skalaen. De beskriver, at skalaens spring omkring midterkaraktererne samt de to dumpekarakterer vækker en umiddelbar undren og kan behøve forklaring. De internationale medarbejdere beskriver dog, at det største forklaringsarbejde opstår i kraft af 7-trins-skalaens opbinding på ECTS-skalaen.

> Vi bruger ikke ECTS-skalaen i udvekslingssammenhæng. Den generer mest. Jeg bruger meget tid på at forklare USA, at den ikke har noget med A, B og C at gøre. Jeg ville helst undgå den.
>
> International medarbejder, erhvervsakademi

> Jeg synes ikke, at det er 7-trins-skalaen, der er problemet, men ECTS-skalaen, fordi den har det her problem med *letter grades*, som kan forvirre. Inden for Europa fungerer det fint, der er det dejligt, at vi har fået en fælles skala, som vi kan snakke ind i, men det er uden for Europa, at det er et problem.
>
> International medarbejder, kunstnerisk uddannelsesinstitution

De internationale medarbejdere beskriver, at det ofte er nemmere at tage udgangspunkt i 7-trins-skalaen end i ECTS-skalaen, når danske karakterer skal formidles i forbindelse med samarbejdsaftaler. De internationale medarbejdere beskriver i den forbindelse, at formidling af karakterskalaen er en del af den samlede formidlingsopgave, der altid vil ligge til grund for at indgå samarbejdsaftaler med udenlandske uddannelsesinstitutioner. Andre forhold er også afgørende for indgåelse af samarbejdsaftaler. Blandt andet nævner de internationale medarbejdere det, at det er vigtigt, at der er overensstemmelse mellem indholdet i kurset på samarbejdsinstitutionen og det, som den studerende skal lære på uddannelsen herhjemme.

> Det handler om, hvor modtagelig modtagerinstitutionen er for forklaringer. [...] Det er ikke så meget anderledes end kultur, kan man næsten sige. Man er afhængig af en eller anden oversættelse, når man er to parter, og så er der en eller anden åbenhed for forståelse, eller også er der ikke en åbenhed for forståelse. Skalaen som sådan giver ikke bare *per default* anledning til problemer. Det ligger i oversættelsen, og det ville du jo skulle lige meget hvad.
>
> International medarbejder, universitet

Danske studerende, som rejser ud på selvarrangerede udlandsophold eller gennemfører en hel uddannelse i udlandet, skal derimod selv formidle deres karakterer til de institutioner, som de søger om optagelse på. De danske studerende beskriver ligesom de internationale medarbejdere beskriver, at ECTS-skalaen kan skabe unødig forvirring uden for Europa, men at den inden for Europa overordnet er et velfungerende formidlingsværktøj. Flere af de danske studerende, der søger om optagelse på hele uddannelser i udlandet, bliver optaget på britiske universiteter, som har så stort et internationalt optag, at de har landespecifikke krav til minimumsgennemsnit efter de respektive nationale karakterskalaer. Derfor bliver oversættelse af karaktererne i disse tilfælde ikke relevant. Ved optagelse på udenlandske universiteter er der dog ofte også andre faktorer end karakterer, der bliver taget i betragtning. Motiverede ansøgninger, udtalelser fra professorer og tests indgår flere steder i optagelsesprocessen.

> Der er meget formalia, og det er meget besværligt. Men selve det med karaktererne er ret nemt. Jeg fik underskrevet mit karakterblad og sendte det med min ansøgning. Der var både de danske karakterer og oversættelsen til A, B [ECTS-skalaen]. Men de har nok taget stilling til det danske, eftersom de har et minimumskrav til de danske karakterer.
>
> Studerende på udveksling under samarbejdsaftale og hel uddannelse i udlandet

De studerende, der søger om optagelse på hele uddannelser og aktivt skal bruge deres danske karakterer i en optagelsesproces, problematiserer i højere grad uigennemskueligheden ved ansøgnings- og optagelsesprocessen i udlandet, end de problematiserer den internationale anvendelighed af karaktererne på deres eksamensbevis. De efterspørger derfor i højere grad at kunne blive vejledt i, hvad der er afgørende i ansøgnings- og optagelsesprocesserne i de mest populære lande for uddannelse i udlandet.

