

Bedømmernes perspektiver på 7-trins-skalaen

Bilagsrapport til Evalueringen af 7-trins-skalaen

INDHOLD

Bedømmernes perspektiver på 7-trins-skalaen

1	Indledning	5
1.1	Baggrund	5
1.1	Formål med evalueringen	8
1.2	Metodisk grundlag	9

2	Forståelse og anvendelse af skalaen	11
2.1	Baggrund og undersøgelsesspørgsmål	11
2.2	Survey blandt bedømmere på grundskoleområdet	12
2.3	Survey blandt bedømmere på de gymnasiale uddannelser	17
2.4	Survey blandt bedømmere på erhvervsuddannelserne	24
2.5	Survey blandt bedømmere på de videregående uddannelser	29
2.6	Fokusgruppeinterview med bedømmere	36
2.7	Enkeltinterviews med bedømmere	45

3	Bedømmelsesredskaber til rådighed	53
3.1	Baggrund og undersøgelsesspørgsmål	53
3.2	Survey blandt bedømmere på grundskoleområdet	54
3.3	Survey blandt bedømmere på de gymnasiale uddannelser	67
3.4	Survey blandt bedømmere på erhvervsuddannelserne	86
3.5	Survey blandt bedømmere på de videregående uddannelser	98
3.6	Fokusgruppeinterviews med bedømmere	111
3.7	Enkeltinterviews med bedømmere	117

4	Understøttelse af bedømmerne	122
4.1	Baggrund og undersøgelsesspørgsmål	122

4.2	Survey blandt bedømmere på grundskoleområdet	122
4.3	Survey blandt bedømmere på de gymnasiale uddannelser	124
4.4	Survey blandt bedømmere på erhvervsuddannelserne	127
4.5	Survey blandt bedømmere på de videregående uddannelser	129
4.6	Fokusgruppeinterviews med bedømmere	131
5	Kvalitetssikring	136
5.1	Baggrund og undersøgelsesspørgsmål	136
5.2	Fokusgruppeinterviews med bedømmere	137
6	Karakterskalaens udformning	143
6.1	Baggrund og undersøgelsesspørgsmål	143
6.2	Survey blandt bedømmere på grundskoleområdet	144
6.3	Survey blandt bedømmere på de gymnasiale uddannelser	148
6.4	Survey blandt bedømmere på erhvervsuddannelserne	154
6.5	Survey blandt bedømmere på de videregående uddannelser	158
6.6	Fokusgruppeinterviews med bedømmere	162
6.7	Enkeltinterviews med bedømmere	177
7	Belønning af særligt dygtige studerende	186
7.1	Baggrund og undersøgelsesspørgsmål	186
7.2	Survey blandt bedømmere på de videregående uddannelser	187
7.3	Fokusgruppeinterviews med bedømmere	188
Appendiks A – Metode		190

1 Indledning

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://pub.uvm.dk/2004/karakterer/>).

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 middelmådig karakter), 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 absolute 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 beståede 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-beståede karakterer, 00 og -3. Talværdierne er fastsat, så afstanden mellem de ikke-beståede karaktererne er den samme som afstanden mellem de laveste beståede 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 beståede 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.

1.1 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 karaktergivningens 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? (UFM)

Nedenfor ses evalueringsspørgsmålene og de hovedspørgsmål, der besvares ud fra hvert evalueringsspørgsmål.

TABEL 1.1

Oversigt over evalueringsspørgsmål og hovedspørgsmål

Evalueringsspørgsmål	Hovedspørgsmål
Evalueringsspørgsmål 2: Anvendes skalaen, som den var tænkt af Karak- terkommissionen, og er kommissio- nens krav til en moderne karakter- skala fortsat aktuelle?	2.a) Forståelsen af skalaen: - Hvordan forstår bedømmerne de absolutte og relative elementer i karak- terskalaen? Og hvilken betydning har det for deres karaktergivningsprak- sis? - Hvilken betydning har trinbeskrivelsen for anvendelsen af skalaen? 2.b) Anvendelsen af skalaen: - Hvordan anvendes bedømmelsesværktøjer i karaktergivningen? - Anvendes skalaen som en mangelskala i karaktergivningen? 2.c) Bedømmelsesværktøjer til rådighed (før karaktergivning): - Hvilke bedømmelsesværktøjer er til rådighed ved mundtlige og skriftlige prøver/eksaminer? - Hvordan klædes bedømmere på til at anvende bedømmelsesværktø- jerne? - Særligt for videregående uddannelser: Er der læringsmålbeskrivelser for alle prøver? Er der behov for andre redskaber ved afsluttende opga- ver? - Hvordan sikres en ensartet bedømmelse på tværs af fag/uddannel- ser/udbud? 2.d) Kvalitetssikringen af karaktergivningen (efter karaktergivning): - Hvad gør forskellige aktører for at opnå den forventede karakterfordeling på de relevante aggregerede niveauer? 2.e) De fem krav til en moderne karakterskala: - Er de fem krav til 7-trins-skalaen stadig relevante?
Evalueringsspørgsmål 3: Hvad bety- der skalaens karakterspring og ka- rakteren -3 for karaktergivningen?	Hvad betyder karakterspringene og karakteren -3 for karaktergivningen?
Evalueringsspørgsmål 4: Er der til- strækkelige muligheder for at identi- ficere særligt dygtige elever og stu- derende og belønne den ekstraordi- nære præstation?	Er der tilstrækkelige muligheder for at identificere særligt dygtige stude- rende og belønne den ekstraordinære præstation?

1.2 Metodisk grundlag

Denne bilagsrapport bygger på følgende datakilder:

- Spørgeskemaundersøgelse blandt bedømmere på grundskoleområdet, ungdomsuddannelserne og de videregående uddannelser
- Tre fokusgruppeinterviews med samlet set 15 bedømmere fra grundskoleområdet
- Fire fokusgrupper med samlet set 22 bedømmere på de gymnasiale uddannelser
- Ti fokusgrupper med samlet set 66 bedømmere på de videregående uddannelser
- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på grundskole-
området
- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på stx

- Fire dobbeltinterviews med bedømmere på de videregående uddannelser
- Fem enkeltinterviews med bedømmere på de videregående uddannelser.

For mere uddybende information om datakilderne, se appendiks A.

Datagrundlaget i afrapporteringen

For at gøre det muligt at vurdere datagrundlaget, herunder udsagnskraften, er alle kapitler bygget op sådan, at vi altid redegør for hver datakilde separat, således at vi først redegør for resultater fra spørgeskemaundersøgelsen, dernæst redegør for fokusgruppeinterviews med bedømmere og til sidst redegør for enkeltinterviews med bedømmere.

Bemærk, at vi ved gennemgangen af spørgeskemadata altid opdeler og kun præsenterer forskelle på tværs af fag (grundskoler og ungdomsuddannelser), typer, fx stx versus hf (ungdomsuddannelser), sektorer (videregående uddannelser) og hovedområder (videregående uddannelser), i det omfang der er nævneværdige forskelle.

2 Forståelse og anvendelse af skalaen

2.1 Baggrund og undersøgelsesspørgsmål

2.1.1 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.”¹

2.1.2 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

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

Forskningsministeriet via uddannelsesbekendtgørelserne, mens de institutioner, som udbyder de 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.

2.1.3 Kvalitetssikring forventes for at sikre den relative fordeling

Karakterkommissionen beskrev også, at der skulle ske en kvalitetssikring af karaktergivningingen ved at koordinere målene for faget, prøvens tilrettelæggelse og den population, der skal bestå prøven. Det skulle understøtte, at de beståede 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 undersøgelsesspørgsmål:

- Hvordan forstår bedømmerne de absolutte og relative elementer i karakterskalaen? Og hvilken betydning har det for deres karaktergivningspraksis?
- Hvilken betydning har trinbeskrivelsen for anvendelsen af skalaen?
- Anvendes skalaen som en mangelskala i karaktergivningingen?

2.2 Survey blandt bedømmere på grundskoleområdet

2.2.1 Sammenfatning

Surveyen blandt bedømmere på grundskoleområdet viser overordnet set, at:

- Bedømmerne føler, at de har tilstrækkeligt kendskab til skalaen og ved, hvordan den skal bruges
- Bedømmerne anvender beskrivelsen af karaktertrinene i bedømmelsessituationer og oplever beskrivelserne som klare
- Bedømmerne har delte meninger om, hvorvidt trinbeskrivelserne er anvendelige i en læringssituation, og hvorvidt der er for meget fokus på mangler
- De fleste bedømmere ikke tager udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen, ligegyldigt om det er en skriftlig eller en mundtlig prøve.

2.2.2 Gennemgang af resultater

Tabel 2.1 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om forståelse af og kendskab til 7-trins-skalaen. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 94 % af bedømmerne på grundskoleområdet er enige i udsagnet "Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen".
- 94 % er enige i udsagnet "Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer".

- 80 % er enige i udsagnet ”Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret ”Meget enig” eller ”Enig” ved hvert udsagn, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.1

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen	59 % (395)	35 % (236)	3 % (22)	1 % (7)	1 % (4)	0 % (1)	100 % (665)
Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer	56 % (374)	38 % (255)	3 % (23)	2 % (11)	0 % (1)	0 % (1)	100 % (665)
Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig	22 % (143)	58 % (384)	13 % (86)	6 % (42)	1 % (5)	1 % (5)	100 % (665)

Kilde: Danmarks Evalueringsinstitut (EVA), survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 2.2 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 80 % af bedømmerne på grundskoleområdet er enige i udsagnet ”Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning”.
- 77 % er enige i udsagnet ”De forskellige karakterer i trinbeskrivelsen er klart beskrevet”.
- 62 % er enige i udsagnet ”Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag”.
- 26 % er enige i udsagnet ”Beskrivelsen af karaktertrinene er for simpel”.
- 14 % er enige i udsagnet ”Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med hvert udsagn, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.2

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
De forskellige karakterer i trinbeskrivelsen er klart beskrevet	20 % (136)	57 % (379)	15 % (98)	6 % (41)	1 % (8)	0 % (3)	100 % (665)
Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning	27 % (180)	53 % (354)	15 % (99)	4 % (25)	1 % (5)	0 % (2)	100 % (665)
Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag	14 % (93)	48 % (320)	24 % (161)	12 % (80)	1 % (9)	0 % (2)	100 % (665)
Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer	2 % (15)	12 % (79)	14 % (95)	38 % (250)	33 % (219)	1 % (7)	100 % (665)
Beskrivelsen af karaktertrinene er for simpel	7 % (49)	19 % (126)	35 % (231)	28 % (187)	10 % (64)	1 % (8)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 2.3 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om beskrivelsen af de forskellige trin. I det følgende oplyses andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 55 % er enige i udsagnet "Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation, herunder fx den usædvanligt gode og selvstændige mestring af stoffet".
- 40 % er enige i udsagnet "Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), frem for hvad den studerende har lært (mestring)".
- 39 % er enige i udsagnet "Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler".
- 34 % er enige i udsagnet "Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med hvert udsagn, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.3

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevens læring	4 % (27)	30 % (201)	38 % (255)	18 % (117)	5 % (30)	5 % (35)	100 % (665)
Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), frem for hvad den studerende har lært (mestring)	10 % (69)	30 % (199)	32 % (215)	21 % (137)	5 % (31)	2 % (14)	100 % (665)
Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation, herunder fx den usædvanligt gode og selvstændige mestring af stoffet	23 % (150)	32 % (215)	18 % (123)	20 % (130)	5 % (36)	2 % (11)	100 % (665)
Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler	10 % (69)	29 % (195)	28 % (189)	25 % (168)	5 % (36)	1 % (8)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 2.4 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en skriftlig prøve?". Tabellen viser, at:

- 73 % angiver: "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen".
- 12 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel".
- 10 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse".
- 5 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har angivet "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen", er statistisk signifikant forskellig mellem bedømmere fra engelsk (85 %) og bedømmere fra matematik (65 %).

TABEL 2.4

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	12 % (45)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	10 % (38)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen	5 % (18)
Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen	73 % (273)
Total	100 % (374)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

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.

Tabel 2.5 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en mundtlig prøve?". Tabellen viser, at:

- 83 % angiver: "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen".
- 9 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel".
- 6 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen".
- 2 % angiver: "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har angivet "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel", er statistisk signifikant forskellig mellem bedømmere fra dansk (16 %) og bedømmere fra engelsk (2 %) såvel som mellem bedømmere fra dansk (16 %) og bedømmere fra matematik (6 %).
- Andelen, der har angivet "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen", er statistisk signifikant forskellig mellem bedømmere fra dansk (73 %) og bedømmere fra engelsk (91 %).

TABEL 2.5

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	9 % (37)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	2 % (8)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	6 % (25)
Når jeg tildeler karakterer, tager jeg tager ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen	83 % (336)
Total	100 % (406)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

2.3 Survey blandt bedømmere på de gymnasiale uddannelser

2.3.1 Sammenfatning

Surveyen blandt bedømmere på de gymnasiale uddannelser viser overordnet set, at:

- Bedømmerne føler, at de har tilstrækkeligt kendskab til skalaen og ved, hvordan den skal bruges.
- Bedømmerne anvender beskrivelsen af karaktertrinene i bedømmelsessituationer og oplever beskrivelserne som klare.
- Bedømmerne har delte meninger om, hvorvidt trinbeskrivelserne er anvendelige i en læringssituation, og hvorvidt der er for meget fokus på mangler.
- De fleste bedømmere ikke tager udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen ligegyldigt om det er en skriftlig eller mundtlig prøve.

2.3.2 Gennemgang af resultater

Tabel 2.6 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om forståelse af og kendskab til 7-trins-skalaen. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 94 % af bedømmerne på de gymnasiale uddannelser er enige i udsagnet: "Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen"

- 94 % er enige i udsagnet ”Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer”
- 85 % er enige i udsagnet ”Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig”, er statistisk signifikant forskellig mellem bedømmere i dansk (81 %) og bedømmere i matematik (88 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig”, er statistisk signifikant forskellig mellem bedømmere på stx (87 %) og bedømmere på htx (74 %), såvel som mellem bedømmere på hf (91 %) og bedømmere på htx (74 %).

TABEL 2.6

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen	58 % (686)	36 % (431)	3 % (39)	2 % (25)	0 % (5)	0 % (1)	100 % (1187)
Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer	56 % (669)	38 % (452)	4 % (44)	2 % (20)	0 % (1)	0 % (1)	100 % (1187)
Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig	32 % (374)	53 % (634)	10 % (119)	4 % (48)	0 % (5)	1 % (7)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere i gymnasiet.

Tabel 2.7 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om beskrivelsen af de forskellige trin. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 75 % er enige i udsagnet ”Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning”
- 72 % er enige i udsagnet ”De forskellige karakterer i trinbeskrivelsen er klart beskrevet”
- 53 % er enige i udsagnet ”Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag”

- 32 % er enige i udsagnet "Beskrivelsen af karaktertrinene er for simpel"
- 16 % er enige i udsagnet "Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor opoplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "De forskellige karakterer i trinbeskrivelsen er klart beskrevet", er statistisk signifikant forskellig mellem bedømmere i dansk (68 %) og bedømmere i matematik (75 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Beskrivelsen af karaktererne er en støtte i forbindelse med karaktergivning", er statistisk signifikant mellem bedømmere i dansk (72 %) og bedømmere i engelsk (80 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag", er statistisk signifikant forskellig mellem bedømmere i dansk (47 %) og engelsk (64 %), såvel som mellem bedømmere i dansk (47 %) og bedømmere i matematik (58 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer", er statistisk signifikant forskellig mellem bedømmere i dansk (16 %) og bedømmere i engelsk (11 %), såvel som mellem bedømmere i engelsk (11 %) og bedømmere i matematik (19 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Beskrivelsen af karaktertrinene er for simpel", er statistisk signifikant forskellig mellem bedømmere i dansk (36%) og bedømmere i matematik (22 %), såvel som mellem bedømmere i engelsk (37 %) og bedømmere i matematik (22 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hhx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor opoplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Beskrivelsen af karaktertrinene er for simpel", er statistisk signifikant forskellig mellem bedømmere på stx (32 %) og bedømmere på hhx (45 %), såvel som mellem bedømmere på hf (25 %) og bedømmere på hhx (45 %).

TABEL 2.7

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
De forskellige karakterer i trinbeskrivelsen er klart beskrevet	19 % (225)	53 % (626)	17 % (197)	10 % (113)	2 % (21)	0 % (5)	100 % (1187)
Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning	23 % (269)	52 % (623)	18 % (212)	5 % (65)	1 % (15)	0 % (3)	100 % (1187)
Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag	12 % (147)	41 % (490)	28 % (330)	16 % (185)	3 % (31)	0 % (4)	100 % (1187)
Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer	3 % (34)	13 % (152)	19 % (220)	40 % (477)	25 % (298)	1 % (6)	100 % (1187)
Beskrivelsen af karaktertrinene er for simpel	9 % (112)	23 % (276)	30 % (359)	28 % (337)	6 % (69)	3 % (34)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere i gymnasiet.

Tabel 2.8 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om de forskellige trin på skalaen. I det følgende opoplistes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 57 % er enige i udsagnet "Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet"
- 49 % er enige i udsagnet "Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)"
- 44 % er enige i udsagnet "Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler"
- 33 % er enige i udsagnet "Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor opoplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring", er statistisk signifikant forskellig mellem bedømmere i dansk (29 %) og bedømmere i matematik (37 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)", er statistisk signifikant forskellig

mellem bedømmere i dansk (56 %) og bedømmere i engelsk (45 %), såvel som mellem bedømmere i dansk (56 %) og bedømmere i matematik (43 %).

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet”, er statistisk signifikant forskellig mellem bedømmere i dansk (68 %) og bedømmere i engelsk (59 %), såvel som for bedømmere i dansk (68 %) og bedømmere i matematik (44 %), såvel som for bedømmere i engelsk (59 %) og bedømmere i matematik (44 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler”, er statistisk signifikant forskellig mellem bedømmere i dansk (53 %) og bedømmere i engelsk (43 %), såvel som for bedømmere i dansk (53 %) og bedømmere i matematik (36 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler”, er statistisk signifikant forskellig mellem bedømmere på stx (45 %) og bedømmere på hf (32 %).

TABEL 2.8

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring	5 % (56)	28 % (338)	32 % (379)	21 % (251)	9 % (102)	5 % (61)	100 % (1187)
Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/elev ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)	16 % (186)	33 % (389)	29 % (339)	17 % (206)	4 % (42)	2 % (25)	100 % (1187)
Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet	28 % (335)	29 % (349)	16 % (188)	19 % (228)	5 % (64)	2 % (23)	100 % (1187)
Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler	13 % (160)	31 % (368)	30 % (352)	21 % (248)	4 % (42)	1 % (17)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere i gymnasiet.

Tabel 2.9 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet: "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en skriftlig prøve?":

- 79 % angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen"
- 14 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel"
- 5 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen"
- 2 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor opoplistes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel", er statistisk signifikant forskellig mellem bedømmere i dansk (5 %) og bedømmere i matematik (31 %), såvel som mellem bedømmere i dansk (5 %) og bedømmere i engelsk (10 %), såvel som mellem bedømmere i engelsk (10 %) og bedømmere i matematik (31 %).
- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse", er statistisk signifikant forskellig mellem bedømmere i dansk (0 %) og bedømmere i matematik (9 %), såvel som mellem bedømmere i engelsk (0 %) og bedømmere i matematik (9 %).
- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen", er statistisk signifikant forskellig mellem bedømmere i dansk (7 %) og bedømmere i matematik (1 %), såvel som mellem bedømmere i engelsk (7 %) og bedømmere i matematik (1 %).
- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen", er statistisk signifikant forskellig mellem bedømmere i dansk (88 %) og bedømmere i matematik (59 %), såvel som mellem bedømmere i engelsk (84 %) og bedømmere i matematik (59 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor opoplistes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse", er statistisk signifikant forskellig mellem bedømmere stx (2 %) og bedømmere på htx (6 %).

TABEL 2.9

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	14 % (98)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	2 % (17)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen	5 % (35)
Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen	79 % (571)
Total	100 % (721)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

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

Tabel 2.10 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet: "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en mundtlig prøve?".

- 86 % angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen"
- 7 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel"
- 7 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen"

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel", er statistisk signifikant forskellig mellem bedømmere fra dansk (4 %) og bedømmere fra matematik (11 %) såvel som mellem bedømmere fra engelsk (5 %) og bedømmere fra matematik (11 %).
- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen", er statistisk signifikant forskellig mellem bedømmere fra dansk (89 %) og bedømmere fra matematik (80 %) såvel som mellem bedømmere fra engelsk (88 %) og bedømmere fra matematik (80 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/ trækker fra hen ad vejen", er statistisk signifikant forskellig mellem bedømmere på htx (7 %) og bedømmere på htx (21 %), såvel som mellem bedømmere på hf (6%) og bedømmere på htx (21 %), såvel som mellem bedømmere på stx (8 %) og bedømmere på htx (21 %).

TABEL 2.10

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	7 % (58)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	1 % (6)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	7 % (60)
Når jeg tildeler karakterer, tager jeg tager ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen	86 % (758)
Total	100 % (882)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018

2.4 Survey blandt bedømmere på erhvervsuddannelserne

2.4.1 Sammenfatning

Surveyen blandt bedømmere på erhvervsuddannelserne viser overordnet set, at:

- Bedømmerne føler, at de har tilstrækkeligt kendskab til skalaen, og ved hvordan den skal bruges
- Bedømmerne anvender beskrivelsen af karaktertrinene i bedømmelsessituationer, og oplever beskrivelserne som klare
- Bedømmerne har delte meninger om, hvorvidt trinbeskrivelserne er anvendelige i en læringssituation, og hvorvidt der er for meget fokus på mangler
- De fleste bedømmere tager ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen ligegyldigt om det er en skriftlig eller mundtlig prøve.

2.4.2 Gennemgang af resultater

Tabel 2.11 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om forståelse af og kendskab til 7-trins-skalaen. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 96 % af bedømmerne på erhvervsuddannelserne er enige i udsagnet "Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen"
- 94 % er enige i udsagnet "Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer"
- 85 % er enige i udsagnet "Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen der har svaret "Meget enig" eller "Enig" i forbindelse med hvert udsagn er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.11

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen	59 % (184)	37 % (115)	3 % (8)	2 % (6)	0 % (1)	0 % (0)	100 % (314)
Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer	53 % (166)	41 % (130)	3 % (10)	2 % (7)	0 % (1)	0 % (0)	100 % (314)
Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig	34 % (106)	51 % (161)	11 % (35)	2 % (7)	1 % (3)	1 % (2)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Tabel 2.12 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om trinbeskrivelsen. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 83 % er enige i udsagnet "Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning".
- 74 % er enige i udsagnet "De forskellige karakterer i trinbeskrivelsen er klart beskrevet".
- 60 % er enige i udsagnet "Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag".
- 27 % er enige i udsagnet "Beskrivelsen af karaktertrinene er for simpel"

- 10 % er enige i udsagnet ”Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Beskrivelsen af karaktertrinene er for simpel”, er statistisk signifikant forskellig mellem bedømmere i dansk (36 %) og bedømmere i matematik (18 %).

TABEL 2.12

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
De forskellige karakterer i trinbeskrivelsen er klart beskrevet	23 % (72)	51 % (161)	15 % (47)	9 % (28)	2 % (6)	0 % (0)	100 % (314)
Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning	31 % (97)	52 % (164)	13 % (41)	3 % (10)	1 % (2)	0 % (0)	100 % (314)
Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag	18 % (55)	42 % (131)	27 % (86)	11 % (35)	2 % (7)	0 % (0)	100 % (314)
Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer	2 % (5)	8 % (24)	14 % (44)	38 % (118)	39 % (121)	1 % (2)	100 % (314)
Beskrivelsen af karaktertrinene er for simpel	7 % (23)	20 % (63)	25 % (80)	34 % (107)	12 % (38)	1 % (3)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Tabel 2.13 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om trinbeskrivelsen. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 61 % er ”enige” i udsagnet ”Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet”
- 45 % er enige i udsagnet ”Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)”
- 44 % er enige i udsagnet ”Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler”
- 39 % er enige i udsagnet ”Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)”, er statistisk signifikant forskellig mellem bedømmere i engelsk (60 %) og bedømmere i matematik (40 %).

TABEL 2.13

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring	7 % (23)	32 % (99)	34 % (107)	20 % (62)	4 % (13)	3 % (10)	100 % (314)
Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)	12 % (38)	33 % (104)	31 % (96)	18 % (58)	4 % (11)	2 % (7)	100 % (314)
Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet	32 % (101)	29 % (92)	18 % (56)	14 % (45)	5 % (17)	1 % (3)	100 % (314)
Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler	14 % (45)	30 % (94)	30 % (94)	18 % (58)	5 % (16)	2 % (7)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Tabel 2.14 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet ”Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en skriftlig prøve?”.

- 61 % angiver ”Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen”
- 33 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel”
- 4 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse”
- 2 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen der har svaret ”Meget enig” eller ”Enig” til hvert udsagn er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.14

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	33 % (19)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	4 % (2)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	2 % (1)
Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen	61 % (35)
Total	100 % (57)

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

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne

Tabel 2.15 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet ”Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en mundtlig prøve?”. Tabellen viser, at:

- 71 % angiver ”Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen”
- 24 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel”
- 3 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen”
- 2 % angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Besvarelsene på hvert bedømmelsesværktøj er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 2.15

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	24 % (69)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	2 % (5)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	3 % (10)
Når jeg tildeler karakterer, tager jeg tager ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen	71 % (204)
Total	100 % (288)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

2.5 Survey blandt bedømmere på de videregående uddannelser

2.5.1 Sammenfatning

Surveyen blandt bedømmere på de videregående uddannelser viser overordnet set, at:

- Bedømmerne føler, at de har tilstrækkeligt kendskab til skalaen, og ved hvordan den skal bruges
- Bedømmerne anvender beskrivelsen af karaktertrinene i bedømmelsessituationer, og oplever beskrivelserne som klare
- Bedømmerne har delte meninger om, hvorvidt trinbeskrivelserne er anvendelige i en læringssituation og hvorvidt der er for meget fokus på mangler
- De fleste bedømmere tager ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen ligegyldigt om det er en skriftlig, mundtlig eller afsluttende prøve.

2.5.2 Gennemgang af resultater

Tabel 2.16 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om forståelse af og kendskab til skalaen. Nedenfor oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 94 % af bedømmerne på de videregående uddannelser er enige i udsagnet ”Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen”
- 93 % er enige i udsagnet ”Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer”
- 81 % er enige i udsagnet ”Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig”.

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen”, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (94 %) og bedømmere på professionsbacheloruddannelser (96 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.16

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg oplever at have tilstrækkeligt kendskab til beskrivelserne af de forskellige karaktertrin i 7-trins-skalaen	57 % (1689)	37 % (1100)	3 % (97)	2 % (49)	0 % (10)	0 % (1)	100 % (2946)
Jeg oplever at have et tilstrækkeligt kendskab til, hvordan karakterskalaen skal anvendes i bedømmelsessituationer	52 % (1520)	41 % (1200)	5 % (148)	2 % (66)	0 % (10)	0 % (3)	100 % (2946)
Jeg oplever, at mine medbedømmere forstår karakterskalaen på samme måde som mig	30 % (893)	51 % (1503)	13 % (370)	5 % (141)	1 % (19)	1 % (19)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 2.17 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om beskrivelsen af de forskellige karaktertrin. Nedenfor oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 79 % er enige i udsagnet ”Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning”
- 73 % er enige i udsagnet ”De forskellige karakterer i trinbeskrivelsen er klart beskrevet”
- 49 % er enige i udsagnet ”Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag”

- 32 % er enige i udsagnet "Beskrivelsen af karaktertrinene er for simpel"
- 11 % er enige i udsagnet "Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag" er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelser (46 %) og bedømmere på erhvervsakademiuddannelser (54 %).
- Andelen der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (12 %) og for bedømmere på professionsbacheloruddannelser (5 %), såvel som mellem bedømmere på professionsbacheloruddannelser (5 %) og erhvervsakademiuddannelser (9 %).
- Andelen der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Beskrivelsen af karaktertrinene er for simpel", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (30 %) og for bedømmere på professionsbacheloruddannelser (35 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.17