6.3.2 Formidling af 7-trins-skalaen til udenlandske studerende i Danmark

Internationale medarbejdere fra uddannelsesinstitutionerne beskriver, at der også er en formidlingsopgave forbundet med 7-trins-skalaen, når udenlandske studerende kommer til Danmark. Medarbejderne beskriver, at de to dumpekarakterer og springene omkring midterkaraktererne vækker undren, men kan forklares. Udfordringen beskrives i højere grad at ligge i 7-trins-skalaens opbinding på ECTS-skalaen, som kan skabe udfordringer for de udenlandske studerende, når de vender hjem. Nedenfor gives et eksempel på en gæstestuderende, der risikerede ikke at kunne gennemføre sin uddannelse i USA grundet et 4-tal, som vedkommende havde fået under sit ophold i Danmark:

> Vi har haft en gæstestuderende fra USA, som fik et 12-tal. Det var jo ikke noget problem. Men så fik hun et 4-tal også. Det står jo som D. Hendes universitet kræver, og det stod også i papirerne, som hun var blevet optaget på, at for at hun kunne gennemføre sin uddannelse, skulle hun have mindst C i snit. De ville ikke godkende et 4-tal som et C derovre. Og vedkommende har jo smidt penge ud og er hamrende tosset.
>
> International medarbejder, universitet

De internationale medarbejdere giver flere eksempler på, at udenlandske studerende oplever udfordringer, når de vender hjem, hvis afsenderinstitutionen kun accepterer A eller B i karakter. Her kan et 7-tal som oversættes til et C ved ECTS-skalaen altså skabe problemer. Derudover beskrives også en udfordring med, at mange danske kurser prøves med bedømmelsen bestået/ikke bestået, da flere afsenderinstitutionen kræver, at der skal foreligge en karakter i de fag, deres studerende tager udlandet.

Hvis det skal understøttes, at danske studerende ikke stilles uretmæssigt dårligere end studerende fra andre lande, og at det i Danmark er muligt at forstå og evt. oversætte karakterer fra udlandet til den danske skala, er det derfor stadig relevant at arbejde for at understøtte, at de danske karakterer kan formidles og oversættes til udenlandske skalaer. Og tilsvarende er det vigtigt for de studerende, som kommer hertil og evt. betaler for en uddannelse, at de kan forklare deres karakterer, når de kommer hjem.

6.3.3 Oversættelse af udenlandske karakterer i Danmark

De internationale medarbejdere fra de videregående uddannelsesinstitutioner beskriver, at der ikke eksisterer en praksis for at oversætte udenlandske karakterer til danske eksamensbevis. På tværs af institutionerne beskrives der samstemmende, at oversættelse af karakterer ikke er en eksakt videnskab, og at der på nuværende tidspunkt ikke eksisterer en praksis, der sikrer, at det bliver konsistent, meningsfuldt og retvisende. En international medarbejder fra en kunstnerisk uddannelsesinstitution udtaler:

> Jeg har det sådan, at jeg er virkelig glad for, at vi ikke skal sidde og omregne, når de kommer hjem. For det kan nærmest kun gå galt. Så jeg er virkelig glad for, at vi ikke skal sidde og omregne karakterer. En del af internationaliseringen er også at anerkende, at vi er forskellige.
>
> International medarbejder, kunstnerisk institution

Blandt institutionerne og de studerende ses der både fordele og ulemper ved den manglende oversættelse af de udenlandske karakterer. Særligt medarbejderne fra erhvervsakademierne ser den manglende oversættelse som en ulempe for den studerende, da den relativt korte uddannelse betyder, at den studerende står med få karakterer på sit eksamensbevis ved uddannelsens afslutning. Det kan således være et skrøbeligt grundlag for den studerende at søge videre på en top-up-uddannelse på baggrund af. Enkelte universitetsstuderende nævner også selv en uretfærdighed i

at blive straffet for at tage udvekslingsophold i stedet for valgfag hjemme på grund af den manglende mulighed for at tælle gode karakterer i udlandet med i det samlede gennemsnit. Det modsatte perspektiv er dog også udbredt blandt en del af de studerende. De mener, at usikkerheden i oversættelsen vil være så stor, at det kan ende med at være den studerende til last. Dette er også den gennemgående holdning blandt medarbejderne på universiteterne og de kunstneriske uddannelsesinstitutioner.

> Jeg synes bare, at det er lidt farligt at lave den der med bare at se på en eller anden skala og så bare regne det om, for er det ikke inderne, der har det sådan, at det er ret nemt at få en god karakter, mens det er ret svært i Kina? Man skal kende sine lande, og man skal kende sine institutioner, nærmest. For der er nogle institutioner, hvor det kan være svært at få en god karakter.
> International medarbejder, universitet

Uddannelsesinstitutionerne beskriver, at der på nuværende tidspunkt eksisterer forskellig praksis fra institution til institution, men også internt på den enkelte institution med hensyn til omregning og oversættelse af karakterer. Der gives eksempler på forskellige metoder ved vurdering af studerende, der søger optag, understøttelse af egne dimittender, der søger videre eller ved tildeling af udlandsophold til egne studerende ud fra tidligere karakterer. Uddannelsesinstitutionerne efterspørger et centralt udarbejdet værktøj til oversættelse af karakterer mellem videregående uddannelser internationalt, hvis sammenligneligheden skal sikres. På nuværende tidspunkt har ministeriet kun udarbejdet et omregningsværktøj til brug for udenlandske eksamensbeviser fra ungdomsuddannelser i forbindelse med optag på videregående uddannelse. Som illustreret i citatet ovenfor udtrykkes der dog også en generel skepsis over for omregningsværtøjer, da der er forskel fra land til land med hensyn til andelen af topkarakterer, og hvad der skal til for at udløse den.