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
De forskellige karakterer i trinbeskrivelsen er klart beskrevet	21 % (606)	52 % (1541)	15 % (454)	9 % (279)	2 % (65)	0 % (2)	100 % (2946)
Beskrivelsen af karaktertrinene er en støtte i forbindelse med karaktergivning	26 % (781)	53 % (1571)	13 % (391)	5 % (157)	1 % (41)	0 % (5)	100 % (2946)
Det er nemt at oversætte beskrivelsen af karaktertrinene til kriterierne for bedømmelse i mit/mine fag	12 % (349)	37 % (1092)	28 % (832)	18 % (543)	4 % (117)	0 % (14)	100 % (2946)
Jeg bruger ikke beskrivelsen af karaktertrinene i bedømmelsessituationer	3 % (81)	8 % (230)	12 % (365)	38 % (1120)	38 % (1131)	1 % (20)	100 % (2946)
Beskrivelsen af karaktertrinene er for simpel	10 % (289)	22 % (637)	30 % (886)	29 % (851)	8 % (241)	1 % (42)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 2.18 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om beskrivelsen af de forskellige trin. Nedenfor oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 64 % er enige i udsagnet "Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet"
- 51 % er enige i udsagnet "Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler"
- 50 % er enige i udsagnet "Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)"
- 35 % er enige i udsagnet "Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevernes læring", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (36 %) og bedømmere på professionsbacheloruddannelser (31 %), såvel som mellem bedømmere på professionsbacheloruddannelser (31 %) og bedømmere på erhvervsakademiuddannelser (39 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (48 %) og bedømmere på professionsbacheloruddannelser (58 %), såvel som mellem bedømmere på professionsbacheloruddannelser (58 %) og bedømmere på erhvervsakademiuddannelser (46 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (49 %) og bedømmere på professionsbacheloruddannelser (58 %) såvel som mellem bedømmere på professionsbacheloruddannelser (58 %) og bedømmere på erhvervsakademiuddannelser (48 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.18

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Trinbeskrivelsernes fokus på bedømmelser ud fra mangler understøtter en retvisende vurdering af de studerendes/elevens læring	6 % (166)	29 % (868)	29 % (858)	24 % (694)	10 % (282)	3 % (78)	100 % (2946)
Ud fra et læringsperspektiv har trinbeskrivelserne for stort fokus på, hvad den studerende/eleven ikke har lært (mangler), fremfor hvad den studerende har lært (mestring)	16 % (471)	34 % (994)	25 % (741)	20 % (595)	3 % (99)	2 % (46)	100 % (2946)
Trinbeskrivelsernes fokus på mangler gør det for svært at honorere særlige aspekter af den faglige præstation herunder fx den usædvanligt gode og selvstændige mestring af stoffet	37 % (1102)	27 % (786)	13 % (379)	17 % (508)	5 % (144)	1 % (27)	100 % (2946)
Beskrivelsen af karaktertrinene lægger i for høj grad op til, at bedømmelsen foretages på baggrund af mangler	19 % (552)	32 % (946)	23 % (685)	21 % (618)	4 % (116)	1 % (29)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

2.5.3 Resultater vedr. tildeling af karakter ved skriftlig, mundtlig og afsluttende prøve

Tabel 2.19 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en skriftlig prøve?". Tabellen viser, at:

- 63 % angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen"
- 28 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel"
- 7 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen"

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel", er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelserne (31 %) og bedømmere på universitetsuddannelserne (26 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.19

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	28 % (598)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	2 % (43)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	7 % (153)
Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen	63 % (1374)
Total	100 % (2169)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 2.20 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet: "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en mundtlig prøve?". Tabellen viser, at:

- 67 % angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen"
- 26 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel"
- 7 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen"
- 0 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel", er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelserne (29 %) og bedømmere på universitetsuddannelserne (25 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.20

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	26 % (659)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	0 % (7)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	7 % (170)
Når jeg tildeler karakterer, tager jeg tager ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen	67 % (1676)
Total	100 % (2513)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve i løbet af 2016, 2017 eller 2018.

Tabel 2.21 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "Hvilket af følgende udsagn passer bedst på den måde, du tildeler karakterer på ved en afsluttende prøve?". Tabellen viser, at:

- 68 % angiver "Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen, men laver en overordnet vurdering af præstationen"
- 26 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel"
- 6 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen"
- 0 % angiver "Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der angiver ”Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel”, er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelserne (29 %) og bedømmere på universitetsuddannelserne (24 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 2.21

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

	Procent (antal)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i topkarakteren og trækker fra for hver fejl eller mangel	26 % (654)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i bundkarakteren og lægger til for hvert rigtigt svar eller udmærkelse	0 % (9)
Når jeg tildeler karakterer, tager jeg som regel udgangspunkt i en midterkarakter og lægger til/trækker fra hen ad vejen	6 % (164)
Når jeg tildeler karakterer, tager jeg ikke udgangspunkt i et bestemt punkt på skalaen men laver en overordnet vurdering af præstationen	68 % (1729)
Total	100 % (2556)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en afsluttende prøve i løbet af 2016, 2017 eller 2018.

2.6 Fokusgruppeinterview med bedømmere

I det følgende beskrives bedømmernes forståelse af og anvendelse af skalaen, som det er kommet til udtryk i fokusgruppeinterviewene.

2.6.1 Bedømmere på grundskoleområdet

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

I fokusgruppeinterviewene identificeres en række forskellige tilgange, som bedømmerne forklarer, at de anvender, når de lægger sig fast på en karakter. En tilgang er fx kendetegnet ved, at bedøm-

merer starter helt åbent ud og ikke tager udgangspunkt i et bestemt punkt på skalaen i sin bedømmelse, mens andre bedømmere beskriver, at de tager udgangspunkt i et bestemt punkt på skalaen.

Nogle bedømmere beskriver, at de starter helt åbent ud, når de skal bedømme eleverne, og tager udgangspunkt i elevernes præstation. En bedømmer fortæller:



Jeg starter ikke på 12 og trækker fra. Jeg ved ikke, om det er erfaring, men jeg vil hellere tage udgangspunkt i det, jeg hører fra eleven. Så altså, når man har hørt eleven, så finder man ud af, hvor vedkommende ligger. Jeg tager mest udgangspunkt i det, jeg synes, eleven er god til, og ikke i en bestemt karakter. Det er en helhedsvurdering.

Nogle beskriver det, at de starter åbent ud, på den måde, at de eksempelvis sporer sig ind på, hvilket interval på skalaen eleven befinder sig inden for, fx ved brug af de taksonomiske niveauer, og så derfra finder frem til en konkret karakter. En bedømmer fortæller, at hun starter med at spore sig ind på det taksonomiske niveau: "Når jeg så har indplaceret eleven, så kan jeg begynde at forholde mig til karakterskalaen. Det er langt mere ud fra elevens præstation [end ud fra et punkt på skalaen]". Nogle af de bedømmere, der gør brug af denne tilgang, beskriver, at det blandt andet handler om, at karakterer ikke er det første, de har fokus på, når de eksempelvis sidder i en mundtlig eksamenssituation, men derimod den konkrete elevs præstation.

En anden del af bedømmerne beskriver deres bedømmelse på den måde, at de tager udgangspunkt i et bestemt sted på skalaen, og så justerer de karakteren ud fra elevens præstation. En bedømmer fortæller:



Jeg tager udgangspunkt i en middelkarakter. Jeg har en eller anden formodning om, at eleven skal have det og det med, og i så fald er eleven middel. Hvis eleven klarer det bedre, ryger karakteren op, og omvendt.

Bedømmerne beskriver, at det i den mundtlige prøve- og bedømmelsessituation har betydning, om man er censor eller eksaminator. Blandt andet ift., om man starter ud helt åbent eller med en forventning om at ligge på en bestemt del af skalaen. Som eksaminator kender man eleverne, der kommer ind, og man kan derfor have en idé om, hvor de ligger, hvilket man ikke har som censor. En bedømmer fortæller:



Hvilken tilgang jeg bruger til at bedømme, afhænger meget af, om det er min egen klasse, eller om jeg er censor. Med min egen klasse ved jeg nogenlunde, hvor de forskellige elever ligger, og så giver det mening at tage udgangspunkt i middelkarakteren og så lægge til eller fra. Men som censor har man ingen idé om, hvor eleverne ligger til daglig eller ift. hinanden, så der giver det ikke mening at bruge middelkarakteren.

Men det har også betydning, om man er censor eller eksaminator, fordi eksaminator i prøvesituationen ikke alene har fokus på at bedømme eleven, men også på at rammesætte og gennemføre prøven på bedst mulig vis, mens censor i højere grad kan have sit hovedfokus på bedømmelsen undervejs i prøven og derfor blandt andet kan få noteret langt mere ned. En bedømmer forklarer:



Det er selvfølgelig nemmere at være censor end at være lærer [eksaminator], for som censor kan man læne sig tilbage. Det er mere presset, når man er lærer, fordi man også skal styre samtalen. Så får man ikke skrevet så meget ned. Og eleverne har brug for, at man kigger på dem og er til stede, så man kan ikke tage noter undervejs.

Nogle bedømmere oplever ikke ensidigt fokus på fejl og mangler som meningsfuldt

Bedømmerne beskriver, at de generelt ikke anvender skalaen som en mangelskala, men som ovenfor beskrevet i højere grad laver 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.

Nogle bedømmere oplever at være blevet præsenteret for 7-trins-skalaen som en mangelskala, hvor man skulle tage udgangspunkt i 12 og så trække fra ved fejl og mangler. Der er en gruppe af bedømmere, som er kritiske over for tanken om at skulle bedømme ud fra fejl og mangler. Dels fordi de ikke finder det meningsfuldt eller hensigtsmæssigt at forsøge at opgøre fejl og mangler på en sådan vis, som det ville kræve at starte på 12 og så tælle ned derfra. Dels fordi de oplever, at det står i modsætning til, hvad det andre steder beskrives, at de skal bedømme eleverne ud fra. En bedømmer siger:



Der blev lagt op til, at man skulle tage udgangspunkt i 12 og så trække fra, dengang skalaen kom. Men vores dokumenter er i splid med sig selv. Det er meget i modstrid med det, de skriver i prøvevejledningen. Der står udtrykkeligt, at man skal fokusere på det gode i engelsk. Det går jo ikke ift. den forfærdelige tankegang med at trække ned fra 12.

Bedømmelsen foregår forskelligt afhængigt af fag og prøveform

I denne evaluering er fokus på grundskoleområdet på bedømmelse i fagene dansk, engelsk og matematik. Af fokusgrupperne med bedømmere i disse fag fremgår det tydeligt, at fagenes forskelligartethed betyder, at bedømmelse i fagene også foregår forskelligt, særligt når det kommer til den skriftlige bedømmelse. Hvor man i matematik og dele af engelsk har en egentlig rettenøgle, så gælder det samme ikke for dansk skriftlig fremstilling. En bedømmer siger: "Det er det nemmeste i verden at give karakterer i skriftlig matematik. Vi har en rettevejledning, som vi skal følge." Mens en anden bedømmer fortæller om skriftlig engelsk:



Man skal lige huske, at når man retter skriftlig engelsk, så er der tre ud af fire delopgaver, hvor man har en decideret rettenøgle. Der sidder man og tæller point. Det er kun den sidste del, der er fri.

Ud over at forskelle fagene imellem har betydning for, hvordan bedømmelsen foregår, så har prøveformen, herunder om der er tale om en mundtlig eller en skriftlig bedømmelse, også betydning.

Man bedømmer noget forskelligt med standpunktskarakterer, årskarakterer og prøvekarakterer

Bedømmerne beskriver, at standpunkts- og årskarakterer er forskellige fra prøvekarakterer, i den forstand at standpunkts- og årskarakterer giver et helhedsbillede, mens prøvekarakterer giver et øjebliksbillede. Men ud over at man bedømmer noget forskelligt, så beskriver nogle bedømmere også, at karakterer også kan bruges forskelligt. Fx kan de løbende standpunktskarakterer, særligt i 8. klasse, bruges pædagogisk som et motivationsredskab ift. eleverne, mens de afsluttende års- og prøvekarakterer alene er en faglig vurdering af eleverne. Brugen af standpunktskarakterer som et motivationsredskab kan enten foregå ved at give eleverne en lidt lavere karakter, end deres faglige niveau egentlig tilsiger, for at holde dem til ilden eller ved at belønne eleverne med en lidt højere karakter, end deres faglige niveau berettiger til, for at understøtte deres motivation, hvis de eksempelvis har arbejdet hårdt på at blive bedre. Disse bedømmere vurderer dermed, hvad der vil virke

motiverende eller demotiverende for de enkelte elever, og tager højde for det, når de giver standpunktskarakterer. En bedømmer fortæller, hvordan hun lægger nogle elevers standpunktskarakterer lidt lavt for at holde dem til ilden: ”Hvis jeg ved, at deres niveau er til 12, men de ikke deltager, så giver jeg 10 for at motivere dem”. Mens en anden bedømmer fortæller:



Jeg synes jo, at man skal have noget at arbejde efter, så jeg giver ikke så meget 12 [i standpunktskarakter] ud over 12-tallet sidste år. Det er jo egentlig lidt forkert, men de skal arbejde efter noget.

Omvendt vurderer bedømmerne af andre elever, særligt fagligt udfordrede elever, der befinder sig i den lave ende af skalaen, men som arbejder hårdt, at vedblivende lave karakterer kan have en demotiverende effekt. En bedømmer fortæller, hvordan hun angående en konkret elev vurderer, at en lav karakter vil være demotiverende:



Jeg har lige nu en dreng, som jeg tvivler på, vil få mere end 00 til danskeksamen. Men jeg giver ham ikke 00 i standpunktskarakter, for han har arbejdet så hårdt og vil virkelig gerne udvikle sig. Det ville være at tage motivationen fra ham, hvis jeg gav ham 00.

Bedømmerne betragter ikke karakterskalaens relative element som noget, de skal forholde sig til, når de bedømmer

Bedømmerne er ikke af den opfattelse, at de skal forholde sig til karakterskalaens relative element, det vil sige forventningen om en normalfordeling over tid og for store populationer, når de bedømmer, men at de alene skal anvende skalaen absolut. En bedømmer siger:



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.

En anden bedømmer fortæller, at man dog kan møde bedømmere, der er af den opfattelse, at man skal tilstræbe en fordeling af karaktererne:



Man kan opleve, at læreren til prøven siger: ”Vi skal også huske at bruge hele skalaen”. Det kan jeg ikke forstå. Det der med, at hvis der er for mange, der har fået noget, så skal man sørge for, at nogle elever får noget andet. Der er nogle, der stadig hænger fast i den mærkelige holdning.

Dog peger nogle bedømmere på, 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:



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.

2.6.2 Bedømmere på det gymnasiale område

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

Bedømmerne har forskellige tilgange til bedømmelse, når det kommer til, hvordan de sporer sig ind på, hvilken karakter eleven skal have til en mundtlig eller skriftlig prøve. Nogle bedømmere beskriver, at de ikke tager udgangspunkt i et bestemt punkt på skalaen i deres bedømmelse, mens andre bedømmere beskriver, at de tager udgangspunkt i et bestemt punkt på skalaen.

En gruppe af bedømmerne beskriver, at de starter helt åbent ud, når de skal bedømme eleverne, og at de tager udgangspunkt i elevernes præstation. Nogle beskriver det eksempelvis på den måde, at de sporer sig ind på, hvilket interval på skalaen eleverne befinder sig inden for, fx ved brug af de taksonomiske niveauer, og så derfra finder frem til en konkret karakter. Nogle af de bedømmere, der gør brug af denne tilgang, beskriver, at det blandt andet handler om, at karakterer ikke er det første, de har fokus på, når de eksempelvis sidder i en mundtlig eksamenssituation, men den konkrete elevs præstation.

En anden del af bedømmerne beskriver deres bedømmelse på den måde, at de tager udgangspunkt i et bestemt sted på skalaen, og så justerer de karakteren ud fra elevens præstation. For nogle bedømmere betyder det, at de tager udgangspunkt i topkarakteren 12 og derfra trækker fra for fejl og mangler. For andre betyder det, at de tager udgangspunkt i midten af skalaen og derfra både lægger til og trækker fra. Mens det for andre igen betyder, at de starter i bunden af skalaen og derfra ser, hvor langt opad eleven kan bevæge sig, eksempelvis med hensyn til de taksonomiske niveauer.

Bedømmerne beskriver, at det i den mundtlige prøve- og bedømmelsessituation har betydning, om man er censor eller eksaminator. Dels ift., om man starter ud helt åbent eller med forventning om at ligge på en bestemt del af skalaen, fordi man som eksaminator kender eleverne, der kommer ind, og derfor kan have en gisning om, hvor de ligger, hvilket man ikke har som censor. Dels fordi eksaminator i prøvesituationen ikke alene har fokus på at bedømme eleven, men også på at rammesætte og gennemføre prøven på bedst mulig vis, mens censor i højere grad kan have sit hovedfokus på bedømmelsen undervejs i prøven og derfor blandt andet kan få noteret langt mere ned.

Bedømmelsen foregår forskelligt afhængigt af fag og prøveform

I denne evaluering er fokus på det gymnasiale område på bedømmelse i fagene dansk, engelsk og matematik. Af fokusgrupperne med bedømmere i disse fag fremgår det tydeligt, at fagenes forskellige karakter betyder, at bedømmelse i fagene også har forskellig karakter, særligt når det kommer til den skriftlige bedømmelse. Hvor man i skriftlig matematik tildeler point for hver delopgave og via en omsætningstabel omregner pointene til en karakter, så kan bedømmelsen i dansk og engelsk ikke på samme måde sættes på formel. Derfor peger nogle bedømmere på, at det af denne grund heller ikke på samme måde er meningsfuldt at tælle fejl og mangler i dansk og engelsk. Derudover er det også forskelligt fagene imellem, hvor mange forskellige parametre der skal i spil i bedømmelsen, og også hvor konkrete bedømmelseskriterierne i fagene er. Ud over at forskelle fagene imellem har betydning for, hvordan bedømmelsen foregår, så har prøveformen, herunder om der er tale om en mundtlig eller skriftlig bedømmelse, også betydning.

En del af bedømmerne oplever det som uhensigtsmæssigt, at skalaen anvendes som en mangelskala

En del af bedømmerne oplever at være blevet præsenteret for 7-trins-skalaen som en mangelskala. Blandt en gruppe af bedømmerne går det igen, at de er kritiske ift. at skulle bedømme ud fra – og dermed have fokus på – fejl og mangler hos eleverne. De beskriver, at fejl og mangler naturligvis indgår i en samlet bedømmelse af, i hvilken grad eleverne opfylder fagets mål og krav. Men at hovedfokus ud fra et pædagogisk perspektiv bør være på at afdække og give eleven mulighed for at vise alt, hvad eleven kan. I eksempelvis en mundtlig eksamenssituation understreger bedømmerne, at det må handle om at få det bedste frem i eleven – ikke at finde frem til alt det, eleven ikke kan. En bedømmer fortæller, hvordan hun oplever, at bedømmelse ud fra fejl og mangler har negativ indvirkning på eleverne:



Det er ikke et godt lærings- og menneskesyn. Det fremmer ikke elevernes lyst til at lære. Det fremmer deres iver efter, at de vil lave nul fejl og få et 12-tal. Det er en tendens, jeg både ser og læser om forskellige steder. Jeg siger ikke, at karakterskalaen er den eneste synder i den forbindelse, der er givetvis mange faktorer, der spiller ind, men jeg synes, at den er udtryk for samme tendens.

Ud fra et læringsperspektiv er der bedømmere, der oplever, at det ikke er meningsfuldt at skulle bedømme eleverne ud fra fejl og mangler og dermed have fejl og mangler som hovedfokus. Dels fordi fejl og mangler, ifølge bedømmerne, ikke er noget entydigt negativt, men også kan rumme et udviklingspotentiale. Eksempelvis hvis man til en mundtlig eksamen selv kan identificere og rette sine egen fejl. Dels fordi bedømmerne beskriver, at når det drejer sig om at bygge mennesker op, så er det centralt at bygge på det, de kan, og ikke alene på det, de ikke kan.

Ligger eleven på vippen mellem to karakterer, kommer tvivlen ofte eleven til gode

Hvis en elev ligger på vippen mellem to karakterer ved en mundtlig eksamen, så drøfter bedømmerne, hvad der taler for hhv. den ene og den anden karakter. Er der uenighed om karakteren, så vejer censors ord tungest. Men hvis der i højere grad er tale om en fælles tvivl frem for en uenighed om karakteren, så fortæller bedømmerne, at de som hovedregel lader tvivlen komme eleven til gode og vælger den højeste karakter. Hvis en elev ligger på vippen mellem to karakterer til den skriftlige eksamen, så kommer tvivlen også her eleven til gode, men uden at de to censorer nødvendigvis drøfter, hvad der taler for eller imod, idet tiden til votering ved de skriftlige eksaminer er begrænset.

Man bedømmer noget forskelligt med standpunktskarakterer, årskarakterer og prøvekarakterer

Bedømmerne beskriver, at hvor års- og prøvekarakterer alene gives ud fra en vurdering af elevernes faglige niveau og faglige præstationer, så bruges de løbende standpunktskarakterer også pædagogisk som et motivationsredskab ift. eleverne. Det kan enten ske ved at give eleverne en lidt lavere karakter, end deres faglige niveau egentlig tilsiger, for at holde dem til ilden eller ved at belønne elever med en lidt højere karakter, end deres faglige niveau berettiger til, for at understøtte deres motivation, hvis de eksempelvis har arbejdet hårdt på at blive bedre. En bedømmer siger:



Jeg kan godt gemme lidt på 12-tallerne de første to år. Hvis den helt sikkert er der, så får de jo 12. Men jeg kan være lidt bange for, at det bliver en sovepude, hvorfor jeg godt kan gemme lidt på dem indtil til sidst.

Samtidig fortæller bedømmerne også, at standpunkts- og årskarakterer kan være svære at give, særligt de mundtlige, fordi det er et mindre afgrænset og konkret defineret grundlag, de skal bedømme eleverne på, end til en konkret prøve. Nogle af de elever, bedømmerne beskriver, kan være svære at bedømme, hvad angår standpunkts- og årskarakterer, er elever, der er stille, men godt kan, når de bliver spurgt, eller elever, hvis deltagelse og svarkvalitet svinger. Ligesom det også kan være udfordrende at bedømme eleverne individuelt, når de laver gruppearbejde.

Bedømmerne betragter ikke karakterskalaens relative element som noget, de skal forholde sig til, når de bedømmer

Bedømmerne er ikke af den opfattelse, at de skal forholde sig til karakterskalaens relative element, det vil sige forventningen om en normalfordeling over tid og for store populationer, når de bedømmer, men at de alene skal anvende skalaen absolut. En bedømmer siger: "Det relative er uvedkommende for mig. Jeg har ingen forpligtigelse over for et landsgennemsnit." En anden udtaler:



Jeg opfatter karakterskalaen som absolut. Men jeg kan da godt huske, at da karakterskalaen kom, så var det meldingen fra Undervisningsministeriet, at man forventede, at 10 % skulle have 12. Jojo, 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.

2.6.3 Bedømmere på de videregående uddannelser

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

Bedømmerne har forskellige tilgange til bedømmelse, når det kommer til, hvordan de sporer sig ind på, hvilken karakter den studerende skal have til en mundtlig eller en skriftlig prøve. Nogle bedømmere beskriver, at de ikke tager udgangspunkt i et bestemt punkt på skalaen i deres bedømmelse:



Jeg tager ikke udgangspunkt i noget punkt på skalaen. Det er en helhedsbedømmelse. Det er svært at sige, hvordan man gør. Det er noget med at notere, hvad der er positive ting, og hvad der er mindre godt. Så jeg danner mig et helhedsbillede. De her karakterer udtrykker, om man er i midterste gruppe, øverste gruppe eller nederste gruppe. Det ligger til grund for diskussionen med censor eller eksaminator. Vi har ofte et skriftligt oplæg med mundtlig eksamination med én karakter.



Jeg synes ikke, at fejl og mangler fylder så meget i diskussionen. Virkeligheden er, at når man er en gammel bedømmer, så får man jo ret hurtigt en fornemmelse af, hvor det er, vi er henne i det her system. Og så er det en finjustering.



Jeg tager ikke udgangspunkt i noget punkt på skalaen. Jeg sidder mest med skriftlige opgaver, der skal forsvares mundtligt. Der er klare ting ved de skriftlige opgaver, som de skal leve op til. Fra bekendtgørelser, læringsmål etc. Så får de en karakter, eller de bliver klassificeret som over middel eller under middel. Og det kan så rykke sig ved den mundtlige. Præsterer de bedre, går karakteren op, og omvendt.

En anden del af bedømmerne beskriver deres bedømmelse på den måde, at de tager udgangspunkt i et bestemt sted på skalaen, og så justerer de karakteren ud fra den studerendes præstation. For nogle bedømmere betyder det, at de tager udgangspunkt i topkarakteren 12 og derfra trækker fra ud fra fejl og mangler. For andre betyder det, at de tager udgangspunkt i midten af skalaen og derfra både lægger til og trækker fra. Mens det for andre igen betyder, at de starter i bunden af skalaen og derfra ser, hvor langt opad eleven kan bevæge sig. Tre bedømmere fortæller:



Man lytter åbent til besvarelsens nuancer. Man falder, hvis der er fejl. Så går man fra 12 til 10. I løbet af processen kan man også stige. Det er en dynamisk proces, der går frem og tilbage. Der er forskel på, om man kender den studerende eller ej. Det kan være vanskeligt at starte på 12, hvis der er tale om en studerende, som man ved, har store udfordringer. Det er måske ikke helt korrekt, men det gør man.



Man tager principielt udgangspunkt i 12. Det vil sige, at man kommer til at sige, at man præsenterer det på en negativ måde, hvorfor man trak fra. Nogle gange er det også noget med refleksion, og så kan det godt blive lidt fluffy. Det er svært at få de studerende til at forstå, hvad det handler om. Men altså, med den her karakterskala skal vi jo tage udgangspunkt i, at den studerende kan alt, og så trække fra.



Jeg er gået imod strømmen og har valgt at tage udgangspunkt i bundkarakteren. Jeg har to mundtlige eksamensformer. I de mundtlige eksaminer, hvor de studerende kommer op og trækker et spørgsmål, ser jeg det sådan, at de studerende kommer ind i et tomt rum, som de skal fylde ud med viden. Der kan man jo ikke trække fra. Hvis de ikke får fyldt rummet ud, eller de ikke hæver sig over det elementære, så kan de ikke få fyldt nok viden på. Så på den måde starter jeg nedefra og går op efter.

Et perspektiv i fokusgrupperne er, at prøvetypen har betydning for brugen af skalaen. Ved skriftlige prøver af mere teknisk eller matematisk karakter findes der ofte rettevejledninger, hvor karakteren fastsættes efter det antal point, som den studerende opnår:



Hvis der ved skriftlige opgaver er point og rettevejledning, så er det fælles for censor og eksaminator – der har vi et fælles grundlag. Der sidder vi egentlig og lægger til og trækker fra. Det er typisk de mundtlige, hvor det er lidt sværere at sige ”der er vi der eller der”. De mundtlige eksaminer foregår jo som regel med udgangspunkt i et skriftligt produkt. Der har jeg vurderet produktet først og forventningsafstemmer med censor.



Ved skriftlige eksaminer tæller vi fra bunden og op – ved hjælp af et pointsystem. Der er hovedspørgsmål og underspørgsmål, som hver er tildelt en pointmængde. Vi starter fra 00 og arbejder os op i karaktersystemet. Når vi er i tvivl om to karakterer, ser vi på helhedsindtrykket: Er der store tomme områder eller mindre fejl hist og pist?

Bedømmerne oplever, at fejl og mangler fylder meget

Fokusgruppeinterviewene med bedømmerne på det videregående område giver indtrykket af, at fejl og mangler fylder meget i bedømmernes karaktergivning. Tre bedømmere fortæller:



Hvis jeg nu starter med at eksaminere opad, og vi ender med nogle spørgsmål på 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 haft 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.



Jeg gør dem [de studerende] opmærksomme 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.



Jeg bruger ordlyden fra trinbeskrivelsen og målene, når jeg melder tilbage til de studerende. Så undgår man nemmere klager.

Som citaterne ovenfor illustrerer, bliver bedømmernes fokus på fejl og mangler ofte koblet sammen med eksamensklager fra studerende. Indtrykket fra fokusgruppen er, at eksamensklager fylder for en del bedømmere.

Ud fra et læringsperspektiv oplever en gruppe af bedømmerne det ikke som meningsfuldt at skulle bedømme de studerende ud fra fejl og mangler. Nogle bedømmere begrundet dette med, at læring handler om at vide, hvad man allerede kan, og bygge oven på det. Andre bedømmere lægger vægt på, at fokus på fejl og mangler opbygger en kultur blandt de studerende, hvor deres hovedfokus er at undgå fejl og mangler. De to perspektiver er illustreret i citaterne nedenfor:



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.



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.



Der er jo en konsekvens af, at man har en karakterskala, der sender et signal til de studerende om, at de bare skal have gode karakterer, og måske ikke i lige så høj grad, at de faktisk skal tænke sig om. Fremdriftsreformen betyder også, at der er store konsekvenser ved ikke at bestå. Derfor kommer de studerende og spørger: "Hvad er de absolutte minimumskrav, for at jeg kan bestå?". Det er ren overlevelsestrategi ift. at komme gennem studiet uden sanktioner ift. SU. Det er svært at motivere de studerende til at studere for læringens skyld og ikke bare for at komme gennem studiet med gode karakterer.



Jeg er så ærgerlig over formuleringer med fejl og mangler. Fordi de studerende bliver så bange for at begå fejl. Men man bliver klogere af at begå fejl.