6.4 De nye anbefalinger i ECTS-systemet og fremtidige perspektiver

Anbefalingerne inden for det europæiske samarbejde om videregående uddannelser tilsiger nu i højere grad, at sammenlignelighed mellem karakterskalaer og karakterer sikres gennem transparens i karaktergivningen. Det anbefales derfor nu at anvende karakterfordelinger fra de enkelte uddannelser som oversættelsesværktøj, frem for en-til-en-oversættelse mellem nationale skalaer ved hjælp af ECTS-skalaen. Dette kræver dog, at de danske uddannelsesinstitutioner gør karakterfordelingerne offentligt tilgængelige.

De studerende og de internationale medarbejdere kan se både fordele og ulemper ved transparens i karaktergivningen og anvendelse af karakterfordelinger som oversættelsesværktøj til karakterer. Enkelte universiteter har allerede haft en praksis med at anvende karakterfordelinger til at synliggøre de studerendes niveau eller placering på uddannelsen.

> Der har været en længere historie om, at det har ligget på vores intranet, men vi har haft et enormt karakterpres og karakterfokus. Så derfor har vi faktisk fjernet det, så det er noget, de aktivt skal komme ind og få på studieadministrationen. Så kunne man gå ind og se, hvordan man lå på sit samlede snit i forhold til sin årgang eller sit hold.
> International medarbejder, universitet

Som det fremgår ovenfor, bliver det dog problematiseret, at denne tendens kan være medvirkende til at skabe et stort fokus på karakterer og et oplevet pres for de studerende om at få gode karakterer. Denne bekymring italesætter de studerende ligeledes:

> Jeg synes, der er så meget fokus på karakterer i dag alligevel, at det måske bare ville være at putte en ekstra byrde på. [...] Måske bare noget bedre formidling af, hvad et 7-tal betyder i Danmark og så prøve at gøre forklaringerne overensstemmende i stedet for at prøve at gøre tallene overensstemmende.
>
> Studerende på selvarrangeret udvekslingsophold og hel uddannelse i udlandet

Den studerende påpeger, at bedre formidling er tilstrækkeligt til at skabe international anvendelighed af danske karakterer. Dette er ligeledes holdningen blandt en stor del af de internationale medarbejdere på uddannelsesinstitutionerne, der ikke ser behovet for karakterfordelinger som oversættelsesværktøj. Der eksisterer en frygt for, at det bliver pålagt institutionerne at gøre det, og at det vil medføre en administrativ byrde for institutionerne.

Flere medarbejdere beskriver dog også, at gennemsigtighed med hensyn til karakterfordelinger vil lette nogle samtaler i forbindelse med etableringen af nye samarbejdsaftaler med internationale partnere, fordi det skaber klarhed med hensyn til de studerendes faglige niveau.

> Jeg er mere pro-statistik. Jeg har arbejdet meget med USA og karakterparadigmer som deres, og de går meget op i, hvilken procentandel man ligger i. Det kan spare os for lange samtaler med nye samarbejdspartnere om, hvorvidt de studerende er dygtige nok.
>
> International medarbejder, erhvervsakademi

Dette perspektiv eksisterer også blandt studerende, der ser det som en fordel at kunne fremvise, at de er topstuderende, ved hjælp af karakterfordelinger.

> Man har også et stærkere argument, hvis man er i stand til at sige, at det godt kan være, at jeg ud fra din skala eller din forståelse af den skala, der findes i Danmark, ikke er dygtig nok til at komme ind på det her program. Men prøv at se her. Så skal de selvfølgelig være villige til at tage stilling til det også.
>
> Studerende på udveksling under samarbejdsaftale og hel uddannelse i udlandet

Både studerende og medarbejdere ser altså både fordele og ulemper ved de fremadrettede anbefalinger om offentliggørelse og anvendelse af karakterfordelinger til oversættelse af karakterer.

Appendiks A – Litteraturliste

Undervisningsministeriet (2004). *Betænkning om indførelse af en ny karakterskala til erstatning af 13-skalaen. Afgivet af karakterkommissionen, november 2004.* Betænkning nr. 1453, november 2004.

Dahl, B., Lien, E., & Lindberg-Sand, Å. (2009). Conformity or confusion? Changing higher education grading scales as a part of the Bologna Process: the cases of Denmark, Norway and Sweden. *Learning and Teaching: The International Journal of Higher Education in the Social Sciences, 2* (1), 39-79.