Karakterernes normalfordeling fylder meget lidt hos den enkelte bedømmer

Bedømmerne forholder sig generelt meget kritisk over for den forventede normalfordeling af karaktererne. De fortæller, at det primært er ledelsen, som har fokus på dette, men at det ikke giver mening for dem som bedømmere, da de skal bedømme hver studerende individuelt. De påpeger desuden, at der kan være stor forskel på enkelte årgange, hvilket gør det problematisk at sigte imod en bestemt fordeling.



Hvis man presser karaktererne ind i en normalfordeling, kan det være, at man ikke kan se den rigtig gode årgang, der kan være en gang imellem.



Jeg synes egentlig, at det er noget pjat. Jeg kan godt forstå, at man i gymnasiet eller folkeskolen vil have normalfordelinger. Men hvis man arbejder med absolutte karaktergivning, giver det jo ikke mening, at man arbejder med normalfordelinger. Det er svært at tage alvorligt. Vi bruger i langt højere grad målbeskrivelserne.

En pointe blandt bedømmerne er også den naturlige selektion, som sker på uddannelserne. De fortæller blandt andet, at det ofte er de dygtige studerende, der fortsætter, mens de mindre dygtige falder fra. Dette afspejles således også i karakterernes udvikling i løbet af uddannelsen.

Ligger den studerende på vippen mellem to karakterer, kommer tvivlen ofte den studerende til gode

Hvis en studerende ligger på vippen mellem to karakterer ved en mundtlig eksamen, så drøfter bedømmerne, hvad der taler for hhv. den ene og den anden karakter. Er der uenighed om karakteren, så vejer censors ord tungest.² Men hvis der i højere grad er tale om en fælles tvivl frem for en uenighed om karakteren, så fortæller bedømmerne, at de som hovedregel lader tvivlen komme den studerende til gode og vælger den højeste karakter. En bedømmer siger:

² Der findes ikke regler for, hvorvidt det er eksaminator eller censor, som først spiller ud ved votering. Kutymen ved votering er dog, at censor spiller først ud. Ved uenighed er det også censors stemme, der vejer tungest. Karakterbekendtgørelsen, § 11, bestemmer, at hvis bedømmerne efter drøftelse ikke er enige om en karakter, giver eksaminator og censor hver en karakter, og karakteren for prøven bliver gennemsnittet af disse to karakterer afrundet til nærmeste karakter på karakterskalaen.



Jeg giver hellere lidt for meget end for lidt. For nogle gange kan de blive nervøse til eksamen. Og så er der et meget stort spring fra 7 til 10, fx. Der ligger et stort ansvar hos os bedømmere.

2.7 Enkeltinterviews med bedømmere

I det følgende beskrives bedømmernes forståelse af og anvendelse af skalaen, som det er kommet til udtryk i de individuelle interviews.

2.7.1 Bedømmere på grundskoleområdet

Hovedpointerne ift. forståelse af og anvendelse af skalaen er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere fra grundskolen. I dette afsnit går en del af overskrifterne fra afsnittet om fokusgrupperne således igen, men pointerne illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

Bedømmerne har forskellige tilgange til bedømmelse, når det kommer til, hvordan de sporer sig ind på, hvilken karakter eleven skal have til en mundtlig eller en skriftlig prøve. Nogle bedømmere beskriver, at de ikke tager udgangspunkt i et bestemt punkt på skalaen i deres bedømmelse, mens andre bedømmere beskriver, at de tager udgangspunkt i et bestemt punkt på skalaen.

En del af bedømmerne beskriver således, at de starter helt åbent ud, når de skal bedømme eleven, og tager udgangspunkt i elevens præstation. Derfra sporer de sig typisk ind på, om eleven ligger i toppen, midten eller bunden af skalaen, for så derfra at finde frem til den konkrete karakter. Bedømmerne fortæller, at drøftelsen af elevens karakter med censor/eksaminator efter en mundtlig prøve også ofte tager udgangspunkt i en snak om, hvilket interval på karakterskalaen eleven befinder sig inden for, hvorefter den konkrete karakter bliver fastlagt – eventuelt ved brug af vurderingskriterier og retteskemaer. En bedømmer fortæller:



Vi starter ikke med 12 og trækker fra. Det er en vurdering af, hvordan de ligger, når de er helt færdige. Vi starter typisk med at sige, om det er middel, under middel eller over middel, og så vurderer vi de forskellige dele – om vi kom forbi dem.

Andre bedømmere beskriver deres bedømmelse på den måde, at de tager udgangspunkt i et bestemt sted på skalaen, og så justerer de karakteren ud fra elevens præstation. Nogle bedømmere starter omkring midten og justerer op eller ned. En bedømmer fortæller:



Jeg starter ikke fra toppen og trækker fra. Sådan tænker jeg ikke. Jeg er open-minded. Jeg starter lidt i midten, og så kan jeg høre i starten, om det går op eller ned, og se, hvor det bærer hen ad.

Andre bedømmere beskriver, at de tager udgangspunkt i 12 og så justerer derfra. Det gælder særligt, når bedømmerne er censorer og således ikke kender de elever, de skal bedømme. Når bedømmerne som eksaminatorer kender eleverne, har de et andet kendskab til, på hvilket niveau eleverne cirka kan forventes at ligge. En bedømmer fortæller om tilgangen som censor:



Når der kommer en ny elev ind, tænker jeg 12 med det samme. Det får de så ikke alle, men de har mulighed for det. Som udgangspunkt er det 12, som de skal forsvare.

Nogle bedømmere oplever, at de meget hurtigt kan spore sig ind på elevens niveau, mens andre bedømmere beskriver, at det er noget, de sporer sig ind på løbende undervejs under den mundtlige prøve eller i gennemgangen af en opgave, og som først til slut ligger helt fast. Dette kan dels hænge sammen med, at det er forskelligt, hvorvidt bedømmerne tænker i karakterer og bedømmelse fra starten af, og dels hænge sammen med, at niveauet på elevens mundtlige såvel som skriftlige præstation kan variere undervejs. En bedømmer fortæller:



Det er ikke sådan, at jeg kender karakteren efter et minut. Men det er undervejs i forløbet og måske mest til slut. Der kan godt være stor forskel på deres egen præsentation og samtalen – nogle får talt sig op og andre omvendt.

Bedømmelse af delementer i en samlet helhedsvurdering

Bedømmelsen af eleverne skal foretages som en helhedsvurdering af deres præstation. Bedømmerne beskriver generelt, at deres bedømmelse af eleverne derfor er en vurdering af en række delementer, som udgør fokuspunkter i bedømmelsen af eleverne. Hvilke delementer der er i fokus, hænger blandt andet sammen med, om der er tale om en mundtlig eller en skriftlig prøve. En bedømmer fortæller om den mundtlige prøve i engelsk:



Både som censor og som eksaminator sporer jeg mig ind på niveauet ved at se på ordforråd og grammatik, om de taler sproget, om de er med i samtalen, eller det er mig og eksaminator, der kører den, og om eleven har noget at sige og mene. Kan de gennemføre samtalen og præsentere emnet selv, og bruger de deres disposition?

Nogle bedømmere oplever ikke ensidigt fokus på fejl og mangler som meningsfuldt

Bedømmerne beskriver generelt ikke, at de anvender skalaen som en mangelskala, men som ovenfor beskrevet, at de i højere grad anvender skalaen som en helhedsvurdering af elevens præstation både ift. fejl og mangler og ift. korrekte svar og udmærkelser. Nogle bedømmere beskriver, at hvor meget de kigger efter det ene og det andet, hænger sammen med, hvilket niveau eleven befinder sig på. Ligger eleven i toppen, så kigger man efter, om der er væsentlige fejl og mangler, der gør, at det ikke kan blive en topkarakter. Men ligger eleven i bunden af skalaen, har man fokus på at finde de gode elementer i elevens præstation. En bedømmer siger:



Hvis vi er i toppen, og der er mange fejl og mangler, ryger vi ned, altså fra 12 til 10. Mens det i bunden mere er kvaliteterne, der kan rykke dem fra 2 til 4. Så fokus er forskelligt.

Blandt flere bedømmere går det igen, at de ikke finder det pædagogisk meningsfuldt at fokusere ensidigt på fejl og mangler. En bedømmer fortæller:



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.

Bedømmelsen foregår forskelligt afhængigt af fag og prøveform

I denne evaluering er fokus på grundskoleområdet på bedømmelse i fagene dansk, engelsk og matematik. Af fokusgrupperne med bedømmere i disse fag fremgår det tydeligt, at fagenes forskellige karakter betyder, at bedømmelse i fagene også har forskellig karakter, særligt når det kommer til den skriftlige bedømmelse. Hvor man i matematik og dele af engelsk har en egentlig rettenøgle, så gælder det samme ikke for dansk skriftlig fremstilling. Bedømmerne understreger dog, at selvom bedømmelsen af den skriftlige dimension i eksempelvis matematik sker ved pointgivning, så spiller

bedømmeren stadig en rolle i den konkrete vurdering af elevens præstation, blandt andet fordi eleven ikke blot vurderes ud fra rigtige og forkerte svar, men også ud fra kompetencer såsom formidlingen i opgaven. En bedømmer i matematik fortæller:



Vi bruger en centralt udviklet rettevejledning, som jeg støtter mig til. Jeg tildeler point ud fra rettevejledningen, og så giver det en sum til sidst. Hvis pointene ligger mellem to karakterer, kigger jeg på det kommunikative og ser, om der er noget, der trækker op eller ned.

Ud over at forskelle fagene imellem har betydning for, hvordan bedømmelsen foregår, så har prøveformen også betydning, herunder om der er tale om en mundtlig eller en skriftlig bedømmelse. Bedømmerne understreger, at en væsentlig forskel på skriftlig og mundtlig bedømmelse er, at der kun er én bedømmer ved skriftlige prøver. Det oplever bedømmerne generelt, gør bedømmelsen af de skriftlige prøver sværere, og mener samtidig, at det kan gå ud over ensartetheden i bedømmelsen. En bedømmer siger:



Karaktergivningen er sværere ved skriftlige prøver, fordi vi sidder alene med det. Jeg synes, det skal være både lærer og censor. Det er ærgerligt, at det ikke er sådan, både som lærer og som censor. Det er svært at vurdere det. Der kan være forskellige ting, og vi kan ikke være 100 % objektive. Der er der større risiko for en forkert beslutning pga. manglende sparring.

Man bedømmer noget forskelligt med standpunktskarakterer, årskarakterer og prøvekarakterer

De interviewede bedømmere beskriver, at standpunkts- og årskarakterer er forskellige fra prøvekarakterer, i den forstand at standpunkts- og årskarakterer giver et helhedsbillede, mens prøvekarakterer giver et øjebliksbillede. Bedømmerne beskriver også, at der indgår andre parametre i bedømmelsen ved løbende standpunktskarakterer end ved eksamen. Eksempelvis beskriver bedømmerne, at elevernes arbejdsindsats og deltagelse kan komme til at spille ind på de løbende standpunktskarakterer, for siger en dygtig elev ikke noget, så vil det have betydning for standpunktskarakteren, også selvom elevens faglige niveau er højt. En bedømmer siger:



Det svære er, at nogle elever er stille, men fagligt dygtige. Men så er det svært at nå at høre dem. Så man kommer til at give karakter ud fra, hvor arbejdsomme de er – selvom man skal gøre det ud fra det, de kan – men det er svært. Men jeg prøver at lave fremlæggelser eller gå rundt og snakke med dem, hvis de ikke er glade for snak i plenum. Men de mundtlige standpunktskarakterer er sværest.

Citatet peger også på den gennemgående pointe blandt bedømmerne, at de mundtlige standpunkts- og årskarakterer kan være nogle af de sværeste karakterer at give, fordi det er et mindre afgrænset og konkret defineret grundlag, de skal bedømme eleverne på, end til en konkret prøve eller en skriftlig opgave.

2.7.2 Bedømmere på det gymnasiale område

Hovedpointerne ift. forståelse af og anvendelse af skalaen er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere fra de gymnasiale uddannelser. I dette afsnit går en del af overskrifterne fra afsnittet om fokusgrupperne således igen, men pointerne illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

Bedømmerne har forskellige tilgange til bedømmelse, når det kommer til, hvordan de sporer sig ind på, hvilken karakter eleven skal have til en mundtlig eller en skriftlig prøve. Nogle bedømmere

beskriver, at de ikke tager udgangspunkt i et bestemt punkt på skalaen i deres bedømmelse, mens andre bedømmere beskriver, at de tager udgangspunkt i et bestemt punkt på skalaen.

En del af de bedømmere, der beskriver, at de starter helt åbent ud, fortæller, at de ikke tænker så meget i karakterer til at starte med, men i højere grad i faglige mål og taksonomiske niveauer. To bedømmere siger:



Jeg starter helt åbent ud. Jeg tænker ikke så meget over skalaen undervejs, men foretager devalueringer ift. de spørgsmål, eleven har trukket. Hvor godt gik den her del af eksamen?



Jeg tænker meget sjældent i karakterer og tal. Jeg tænker mere i faglige mål. Der er naturligt nogle faglige mål, der hører til det emne, de har trukket.

Andre bedømmere beskriver, at de tager udgangspunkt i et punkt på skalaen, når de skal bedømme eleverne. En bedømmer fortæller, at han starter lavt på skalaen og ud fra de taksonomiske niveauer giver eleven mulighed for at bevæge sig op ad skalaen:



Som eksaminator starter jeg nedefra med det basale. Hvis de sidder med et digt: Har de forstået, hvad det går ud på? Så er de i hvert fald på 4. Har de fået den forståelse på baggrund af en analysetilgang, de også kan gøre rede for, så ligger vi måske på 7. Kan de reflektere over, hvad man ellers kunne gøre, så er vi lidt højere oppe med hensyn til de taksonomiske niveauer. Så det, jeg giver karakterer efter, er, hvor højt de kommer op med hensyn til de taksonomiske niveauer.

En anden bedømmer fortæller, at han starter højt på skalaen:



Jeg starter forholdsvis højt på skalaen. Særligt hvis eleven kommer med en god disposition, så er jeg som udgangspunkt positivt stemt og ligger i hvert fald på en 10^{er}.

Bedømmelse af delelementer i en samlet helhedsvurdering

Bedømmelsen af eleverne skal ske som en helhedsvurdering af deres præstation. Bedømmerne beskriver generelt, at deres bedømmelse af eleverne derfor foretages som en vurdering af en række delelementer, som udgør fokuspunkter i bedømmelsen af eleverne og er hentet fra fagets faglige mål og fagets genstandsfelter. Det gælder ved både skriftlig og mundtlig bedømmelse. Bedømmerne kan bruge forskellige officielle bedømmelsesværktøjer til at udpege elementer i helhedsvurderingen, men nogle bedømmere fastlægger også selv fokuspunkter, eksempelvis for at gøre fokuspunkterne mere konkrete end de officielt formulerede. En bedømmer fortæller om bedømmelsen af delelementer ved den mundtlige prøve i matematik:



Jeg har tre fokuspunkter ift. at bedømme undervejs i den mundtlige prøve: 1) hvor god eleven er til at argumentere, 2) hvor god eleven er til at koble grafer og figurer til formler og beskrivelser, og 3) hvor god eleven er til at inddrage forskellige andre værktøjer i sin fremlæggelse.