Karran, T. (2005). Pan-European Grading Scales: Lessons from National Systems and the ECTS. *Higher Education in Europe, 30* (1).

Karran, T. (2004). Achieving Bologna Convergence: Is ECTS Failing to Make the Grade?. *Higher Education in Europe, 24* (3).

EU-Kommissionen (2009). *ECTS Users' Guide*.

EU-Kommissionen (1998). *EUROPEAN CREDIT TRANSFER SYSTEM – ECTS USERS' GUIDE*.

EU-Kommissionen (2015). *ECTS Users' Guide*.

Søndergaard, B. D. (2008). International mobilitet med den nye danske karakterskala?. *MONA – Matematik- og Naturfagsdidaktik, 2*.

Appendiks B – 7-trins-skalaen og ECTS-skalaen

Da 7-trins-skalaen blev etableret i 2004, var der stort fokus på international mobilitet i Danmark såvel som internationalt. Forskellighederne i de europæiske landes uddannelsessystemer blev set som en potentiel hindring for, at studerende kunne bevæge sig fra land til land som led i deres uddannelse. Derfor blev ECTS-systemet (European Credit Transfer and Accumulation System) etableret i europæisk regi som en opfølgning på Bolognadeklarationen fra 1999. Formålet var at understøtte gennemsigtighed med hensyn til indholdet af uddannelsesprogrammer og med hensyn til de studerendes aktivitet på uddannelser på tværs af landene.

ECTS-systemet har fokus på tre områder:

- Beskrivelsen af kurserne, herunder læringsmål og arbejdsbelastning (ECTS-point),
- Udvekslingsaftaler (learning agreements)
- Dokumentationen for de studerendes uddannelsesaktivitet (transcript of records), herunder karakterer.

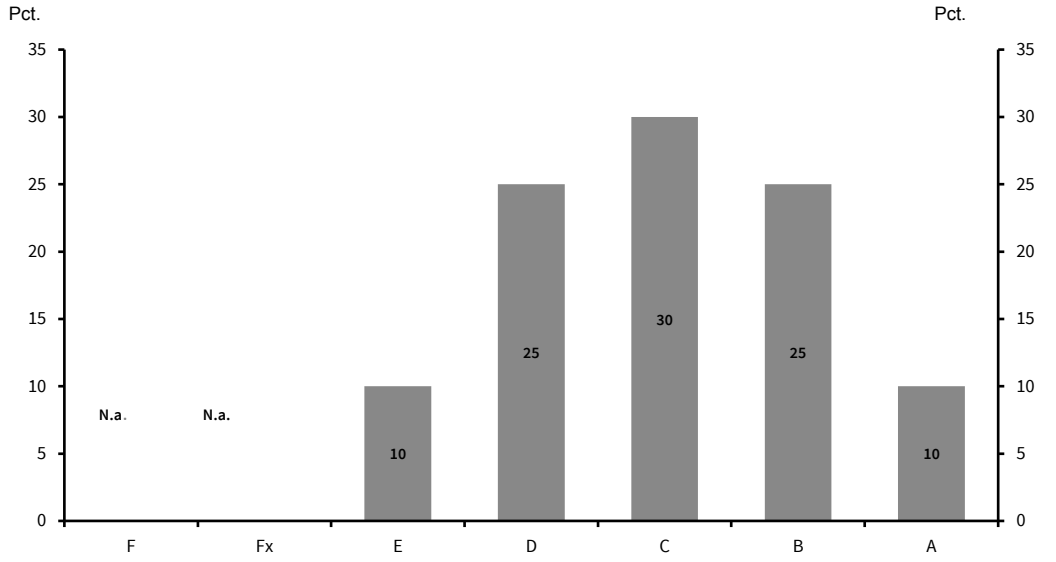
Der er stor forskel på karakterskalaer og karaktergivningspraksis i de enkelte europæiske lande; i nogle lande er der endda forskelle institutionerne imellem. Det gør det svært at sammenligne og formidle karakterer på tværs af landene. Derfor var karakterer en central del af det tredje punkt i ECTS-systemet.

ECTS-systemets oprindelige anbefaling var, at landene brugte den europæiske internationale karakterskala kaldet EGIS eller ECTS-skalaen til at skabe gennemsigtighed om karakterer. ECTS-skalaen var tænkt som et værktøj til at oversætte karakterer fra nationale skalaer til en fælles international karakterskala. Den var ikke tænkt til at erstatte de nationale skalaer.

ECTS-skalaen har fem bestående og to ikke-bestående karakterer. Det er en relativ karakterskala, dvs. at karaktererne tildeles ud fra de studerendes rangering i forhold til hinanden. Skalaen har en fast fordeling for, hvilken andel af de bestående resultater der skal have hver af de fem bestående karakterer. Der er ikke fastsat en forventet fordeling af de ikke-bestående karakterer, jf. figur B.1.

FIGUR B.1

Opbygning af European International Grading Scale – ECTS-skalaen



Kilde: Styrelsen for Forskning og Uddannelse på basis af ECTS Users' Guide 2004.
Note: Procentandelene angiver, hvor stor en andel af de bestående resultater der bør tildeles hver af de bestående karakterer efter ECTS-skalaen. Der er ikke fastsat en forventet fordeling for de ikke-bestående karakterer.

Karaktererne oversættes til ECTS-skalaen ved at se på rangeringen af de studerendes resultat efter den nationale karakterskala. Hvis en studerende fx er blandt de 30 % bedste ved en prøve, men ikke er blandt de 10 % bedste, svarer præstationen til karakteren B efter ECTS-skalaen.

Oversættelse til ECTS-skalaen kræver altså, at den nationale karaktergivning giver tilstrækkeligt detaljerede oplysninger om de studerendes indbyrdes rangering i karakterfordelingen til at kunne identificere, hvilken af de fem percentilgrupper i skalaens karakterfordeling de studerendes resultat ligger i.

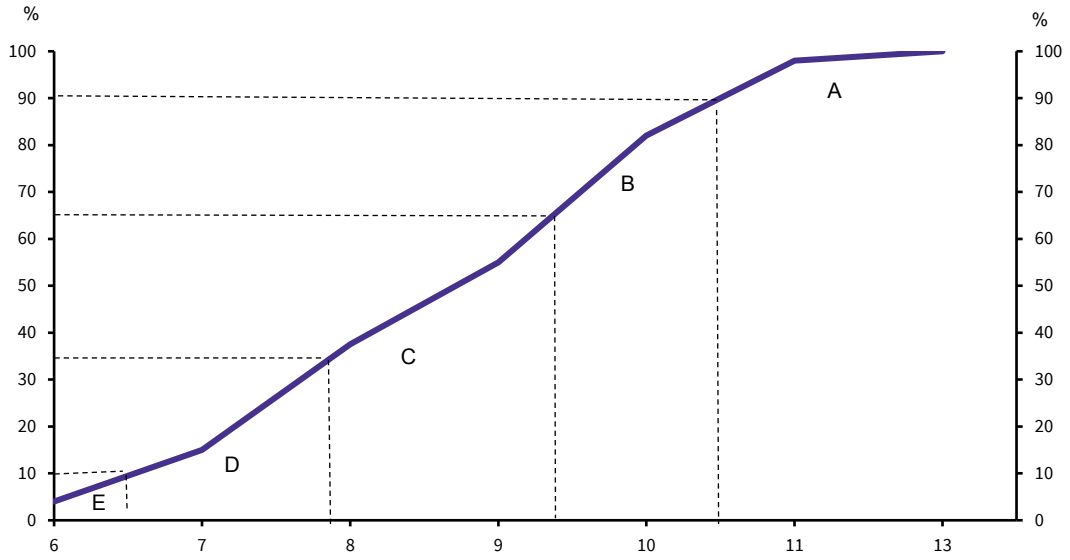
Alternativt vil karakterne ikke kunne oversættes til andet end bestået/ikke-bestået. Det efterlader de studerende med den udfordring at skulle formidle deres karakterer til lande, som benytter en anden karakterskala.

Den danske implementering af ECTS-skalaen

Den asymmetriske opbygning af de bestående karakterer i den tidligere danske karakterskala, 13-skalaen, med to lave karakterer og fire karakterer over middel gjorde det svært at etablere en en-til-en-sammenhæng mellem denne skala og ECTS-skalaens fem bestående karakterer. Det udelukkede dog ikke, at 13-skalaen kunne bruges til oversættelse til ECTS-skalaen ved hjælp af data for fordelingen af de bestående karakterer, jf. figur B.1.

FIGUR B.2

Omregning af bestående karakterer efter 13-skalaen for de afsluttende karakterer på sygeplejeuddannelsen



Kilde: Uddannelses- og Forskningsministeriet på basis af data fra Karakterkommissionen, figur 13, bilag 3.
Note: Omregningen er bestemt af den akkumulerede karakterfordeling efter 13-skalaen.

Karakterfordelingerne kan variere fra uddannelse til uddannelse, hvilket kan betyde, at de enkelte karakterer skal oversættes anderledes inden for andre områder. Karakterkommissionen påpegede, at det kunne give forvirring i oversættelsen, da A kunne svare til 11 eller 13, B, C og D samlet set skulle dække 7, 8, 9 og 10 uden en klar sammenhæng, og E skulle dække 6.

Kommissionen pegede derfor på, at 13-skalaen var:

- Svær at kommunikere til udlandet, blandt andet på grund af undtagelseskarakteren 13
- Ikke kunne oversættes til ECTS-skalaen
- Ikke kunne oversættes til den mest anvendte amerikanske skala.