En anden bedømmer fortæller om bedømmelse ud fra delelementer i dansk:



Vi har fået en tredelt vejledning fra fagkonsulenten om, hvad vi skal bedømme dem på. Der er en skriftlig dimension, en indholdsdimension og en faglig dimension.

Nogle bedømmere oplever ikke ensidigt fokus på fejl og mangler som meningsfuldt

Bedømmerne beskriver generelt, at de ikke anvender skalaen ensidigt som en mangelskala, men i højere grad foretager en helhedsvurdering af elevens præstation, der involverer både positive og

negative elementer. Generelt finder bedømmerne det dog mest konstruktivt at lægge hovedvægten på de ting, eleven kan, frem for det, eleven ikke kan eller ikke siger, for at få det bedste frem i eleven og give eleven mulighed for at vise, hvad vedkommende kan. Dette kan bedømmerne naturligtvis i højere grad understøtte ved en mundtlig prøve, hvor bedømmerne er med til at styre samtalen og stille spørgsmål, end ved en skriftlig prøve, hvor bedømmerne alene kan bedømme eleven på det, som eleven selv har skrevet. En bedømmer forklarer:



Jeg prøver ikke at bore meget i det, eleverne ikke kan. Det er ikke sjovt. Heller ikke for eleverne. De skal også føle, at de får lov at vise, hvad de kan. Så er der nogle steder, hvor de bliver usikre, og så spørger jeg lidt ind til det, skriver en lille note på papiret, og så videre til noget andet. Så jeg fokuserer på at prøve at finde så meget godt som muligt.

En anden bedømmer forklarer, hvorfor han har fokus på det fornuftige og korrekte, eleven siger, frem for på det, eleven ikke får med:



Jeg ser meget på plussiden. Som lærer er ens faglige overblik meget større end elevens, så der er så mange ting, eleven kunne sige. Jeg har ikke en liste over ting, eleven skal sige. Men det handler mere om, hvad eleven kan sige, som er godt og rigtigt og har noget med spørgsmålet at gøre. Eleverne har 20 minutter til at præsentere det, de har lært på to-tre år. Så de har lært meget mere, end de kan nå at sige.

Nogle bedømmere beskriver også, at det er vigtigt med et nuanceret blik på fejl og mangler. For det første drejer det sig om at forholde sig til, hvilken type fejl der er tale om, og på hvilket taksonomisk niveau. En bedømmer fortæller:



Vi holder øje med både fejl og mangler og positive ting. For det trækker i hver sin retning. Men der er forskel på, om du har fejl og mangler på de højeste taksonomiske niveauer, for der kan du godt få en høj karakter alligevel. Så det handler også om, hvilket taksonomisk niveau fejlen ligger på.

En anden bedømmer fortæller, at fejl ikke nødvendigvis bør trække ned, men også kan indgå som et element i, at eleven rent faktisk reflekterer:



Hvis eleverne selv kan se, at der er en fejl, og kan rette den, så er det fint og ikke noget, der trækker ned på nogen måde. For man vil jo gerne frem til en dialog, hvor man taler matematik.

Bedømmelse foregår forskelligt afhængigt af fag og prøveform

I denne evaluering er fokus på det gymnasiale område på bedømmelse i fagene dansk, engelsk og matematik. Af interviewene med bedømmere i disse fag fremgår det tydeligt, at fagenes forskellige karakter betyder, at bedømmelse i fagene også har forskellig karakter, særligt når det kommer til den skriftlige bedømmelse. Hvor man i skriftlig matematik tildeler point for hver delopgave og via en omsætningstabel omregner pointene til en karakter, så kan bedømmelsen i dansk og engelsk ikke på samme måde sættes på formel. Derfor peger nogle bedømmere på, at det af denne grund heller ikke på samme måde er meningsfuldt at tælle fejl og mangler i dansk og engelsk. En bedømmer i dansk siger:



Jeg har hørt det der med fejltælling. Nå, så var der syv fejl, det giver den karakter. Det kan man ikke i dansk. Det giver ikke mening for mig i hvert fald.

Derudover er det også forskelligt fagene imellem, hvor mange forskellige parametre der skal i spil i bedømmelsen, og også hvor konkrete bedømmelseskriterierne er.



I matematik er det næsten nemt at bedømme skriftlige opgaver. I engelsk kan det godt være sværere at sige, hvilke faglige mål der er opfyldt. Der handler det om: Er indholdet i orden, har de foretaget den analyse, de skal, har de taget faglige begreber med ind? Men også formen, altså overholder de formen for den skrivegenre, de er blevet bedt om? Og hvor korrekt er sproget?

Nogle bedømmere i matematik understreger dog, at selvom karaktergivningen sker ud fra pointgivning, så drejer det sig ikke alene om at bedømme rigtige og forkerte svar, men også om at bedømme elevernes kompetencer i bredere forstand.



Skriftlige opgaver i matematik er delt op i spørgsmål med underspørgsmål. Hvert underspørgsmål kan give fra 0 til 10 point. Man vurderer, om eleven har regnet rigtigt, men også, om der er en fornuftig tekst, om figurer og symboler er anvendt korrekt, og hvordan vedkommende har været til at bruge og forklare hjælpemidler.

2.7.3 Bedømmere på de videregående uddannelser

Hovedpointerne ift. forståelse og anvendelse af karakterskalaen er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere på de videregående uddannelser. I dette afsnit vil der derfor være pointer, der går igen, men de illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere tager udgangspunkt i et punkt på skalaen, andre gør ikke

Bedømmerne har forskellige tilgange til bedømmelse, når det kommer til, hvordan de sporer sig ind på, hvilken karakter den studerende skal have til en mundtlig eller skriftlig prøve. Nogle bedømmere beskriver, at de ikke tager udgangspunkt i et bestemt punkt på skalaen i deres bedømmelse, mens andre tager udgangspunkt i et bestemt sted på skalaen og trækker fra og lægger til:



Jeg begynder ikke med et afsæt i toppen af karakterskalaen. Jeg tror, at jeg ligger på et gennemsnit, og så noterer jeg plusser og minusser i margenen imens. Og så tegner jeg mentalt et billede af, om det er højt eller lavt.



Det gjorde jeg i starten, da jeg begyndte at undervise. Så startede jeg ligesom på et 12-tal, og så måtte jeg så trække fra. Nu er det mere sådan, at jeg læser det og vurderer det ud fra det, de skal kunne, ud fra studieordningen og opgavebeskrivelsen, og så har jeg måske allerede dannet mig et billede, for jeg kender efterhånden skalaen så godt, at jeg allerede har dannet mig en idé – og så går jeg bare ind og dobbelttjekker mig selv. Så stiller jeg tit mig selv det spørgsmål, at hvis jeg synes, at det ligger på en 7'er, hvorfor er det så ikke en 10'er? Så egentlig kan man godt sige, at jeg starter fra toppen af. Så siger jeg "hvorfor kan det ikke være en 10'er?", og så læser jeg ligesom beskrivelsen af både 10 og 7 for at se, hvorfor det ikke skulle være det.

Bedømmelse ved mundtlige prøver med et skriftligt element

Mange prøver på de videregående uddannelser er mundtlige prøver med et skriftligt element. Et perspektiv i interviewene er, at det er forskelligt, hvorvidt bedømmerne går ind til den mundtlige prøve med en karakter for det skriftlige, eller om de laver en mere grov kategorisering af det skriftlige produkt. Nedenfor er et eksempel fra en bedømmer, som skriver en konkret karakter ned på baggrund af gennemgangen af det skriftlige produkt, og en, som benytter sig af en kombination af de to tilgange:



Og så skriver jeg selvfølgelig nogle kommentarer og uddybende spørgsmål til de studerende for at se, om jeg kan trække dem op. Hvis jeg er i tvivl om, hvorvidt de har forstået et eller andet, eller om de opfylder et eller andet krav, så skriver jeg et spørgsmål ned, som jeg kan stille for at afdække, om de har forstået det. Så de på den måde kan trække sig selv op, kan man sige. Og så skriver jeg en karakter ned, hvis jeg skulle give dem karakter her og nu.



Først tænker jeg: "Er det her middel eller er det over eller under?". En grov kategorisering. Og når jeg så har lavet den, siger jeg så: "Okay, det er i hvert fald ikke mere end den her karakter." Eller også tænker jeg: "Det er i hvert fald mere end den her." Tit har jeg faktisk skrevet et tal på mit papir, inden jeg går ind. Og det har jeg lidt for at holde mig selv fast på det. Jeg kan godt rykkes, men jeg vil gerne holde fast i, hvad det var jeg tænkte, da jeg læste rapporten.

Bedømmerne fortæller også, at de løbende i deres læsning af det skriftlige produkt skriver spørgsmål ned, som skal afdækkes i den mundtlige del for at afdække, om den studerende lever op til læringsmålene. En bedømmer fortæller:



Det er en mundtlig fremstilling på baggrund af et projekt. Det betyder, at vi egentlig ikke skal give projektet en karakter, men der ligger et forventningsniveau. Man afprøver projektet mundtligt. Ofte får man ikke den karakter, som projektet svarede til. Man får ofte en karakter lige et niveau højere end det, man afleverede, fordi man til eksamen netop afprøver fejl og mangler. Når jeg læser projektet, så prøver jeg at lede efter fejl og mangler. Det vil sige, at jeg gennem den proces får et indtryk af niveauet på projektet. Ved en af de svagere opgaver i går noterede jeg nogle ting, der ikke fungerede ved projektet.

Bedømmerne fortæller videre, at de ofte lige snakker kort sammen om deres vurdering af det skriftlige produkt inden den mundtlige prøve, herunder hvad der er særligt vigtigt at afdække i den mundtlige prøve, hvilket er illustreret i citaterne nedenfor:



Vi har en meget kort proces. Der ""aligner vi. Eller vi ""aligner ikke, men der er det sådan lige "okay, hvor er du henne i universet?". Det taler vi om. Og hvis der så er et eller andet, hvor vi tænker: "Ej, det der, det manglede da, gjorde det ikke?". Så vi lige er lidt ""alignet, når vi går ind, men det er meget kort.



Vi taler sammen inden. Så er det lidt forskelligt, om det er eksaminator eller censor, der lægger ud og siger: "Det ligger nok der". Men de studerende må ikke få nogen karakter for den skriftlige eksamen. Det skal være en mundtlig eksamen med udgangspunkt i projektet. Men vi taler om niveauet, inden vi går i gang.

Bedømmerne er opmærksomme på, at skalaen har fokus på fejl og mangler

De interviewede bedømmere er meget bevidste om, at trinbeskrivelsen tager udgangspunkt i fejl og mangler:



Den er lagt an på, at man starter fra toppen og tæller ned derfra. Det er sådan set også det, jeg har prøvet at gøre. Men der er nogle ting, der kan tælle op igen. Men med sådan et fag som det her med et relativt givet pensum fungerer det vældig godt at sige: "Det og det og det skal med. Og hvis det og det mangler, så bliver det den her karakter."



Jeg starter jo på 12, som man skal. Derfra trækker jeg så fra. Jo flere fejl og mangler, og jo værre de performer, des mere rykker jeg ned ad skalaen.

Ud fra et læringsperspektiv giver en gruppe af bedømmerne udtryk for, at de ikke oplever det som meningsfuldt at skulle bedømme de studerende ud fra fejl og mangler. Dette begrundes blandt andet med, at læring handler om at vide, hvad man allerede kan, og bygge oven på det. En bedømmer fortæller fx:



Jeg synes, at det er helt forfærdeligt. Det er en forkert måde at tænke på. Jeg synes, at det skulle være omvendt. Man skulle lede efter, hvad der er godt. Det gør vi jo i og for sig også lidt, idet jeg også leder efter, hvad de kan. Hvis der er noget, de ikke kan, er det ikke sådan, at jeg bliver ved med at bore i det. Så går jeg videre til et andet område. Det er en destruktiv måde at gå til eksamen på. Jeg synes også, der er for meget kontrol til eksamen nu. Man skal lære noget til eksamen, men det er blevet sværere. Når man leder efter fejl og mangler, så er det ikke samme type dialog. Det er meget kontrollerende.

I interviewene med bedømmerne og censorerne tog vi udgangspunkt i en bestemt eksamen, som de lige havde afholdt. Et tema i interviewene var derfor votering. Det er et gennemgående perspektiv i interviewene, at censor og eksaminator starter med at drøfte, hvad der gik godt, og hvad der gik dårligt, i den mundtlige prøve, for derefter at diskutere en konkret karakter. Bedømmerne fremhæver, at det er vigtigt, at de har snakket om, hvad der specifikt taler for en bestemt karakter, så den studerende får en konkret og præcis tilbagemelding. En bedømmer fortæller:



Vi har jo begge lagt os fast på et eller andet. Vi har en idé om en karakter. Så snakker vi om, hvad der var godt og skidt, og hvor meget de har forstået, etc. Og så snakker vi om, hvilke karakterer vi begge to har givet. Og det er jo censor, der har det sidste ord, men vi prøver selvfølgelig at diskutere ud fra, hvordan vi har forstået det, og så bruger jeg faktisk altid trinbeskrivelsen. Vi kigger meget på de der formuleringer i vores diskussion, særligt hvis vi bare er den mindste smule i tvivl.



Hvis vi tager den skala, som vi har nu, ift. den gamle skala: Fordi karaktererne er så relativt brede, og der er ret stor forskel på at få 7 eller 10 eller 4 eller 7, så skal vi være skarpere i vores tilbagemelding. For springet er jo stort. Så vi bruger trinbeskrivelsen i voteringen, fordi vi skal være skarpe på, at vi rammer rigtigt, også af hensyn til den studerende. Det nytter jo ikke noget, at vi siger til en studerende, at vedkommende har fået 7, og at vi så ikke kan argumentere for det, når vedkommende spørger hvorfor. Så må vi jo være skarpe på, hvad det betyder, hvis vi giver et 7-tal, et 4-tal etc.

Ligger den studerende på vippen mellem to karakterer, kommer tvivlen ofte den studerende til gode

Hvis en studerende ligger på vippen mellem to karakterer ved mundtlig eksamen, så drøfter bedømmerne, hvad der taler for hhv. den ene og den anden karakter. For nogle bedømmere kommer tvivlen den studerende til gode.