Kommissionen vurderede derfor, at 13-skalaen ikke var internationalt anvendelig.

7-trins-skalaen er derfor opbygget med udgangspunkt i ECTS-skalaen, så karaktererne kan oversættes en-til-en mellem de to skalaer. Karaktertrinene er valgt ud fra procentfordelingen af de bestående karakterer i ECTS-skalaen, jf. figur B.3.

I udgangspunktet er den laveste bestående karakter sat til at være 0-punktet, den højeste til 100 %. De midterste karakterer er sat til den percentil midt mellem deres procentgrænser. Dernæst er værdierne divideret med 10 og rundet op for at få tal, der i højere grad svarer til de karakterer, som er kendt og brugt i Danmark.

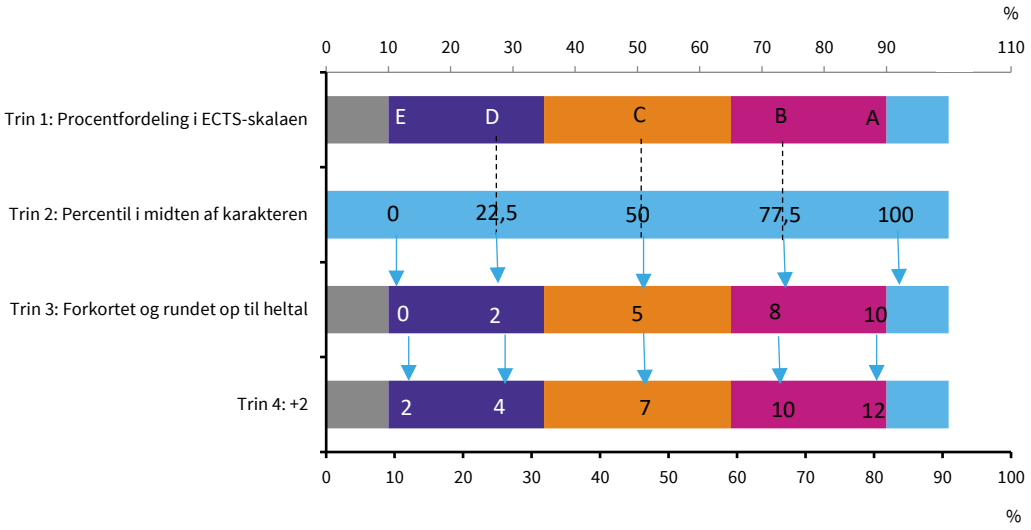
Kommissionen vurderede, at de bestående og ikke-bestående karakterer skulle ligge symmetrisk omkring den laveste karakter (altså -2 og -5). Baggrunden for kommissionens beslutning var blandt

andet, at opbygningen af skalaen skulle understøtte, at andelen af elever, der samlet set (ud fra deres karaktergennemsnit) ikke bestod deres uddannelse, skulle være den samme efter 7-trins-skalaen som efter 13-skalaen. Ved at omregne karakterer mellem de to skalaer og udregne gennemsnit påviste kommissionen, at denne afstand mellem de ikke-bestående karakterer var passende.

For at undgå at have en bestået karakter på 00, valgte kommissionen at lægge 2 til alle karakterer, så der kun var én negativ karakter (-3)[9].

FIGUR B.3

Sammenhæng mellem ECTS-skalaen og 7-trins-skalaen



Kilde: Uddannelses- og Forskningsministeriet på grundlag af Karakterkommissionen, 2004.

7-trins-skalaen holder derved fast i den danske tradition med at bruge talkarakterer, som ikke i sig selv udtrykker opnåede point, som det fx gælder ved den mest udbredte franske skala, eller er knyttet til et bagvedliggende pointsystem, som det fx gælder med karaktersystemet på de videregående uddannelser i Storbritannien. Derved adskiller 7-trins-skalaen sig fra karakterskalaen i alle andre lande, hvilket er blevet belyst som en del af evalueringen.

En forudsætning for, at karakterer givet efter 7-trins-skalaen kan oversættes en-til-en til ECTS-skalaen er dog, at de i det mindste over tid og for store populationer følger den relative karakterfordeling efter ECTS-skalaen (10 % af de bestående prøver får 2, 25 % får 4, 30 % får 7, 25 % får 10 og 10 % får 12). Som beskrevet i kapitel 3 af evalueringen er det ikke tilfældet på de videregående uddannelser, hvor der med undtagelse af de maritime uddannelser og erhvervsakademierne er blevet givet højere karakterer end forventet samlet set over de seneste ti år.

En direkte en-til-en-oversættelse af de danske karakterer efter ECTS-skalaen kan derfor føre til for høje ECTS-karakterer på hovedparten af de videregående uddannelser.

[9] Opmærksomhedspunkterne i forbindelse med den negative karakter -3 er beskrevet nærmere i evalueringens kapitel 5. Kommissionens betragtninger om skalaens opbygning kan læses i Karakterkommissionen, 2004, kapitel 8 og bilag 8.

Udfasning af ECTS-skalaen i ECTS-systemets anbefalinger

ECTS-skalaen har været udsat for kritik siden dens implementering i ECTS-systemets anbefalinger. En tredje faktor, som gav anledning til debat var den oprindelige version af ECTS-skalaens beskrivelse. Beskrivelserne fremgår af tabellen nedenfor.

TABEL B.1

Oprindelig beskrivelse af ECTS-karakterer i vejledningen til ECTS-systemet

ECTS-karakter	% af beståede resultater	Beskrivelse i ECTS-skalaen
A	10	Excellent – outstanding performance with minor errors
B	25	Very Good – above the average standard but with some errors
C	30	Good – generally sound work but with a number of notable errors
D	25	Satisfactory – fair but with significant shortcomings
E	10	Sufficient –performance meets the minimum criteria
Fx	-	Fail – some more work before credit can be awarded
F	-	Fail – considerable further work is required

Kilde: Karran, 2004.

Kritikere pegede for eksempel på, at de 10 % bedste præstationer ikke nødvendigvis var excellente[10]. Beskrivelserne blev derfor fjernet fra vejledningen i brug af skalaen i 2004.

Samtidig var skalaen en udfordring for lande og institutioner, som benytter karakterskalaer med færre beståede karaktertrin end fem (herunder mange svenske institutioner, som benyttede en skala med tre beståede trin). Her kunne ECTS-skalaen, som beskrevet tidligere, ikke bruges. Sverige overvejede derfor at implementere ECTS-skalaen som bedømmelsesværktøj på alle videregående uddannelser. Det blev dog ikke gennemført, men førte til debat og kritik af ECTS-skalaen[11].

Desuden blev ECTS-skalaen kritiseret for, at oversættelse af karakterer i to led (først oversættelse til ECTS-skalaen efterfulgt af oversættelse videre til en ny skala) var upræcist og kunne føre til problemer. I mange europæiske lande findes karakteroplysningerne ikke med større detaljeringsgrad end de trin, der er i den anvendte karakterskala. I Danmark findes der fx ikke oplysninger om, hvordan de studerende, der har fået 7, har klaret sig i forhold til hinanden. Det registreres kun, at de har fået 7. Der er derfor en usikkerhed i forbindelse med oversættelsen til ECTS-skalaen og ekstra usikkerhed, når der oversættes videre til en anden skala.

Anbefalingen i ECTS Users' Guide var ikke, at alle de europæiske lande skulle udskifte deres nationale karakterskala og implementere ECTS-skalaen direkte som en del af deres uddannelsessystem. I respekt for, at karaktersystemer typisk er dybt forankrede i landenes uddannelseskultur og historie, var tanken derimod at etablere et fælles sprog til at tale om karakterer.

[10] Interview med Anthony Vickers samt Karran, 2004.

[11] Se Dahl et al., 2009.

Ud over Danmark har Norge, Estland og flere institutioner og uddannelser med autonomi til selv at vælge deres karakterskalaer (fx Stockholms universitet og University of Eastern Finland) implementeret skalaer, som skulle understøtte direkte en-til-en-oversættelse til ECTS-skalaen.

Langt de fleste europæiske karakterskalaer bruger absolut (kriteriebaseret) karaktergivning, hvor karaktererne ikke nødvendigvis følger en fast fordeling fra år til år. En en-til-en oversættelse er derfor ikke nødvendigvis retvisende.

Dette er en af de udfordringer, som siden er blevet fremhævet ved ECTS-skalaen, idet det har betydet, at skalaen faktisk ikke blev anvendt ensartet på tværs af lande, institutioner og uddannelser.

I 2009 udgik ECTS-skalaen derfor af anbefalingerne i ECTS-systemet og blev erstattet af en anbefaling om brug af statistik for den faktiske karaktergivning i form af ECTS grading tables til at skabe gennemsigtighed med hensyn til karaktergivningen.

Sigtet med brug af grading tables er det samme som med ECTS-skalaen. Men i stedet for at forsøge at passe karakterer fra forskellige skaler ind i en på forhånd fastsat normalfordeling, oversættes karaktererne direkte fra en skala til en anden ved at bruge de faktiske karakterfordelinger for de skalaer og faglige områder, der oversættes imellem.

Det betyder, at der ikke er krav om antallet af trin i de to skalaer eller den specifikke fordeling af karaktererne, som der oversættes imellem. Det eneste krav er, at der findes oplysninger om karakterfordelingen for de beståede karakterer efter de to karakterskalaer. Anbefalingen i ECTS-systemet er derfor ændret til, at landene skal offentliggøre de faktiske karakterfordelinger, helst per fagligt område (fx ISCED-kode) per institution.