3 Bedømmelsesredskaber til rådighed

3.1 Baggrund og undersøgelsesspørgsmål

Karakterskalaen er ikke det eneste redskab bedømmerne skal forholde sig til i karaktergivning. Skalaen skal anvendes som en absolut skala. Dvs. at bedømmelsen sker ud fra målopfyldelse, og at der som udgangspunkt skal ske en løbende justering af målene med udgangspunkt i den population, som ventes at bestå eksamen. Bedømmerne har en række forskellige værktøjer til rådighed jf. tabel 3.1 nedenfor.

TABEL 3.1

Oversigt over bedømmelsesværktøjer (ud over karakterbekendtgørelse)

Niveau	Bedømmelsesværktøjer
Grundskole	- Fælles Mål - Bekendtgørelse om Folkeskolens Prøver, herunder vurderingskriterier fordelt på fag - Prøvevejledninger med bl.a. 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 i de almene og studieforbereende 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 – fastsat 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 i 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)
Videregående uddannelser	- Studieordning med målbeskrivelser - Uddannelsesbekendtgørelser

I dette kapitel sætter vi fokus på følgende undersøgelsesspørgsmål:

- Hvordan anvendes bedømmelsesværktøjer i karaktergivning?
- Hvilke bedømmelsesværktøjer er til rådighed ved mundtlige og skriftlige prøver/eksaminer?
- Særligt for videregående uddannelser: Er der læringsmålbeskrivelser ved alle prøver? Er der behov for andre redskaber ved afsluttende opgaver?

3.2 Survey blandt bedømmere på grundskoleområdet

Sammenfatning

Surveyen blandt bedømmere på grundskoleområdet viser overordnet set, at:

- At bedømmerne i varierende grad anvender bedømmelsesværktøjerne som er til rådighed ved mundtlige og skriftlige prøver
- At mellem 62 % og 80 % af bedømmerne er meget enig eller enige i, at bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis og at de giver et solidt udgangspunkt for bedømmelsen
- At alle bedømmere vurderer elevernes faglige kompetence ved afgivelse af standpunkts/årskarakter, og mere end halvdelen vurderer også elevernes deltagelse
- Størstedelen af bedømmerne er enige i, at deres medbedømmere typisk lægger det faglige niveau samme sted som dem, men under halvdelen at bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring.

Gennemgang af resultater vedr. skriftlige prøver

Tabel 3.2 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved skriftlige prøver. Nedenfor oplystes andelen, der har svaret, at de bruger værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 87 % bruger i høj grad "Rettevejledningen"
- 70 % bruger i høj grad "Vurderingskriterierne og den vejledende karakterbeskrivelse i prøvevejledningen"
- 62 % bruger i høj grad "Vurderingsskema der tager udgangspunkt i den vejledende karakterbeskrivelse"
- 41 % bruger i høj grad "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 22 % bruger i høj grad "Beskrivelserne i fælles mål af kompetencemål samt de vejledende videns- og færdighedsmål"
- 16 % bruger i høj grad "Den vejledende læseplan"
- 13 % bruger i høj grad "Rettegrupper hvor kollegaer retter skriftlige opgaver sammen".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har angivet, at de i høj grad gør brug af ”Den generelle karakterbeskrivelse af 7-trins-skalaen”, er statistisk signifikant forskellig mellem bedømmere i engelsk (50 %) og bedømmere i matematik (36 %).
- Andelen, der har angivet, at de i høj grad gør brug af ”Vurderingskriterierne og den vejledende karakterbeskrivelse i prøvevejledningen”, er statistisk signifikant forskellig mellem bedømmere i dansk (79 %) og bedømmere i matematik (63 %), såvel som mellem bedømmere i engelsk (79 %) og bedømmere i matematik (63 %).
- Andelen, der har angivet, at de i høj grad gør brug af ”Vurderingsskema der tager udgangspunkt i den vejledende karakterbeskrivelse”, er statistisk signifikant forskellig mellem bedømmere i dansk (52 %) og bedømmere i engelsk (77 %), såvel som mellem bedømmere i engelsk (77 %) og bedømmere i matematik (61 %).
- Andelen, der har angivet, at de i høj grad gør brug af ”Rettevejledningen”, er statistisk signifikant forskellig mellem bedømmere i dansk (76 %) og bedømmere i matematik (97 %), såvel som mellem bedømmere i engelsk (74 %) og bedømmere i matematik (97 %).

TABEL 3.2

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver karakter ved en skriftlig prøve? Procent (antal)

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	14% (53)	27% (102)	34% (126)	17% (64)	8% (29)	0% (0)	100% (374)
Vurderingskriterierne og den vejledende karakterbeskrivelse i prøvevejledningen	33% (123)	37% (139)	21% (79)	6% (23)	3% (10)	0% (0)	100% (374)
Vurderingsskema der tager udgangspunkt i den vejledende karakterbeskrivelse	30% (114)	32% (119)	22% (83)	9% (35)	6% (23)	0% (0)	100% (374)
Rettevejledningen	60% (226)	27% (100)	9% (33)	2% (8)	2% (7)	0% (0)	100% (374)
Beskrivelserne i fælles mål af kompetencemål samt de vejledende videns- og færdighedsmål	7% (27)	15% (55)	33% (125)	30% (111)	15% (56)	0% (0)	100% (374)
Den vejledende læseplan	6% (24)	10% (36)	28% (103)	38% (143)	18% (68)	0% (0)	100% (374)
Rettegrupper hvor kollegaer retter skriftlige opgaver sammen	4% (15)	9% (32)	26% (97)	32% (119)	30% (111)	0% (0)	100% (374)

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Andet	6%	2%	2%	0%	1%	89%	100%
	(22)	(8)	(6)	(1)	(3)	(334)	(374)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

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 2019

Tabel 3.3 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet ”I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende?”. Nedenfor oplystes andelen, der har svaret, at de bruger værktøjet i høj grad (enten ”I meget høj grad” eller ”I høj grad”):

- 81 % bruger i høj grad ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”
- 26 % bruger i høj grad ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”
- 26 % bruger i høj grad ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 20 % bruger i høj grad ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har angivet, at de i høj grad baserer deres bedømmelse på ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”, er statistisk signifikant forskellig mellem bedømmere i dansk (68 %) og bedømmere i engelsk (88 %), såvel som mellem bedømmere i dansk (68 %) og bedømmere i matematik (88 %).

TABEL 3.3

I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	53 % (200)	28 % (106)	13 % (48)	3 % (12)	1 % (4)	1 % (4)	100 % (374)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	11 % (41)	15 % (57)	18 % (67)	16 % (61)	23 % (85)	17 % (63)	100 % (374)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	4 % (15)	16 % (60)	27 % (102)	17 % (62)	33 % (124)	3 % (11)	100 % (374)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	8 % (29)	18 % (66)	30 % (111)	24 % (90)	18 % (66)	3 % (12)	100 % (374)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

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.

Tabel 3.4 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved skriftlige prøver. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 82 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 79 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 59 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 30 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 14 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles"
- 11 % er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen", er statistisk signifikant forskellig mellem bedømmere fra dansk (66 %) og bedømmere fra matematik (89 %), såvel som mellem bedømmere fra engelsk (72 %) og bedømmere fra matematik (89 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes", er statistisk signifikant forskellig mellem bedømmere fra engelsk (39 %) og bedømmere fra matematik (22 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles", er statistisk signifikant forskellig mellem bedømmere fra dansk (26 %) og bedømmere fra engelsk (6 %), såvel som mellem bedømmere fra dansk (26 %) og bedømmere fra matematik (12 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere fra dansk (65 %) og bedømmere fra matematik (87 %), såvel som mellem bedømmere fra engelsk (71 %) og bedømmere fra matematik (87 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse", er statistisk signifikant forskellig mellem bedømmere fra dansk (49 %) og bedømmere fra matematik (69 %).

TABEL 3.4

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	43 % (161)	39 % (145)	13 % (47)	3 % (12)	1 % (3)	2 % (6)	100 % (374)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	8 % (29)	22 % (81)	34 % (126)	24 % (91)	7 % (27)	5 % (20)	100 % (374)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	2 % (9)	12 % (45)	31 % (116)	37 % (139)	11 % (41)	6 % (24)	100 % (374)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	35 % (131)	44 % (164)	13 % (50)	5 % (18)	1 % (3)	2 % (8)	100 % (374)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse	17 % (65)	42 % (158)	27 % (102)	6 % (23)	1 % (4)	6 % (22)	100 % (374)
Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser	1 % (5)	10 % (39)	17 % (64)	27 % (101)	39 % (144)	6 % (21)	100 % (374)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

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

Gennemgang af resultater vedr. mundtlige prøver

Tabel 3.5 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved mundtlige prøver. Tabellen viser at:

- Bedømmere af mundtlige prøver (79 %) bruger ”I høj grad” eller ”I meget høj grad” ”Vurderingskriterierne og den vejledende karakterbeskrivelse i prøvevejledningen”
- Bedømmere af mundtlige prøver (72 %) bruger ”I høj grad” eller ”I meget høj grad” ”Vurderings-skema der tager udgangspunkt i den vejledende karakterbeskrivelse”
- Bedømmere af mundtlige prøver (62 %) bruger ”I høj grad” eller ”I meget høj grad” ”Den generelle karakterbeskrivelse af 7-trins-skalaen”
- Bedømmere af mundtlige prøver (32 %) bruger ”I høj grad” eller ”I meget høj grad” ”Beskrivelserne i fælles mål af kompetencemål samt de vejledende videns- og færdighedsmål”
- Bedømmere af mundtlige prøver (24 %) bruger ”I høj grad” eller ”I meget høj grad” ”Den vejledende læseplan”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der "I høj grad" eller "I meget høj grad" bruger "Den generelle karakterbeskrivelse af 7-trins-skalaen", er statistisk signifikant forskellig mellem bedømmere i dansk (58 %) og bedømmere i matematik (73 %), såvel som mellem bedømmere i engelsk (56 %) og bedømmere i matematik (73 %).
- Andelen, der "I høj grad" eller "I meget høj grad" bruger "Vurderingsskema der tager udgangspunkt i den vejledende karakterbeskrivelse", er statistisk signifikant forskellig mellem bedømmere i dansk (69 %) og bedømmere i matematik (86 %).
- Andelen, der "I høj grad" eller "I meget høj grad" bruger "Beskrivelserne i fælles mål af kompetencemål samt de vejledende videns- og færdighedsmål", er statistisk signifikant forskellig mellem bedømmere i dansk (25 %) og bedømmere i matematik (46 %), såvel som mellem bedømmere i engelsk (23%) og bedømmere i matematik (46 %).

TABEL 3.5

I hvilken grad gør du brug af følgende bedømmelsesværktøjer når du afgiver karakter ved en mundtlig prøve?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	24% (97)	38% (155)	29% (116)	8% (32)	1% (6)	0% (0)	100% (406)
Vurderingskriterierne og den vejledende karakterbeskrivelse i prøvevejledningen	39% (157)	40% (161)	18% (72)	4% (16)	0% (0)	0% (0)	100% (406)
Vurderingsskema der tager udgangspunkt i den vejledende karakterbeskrivelse	36% (147)	36% (147)	19% (77)	6% (26)	2% (9)	0% (0)	100% (406)
Beskrivelserne i fælles mål af kompetencemål samt de vejledende videns- og færdighedsmål	12% (47)	20% (80)	28% (114)	29% (119)	11% (46)	0% (0)	100% (406)
Den vejledende læseplan	7% (30)	17% (68)	28% (113)	35% (142)	13% (53)	0% (0)	100% (406)
Andet	2% (9)	1% (5)	0% (1)	0% (1)	0% (0)	96% (390)	100% (406)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Tabel 3.6 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet: ”I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende?”. I det følgende oplystes andelen, der har svaret, at de bruger værktøjet i høj grad (enten ”I meget høj grad” eller ”I høj grad”):

- 70 % bruger i høj grad ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”
- 31 % bruger i høj grad ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 23 % bruger i høj grad ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”
- 20 % bruger i høj grad ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle

- Andelen, der har angivet at de i høj grad bruger ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”, er statistisk signifikant forskellig mellem bedømmere i dansk (17 %) og bedømmere i matematik (32 %), såvel som mellem bedømmere i engelsk (13 %) og bedømmere i matematik (32 %).
- Andelen, der har angivet at de i høj grad bruger ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”, er statistisk signifikant forskellig mellem bedømmere i dansk (26 %) og bedømmere i matematik (40 %).

TABEL 3.6

I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	33 % (133)	37 % (152)	19 % (77)	7 % (28)	3 % (12)	1 % (4)	100 % (406)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	8 % (32)	15 % (62)	25 % (101)	18 % (74)	22 % (91)	11 % (46)	100 % (406)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	4 % (17)	16 % (64)	22 % (90)	23 % (92)	32 % (129)	3 % (14)	100 % (406)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	10 % (41)	21 % (85)	34 % (137)	19 % (78)	14 % (57)	2 % (8)	100 % (406)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Tabel 3.7 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved mundtlige prøver. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 69 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 62 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 51 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 33 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 22 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen der har svaret "Meget enig" eller "Enig" til hvert udsagn er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.7

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	21 % (87)	48 % (195)	22 % (89)	4 % (16)	1 % (4)	4 % (15)	100 % (406)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	9 % (38)	24 % (99)	43 % (175)	14 % (57)	2 % (8)	7 % (29)	100 % (406)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	4 % (17)	18 % (72)	39 % (157)	26 % (107)	5 % (22)	8 % (31)	100 % (406)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	15 % (60)	47 % (191)	22 % (89)	7 % (30)	3 % (11)	6 % (25)	100 % (406)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse	11 % (44)	40 % (162)	33 % (132)	5 % (20)	3 % (12)	9 % (36)	100 % (406)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Gennemgang af resultater vedr. standpunktskarakter/årskarakterer

Tabel 3.8 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer. I det følgende oplystes andelen, der har svaret, at elementet indgår i høj grad i karakterafgivelsen (enten "I meget høj grad" eller "I høj grad"):

- For 99 % indgår i høj grad "Vurdering af elevernes faglige kompetence"
- For 59 % indgår i høj grad "Vurdering af elevernes deltagelse"
- For 7 % indgår i høj grad "Vurdering af elevernes opførsel".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der angiver at "Vurdering af elevernes deltagelse" indgår i høj grad, er statistisk signifikant forskellig mellem bedømmere i dansk (65 %) og bedømmere i matematik (53 %).