EU's officielle vejledning i udarbejdelse af diploma supplements er derfor også ændret. Diploma supplements er et engelsksproget bilag til eksamensbeviser med oplysninger om uddannelsen, uddannelsessystemet og de opnåede resultater, som alle danske videregående uddannelser er forpligtet til at udarbejde. Ændringen betyder, at der i stedet for oversættelse af karakterer til ECTS-skalaen skal være en beskrivelse af eksamenssystemet og evt. karakterstatistik i diploma supplementet.

Anthony Vickers, som har været med til at udvikle oversættelsesmetoden, og som siden har arbejdet med udviklingen af Egracons (Erasmus+-samarbejdets system til it-understøttelse af karakteroversættelser) og raffinering af University of Essexs interne praksis for oversættelse af karakterer mellem lande, påpeger, at metoden kan være mere grov, når der er tale om oversættelser mellem skalaer med mange trin (fx det franske eller italienske system) og skalaer med få trin (fx den mest brugte svenske skala med kun tre beståede trin). Men han påpeger også, at det betyder, at der bliver transparens med hensyn til, hvordan de nationale karakterskalaer i realiteten bruges, og hvordan de studerende har klaret sig i forhold til andre.

Den nye oversættelsesmetode løser dog ikke udfordringen med, at der kan være stor forskel på de faglige krav i de forskellige lande. Vickers påpeger, at da omregningen kun sker på grundlag af karakterfordelingen for de beståede karakterer, kan andelen af karakterer, som ikke er bestået, være et nyttigt supplement til omregningen. Men udfordringen med forskelle i det faglige niveau gælder

både med hensyn til grænsen for, hvornår man består et fag eller opnår en given karakter, og med hensyn til de enkelte fags indhold, og hvad der lægges vægt på i udprøvningen[12].

Han påpeger derfor, at metoden med at omregne ud fra de bestående karakterer er den mest hensigtsmæssige, men at der kan være behov for at foretage justeringer af beregningerne for lande, hvor det faglige niveau viser sig at være lavere, end karaktererne signalerer.

Dansk implementering af de gældende anbefalinger om transparens i karaktergivningen

Skiftet i ECTS-systemet til en anbefaling om at offentliggøre og bruge grading tables er ikke blevet implementeret i dansk lovgivning.

De danske karakterbekendtgørelser refererer stadig til ECTS-skalaen, og den danske vejledning til udarbejdelse af dokumentation for udlandsophold opfordrer ikke til, at institutionerne skal vedlægge statistik for karakterfordelinger[13].

Institutionerne er i dag forpligtede til at offentliggøre statistik for karaktergivningen på deres uddannelser; det er typisk karaktergennemsnit, der offentliggøres og evt. kun for de afsluttende opgaver. Der er altså ikke krav om offentliggørelse af statistik i form af grade distribution tables[14]. Deltagelse i ECTS-systemet er baseret på frivillighed, og det er ikke et krav, at Danmark skal følge alle systemets anbefalinger. Ændringerne i anbefalingerne har betydet, at andre lande, fx Sverige og Finland, ikke længere refererer til ECTS-skalaen i deres forklaringer af deres karakterskalaer. Referencer til ECTS-skalaen kan dog stadig findes på flere europæiske universiteters hjemmesider, og fokusgrupperne med internationale medarbejdere fra de videregående uddannelser peger på, at nogle stadig finder det nyttigt at referere til ECTS-skalaen, når de formidler de danske karakterer.

Det kan dog være relevant at vurdere, om justeringen af anbefalingerne også skal føre til ændringer i den officielle danske formidling af karakterer i udlandet.

[12] Et eksempel på en sådan forskel fremhæves i Søndergaard, 2008: at gymnasiale uddannelser i Storbritannien krævede, at mindst 70 % af en besvarelse var rigtig for at opnå topkarakter, mens den tilsvarende procentdel i Danmark var 91 %.

[13] https://ufm.dk/uddannelse/anerkendelse-og-dokumentation/dokumentation/europass/diploma-supplement/ds-vejledning/udfyldelse-af-diploma-supplement.

[14] Lov nr. 810 af 2.6.2018 om gennemsigtighed og åbenhed i uddannelserne m.v.

Evalueringen af 7-trins-skalaen

© 2018 Danmarks Evalueringsinstitut

Citat med kildeangivelse er tilladt

Publikationen er kun udgivet i elektronisk form på: www.eva.dk

ISBN (www) 978-87-7182-243-4

Danmarks Evalueringsinstitut (EVA) gør uddannelse og dagtilbud bedre. Vi leverer viden, der bruges på alle niveauer – fra institutioner og skoler til kommuner og ministerier.

DANMARKS EVALUERINGSINSTITUT

T 3555 0101
E eva@eva.dk
H www.eva.dk