TABEL 3.8

I hvilken grad indgår følgende elementer når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Total
Vurdering af elevernes faglige kompetence	76%	23%	1%	0%	0%	100%
	(268)	(80)	(4)	(0)	(0)	(352)
Vurdering af elevernes deltagelse	23%	36%	25%	11%	4%	100%
	(82)	(126)	(89)	(40)	(15)	(352)
Vurdering af elevernes opførsel	2%	5%	16%	32%	45%	100%
	(7)	(17)	(55)	(113)	(160)	(352)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 3.9 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer i hhv. dansk, engelsk og matematik. Det følgende oplystes hvilket bedømmelsesværktøj, bedømmere gør brug af i høj grad (enten "I høj grad" eller "I meget høj grad"):

- 68 % af bedømmere af standpunktskarakter/årskarakterer svarer at de "I meget høj grad" eller "I høj grad" gør brug af "De krav der i sidste ende stilles til den afsluttende prøve i faget"
- 65 % af bedømmere af standpunktskarakter/årskarakterer svarer at de "I meget høj grad" eller "I høj grad" gør brug af "Den generelle beskrivelse af karaktertrin i 7-trins-skalaen"
- 47 % af bedømmere af standpunktskarakter/årskarakterer svarer at de "I meget høj grad" eller "I høj grad" gør brug af "Beskrivelserne af fælles mål/faglige mål i faget"
- 11 % af bedømmere af standpunktskarakter/årskarakterer svarer at de "I meget høj grad" eller "I høj grad" gør brug af "Lokalt udarbejdet hjælpemiddel (fx bedømmelsesskema)".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret "I meget høj grad" eller "I høj grad" til hvert bedømmelsesværktøj, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.9

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle beskrivelse af karaktertrin i 7-trins-skalaen	20% (69)	45% (157)	26% (91)	7% (26)	3% (9)	0% (0)	100% (352)
Beskrivelserne af fælles mål/faglige mål i faget	15% (54)	32% (114)	33% (117)	16% (56)	3% (11)	0% (0)	100% (352)
De krav der i sidste ende stilles til den afsluttende prøve i faget	25% (88)	43% (152)	25% (89)	5% (16)	2% (7)	0% (0)	100% (352)
Lokalt udarbejdet hjælpemiddel (fx bedømmelsesskema)	3% (10)	8% (28)	18% (63)	23% (82)	48% (169)	0% (0)	100% (352)
Andet	1% (4)	2% (6)	0% (0)	0% (0)	0% (0)	97% (342)	100% (352)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 3.10 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet: "I hvilken grad baserer du afgivelse af standpunktskarakterer/årskarakterer på følgende". Det følgende oplystes andelen, der har svaret at de baserer sin karaktergivning på værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 52 % bruger i høj grad: "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)"
- 48 % bruger i høj grad: "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle"
- 21 % bruger i høj grad: "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)"
- 18 % bruger i høj grad: "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)" er statistisk signifikant forskellig mellem bedømmere i dansk (59 %) og bedømmere i matematik (43 %).

- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”, er statistisk signifikant forskellig mellem bedømmere i engelsk (30 %) og bedømmere i matematik (14 %).
- Andelen, der i høj grad bruger ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”, er statistisk signifikant forskellig mellem bedømmere i dansk (38 %) og bedømmere i matematik (55 %).

TABEL 3.10

I hvilken grad baserer du afgivelse af standpunktskarakterer/årskarakterer på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	16 % (57)	36 % (127)	28 % (100)	12 % (42)	5 % (18)	2 % (8)	100 % (352)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	5 % (16)	13 % (47)	22 % (79)	22 % (78)	26 % (93)	11 % (39)	100 % (352)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	3 % (12)	18 % (63)	30 % (107)	17 % (60)	27 % (94)	5 % (16)	100 % (352)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	14 % (49)	34 % (119)	26 % (90)	17 % (60)	7 % (25)	3 % (9)	100 % (352)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Tabel 3.11 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskarakterer. Det følgende oplystes andelen, der har svaret at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 52 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen”
- 45 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis”
- 43 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse”
- 26 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes”
- 20 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad er enig i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen", er statistisk signifikant forskellig mellem bedømmere i engelsk (63 %) og bedømmere i matematik (43 %).
- Andelen, der i høj grad er enig i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere i dansk (39 %) og bedømmere i engelsk (43 %), såvel som mellem bedømmere i engelsk (43 %) og matematik (39 %).

TABEL 3.11

Hvor enig eller uenig er du i følgende udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskarakterer?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	14 % (49)	38 % (132)	32 % (113)	7 % (23)	2 % (6)	8 % (29)	100 % (352)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	6 % (21)	20 % (71)	47 % (166)	13 % (44)	3 % (9)	12 % (41)	100 % (352)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	3 % (11)	17 % (60)	43 % (153)	19 % (68)	4 % (13)	13 % (47)	100 % (352)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	9 % (32)	36 % (125)	35 % (123)	7 % (25)	2 % (6)	12 % (41)	100 % (352)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål og min bedømmelse	7 % (25)	36 % (125)	37 % (129)	6 % (20)	2 % (7)	13 % (46)	100 % (352)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Tabel 3.12 viser, hvad bedømmerne på grundskoleområdet svarer på spørgsmålet "I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?". Det følgende oplistes andelen, der har svaret at de anvender kilden i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 80 % anvender i høj grad "Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)"

- 75 % anvender i høj grad "Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)"
- 50 % anvender i høj grad "Resultater fra standardiserede testredskaber (fx nationale test mv.)"
- 38 % anvender i høj grad "Resultater fra lokalt udviklede test".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad anvender "Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)", er statistisk signifikant forskellig mellem bedømmere i dansk (60 %) og bedømmere i engelsk (27 %), såvel som mellem bedømmere i engelsk (27 %) og bedømmere i matematik (70 %).
- Andelen, der i høj grad anvender "Resultater fra standardiserede testredskaber (fx nationale test mv.)", er statistisk signifikant forskellig mellem bedømmere i engelsk (33 %) og bedømmere i matematik (53 %).
- Andelen, der i høj grad anvender "Resultater fra lokalt udviklede test", er statistisk signifikant forskellig mellem bedømmere i dansk (28 %) og bedømmere i matematik (47 %), såvel som mellem bedømmere i engelsk (25 %) og bedømmere i matematik (47 %).

TABEL 3.12

I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke rele- vant	Total
Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)	37 % (131)	43 % (152)	16 % (56)	3 % (10)	1 % (3)	0 % (0)	100 % (352)
Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)	36 % (125)	39 % (138)	19 % (66)	5 % (19)	1 % (4)	0 % (0)	100 % (352)
Resultater fra standardiserede testredskaber (fx nationale test mv.)	16 % (55)	34 % (120)	27 % (96)	17 % (59)	6 % (22)	0 % (0)	100 % (352)
Resultater fra lokalt udviklede test	11 % (38)	27 % (96)	28 % (97)	15 % (53)	19 % (68)	0 % (0)	100 % (352)
Andet	2 % (6)	2 % (7)	1 % (3)	0 % (0)	0 % (1)	95 % (335)	100 % (352)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Gennemgang af resultater vedr. sammenhængen mellem læringsmål og bedømmelse

Tabel 3.13 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om sammenhængen mellem læringsmål og bedømmelse. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 71 % er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig"
- 38 % er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring"
- 27 % er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret "Meget enig" eller "Enig" til hvert udsagn, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.13

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring	6 % (43)	32 % (211)	33 % (221)	18 % (123)	9 % (57)	1 % (6)	1 % (4)	100 % (665)
Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne	5 % (33)	22 % (148)	23 % (155)	33 % (222)	10 % (64)	3 % (18)	4 % (25)	100 % (665)
Mine medbedømmere lægger typisk det faglige niveau samme sted som mig	8 % (52)	63 % (422)	19 % (129)	6 % (40)	1 % (5)	1 % (8)	1 % (9)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

3.3 Survey blandt bedømmere på de gymnasiale uddannelser

3.3.1 Sammenfatning

Surveyen blandt bedømmere på de gymnasiale uddannelser viser overordnet set, at:

- Bedømmerne i varierende grad anvender bedømmelsesværktøjerne, som er til rådighed ved mundtlige og skriftlige prøver

- Omkring halvdelen af bedømmerne er meget enig eller enige i, at bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis og at de giver et solidt udgangspunkt for bedømmelsen
- Alle bedømmere vurderer elevernes faglige kompetence ved afgivelse af standpunkts/årskarakter, og omkring halvdelen vurderer også elevernes deltagelse
- Størstedelen af bedømmerne er enige i, at deres medbedømmere typisk lægger det faglige niveau samme sted som dem, men under halvdelen at bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på elevernes læring.

3.3.2 Gennemgang af resultater vedr. skriftlige prøver

Tabel 3.14 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved skriftlige prøver. Det følgende oplystes andelen, der har svaret, at de bruger værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 72 % bruger i høj grad "Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier"
- 60 % bruger i høj grad "Vejledning til censorer i faget"
- 54 % bruger i høj grad "Egne retteark til den skriftlige eksamen"
- 53 % bruger i høj grad "Vejledning til fagets læreplan"
- 46 % bruger i høj grad "Andre materialer fra Undervisningsministeriet, herunder fagkonsulenterne"
- 44 % bruger i høj grad "Formelle retteark til den skriftlige eksamen"
- 38 % bruger i høj grad "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 28 % bruger i høj grad "Standardoversættelsesskala fra point til karakter til den skriftlige eksamen" (findes kun i matematik)
- 24 % bruger i høj grad "Karakterbekendtgørelsen - bekendtgørelse om karakterskala og anden bedømmelse"
- 18 % bruger i høj grad "Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelse"
- 18 % bruger i høj grad "Mindstekrav til den skriftlige eksamen" (findes kun i matematik)
- 10 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet internt på skolen"
- 9 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet af faglige foreninger"
- 3 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet på andre skoler".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger "Den generelle karakterbeskrivelse af 7-trins-skalaen", er statistisk signifikant forskellig mellem bedømmere i dansk (41 %) og bedømmere i matematik (25 %), såvel som mellem bedømmere i engelsk (45 %) og bedømmere i matematik (25 %).

- Andelen, der i høj grad bruger ”Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelser”, er statistisk signifikant forskellig mellem bedømmere i dansk (22 %) og bedømmere i matematik (14 %).
- Andelen, der i høj grad bruger ”Vejledning til fagets læreplan”, er statistisk signifikant forskellig mellem bedømmere i dansk (55 %) og bedømmere i matematik (43 %), såvel som mellem bedømmere i engelsk (57 %) og bedømmere i matematik (43 %).
- Andelen, der i høj grad bruger ”Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier”, er statistisk signifikant forskellig mellem bedømmere i dansk (75 %) og bedømmere i matematik (61 %), såvel som mellem bedømmere i engelsk (79 %) og bedømmere i matematik (61 %).
- Andelen, der i høj grad bruger ”Andre materialer fra Undervisningsministeriet, herunder fagkonsulenterne”, er statistisk signifikant forskellig mellem bedømmere i dansk (38 %) og bedømmere i matematik (55 %), såvel som mellem bedømmere i engelsk (45 %) og bedømmere i matematik (55 %).
- Andelen, der i høj grad bruger ”Materialer til brug i bedømmelsen udarbejdet internt på skolen”, er statistisk signifikant forskellig mellem bedømmere i dansk (15 %) og bedømmere i engelsk (8 %), såvel som mellem bedømmere i dansk (15 %) og bedømmere i matematik (6 %).
- Andelen, der i høj grad bruger ”Formelle retteark til den skriftlige eksamen”, er statistisk signifikant forskellig mellem bedømmere i dansk (47 %) og bedømmere i engelsk (60 %), såvel som mellem bedømmere i dansk (47 %) og bedømmere i matematik (19 %), såvel som mellem bedømmere i engelsk (60 %) og bedømmere i matematik (19 %).
- Andelen, der i høj grad bruger ”Egne retteark til den skriftlige eksamen”, er statistisk signifikant forskellig mellem bedømmere i dansk (50 %) og bedømmere i matematik (73 %), såvel som mellem bedømmere i engelsk (53 %) og bedømmere i matematik (73 %).

3.14

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver karakter ved en skriftlig prøve?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	9%	29%	36%	18%	8%	0%	100%
	(62)	(212)	(261)	(130)	(56)	(0)	(721)
Karakterbekendtgørelsen - bekendtgørelse om karakterskala og anden bedømmelse	5%	19%	29%	27%	19%	0%	100%
	(39)	(139)	(207)	(197)	(139)	(0)	(721)
Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelser	3%	15%	26%	29%	26%	0%	100%
	(23)	(110)	(188)	(211)	(189)	(0)	(721)
Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier	33%	39%	18%	6%	3%	0%	100%
	(237)	(284)	(131)	(44)	(25)	(0)	(721)

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke rele- vant	Total
Vejledning til fagets lære- plan	20% (144)	33% (236)	26% (188)	13% (94)	8% (59)	0% (0)	100% (721)
Vejledning til censorer i fa- get	26% (185)	34% (243)	21% (154)	10% (73)	9% (66)	0% (0)	100% (721)
Andre materialer fra Un- dervisningsministeriet, herunder fagkonsulen- terne	20% (141)	26% (186)	29% (209)	16% (112)	10% (73)	0% (0)	100% (721)
Materialer til brug i bedøm- melsen udarbejdet internt på skolen	3% (20)	7% (53)	14% (101)	28% (200)	48% (347)	0% (0)	100% (721)
Materialer til brug i bedøm- melsen udarbejdet på an- dre skoler	1% (4)	2% (12)	8% (55)	27% (192)	64% (458)	0% (0)	100% (721)
Materialer til brug i bedøm- melsen udarbejdet af fag- lige foreninger	3% (19)	6% (46)	14% (101)	25% (177)	52% (378)	0% (0)	100% (721)
Formelle retteark til den skriftlige eksamen	19% (137)	25% (180)	18% (129)	13% (93)	25% (182)	0% (0)	100% (721)
Egne retteark til den skrift- lige eksamen	24% (171)	30% (214)	17% (126)	11% (79)	18% (131)	0% (0)	100% (721)
Mindstekrav til den skrift- lige eksamen	4% (27)	14% (100)	23% (168)	21% (148)	39% (278)	0% (0)	100% (721)
Standardoversættelses- skala fra point til karakter til den skriftlige eksamen	20% (145)	8% (56)	5% (35)	11% (76)	57% (409)	0% (0)	100% (721)
Andet	4% (26)	1% (7)	0% (0)	0% (0)	0% (0)	95% (688)	100% (721)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Tabel 3.15 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet: "I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende?". Nedenfor oplystes andelen, der har svaret, at de baserer deres bedømmelse på værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 67 % bruger i høj grad "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)"
- 47 % bruger i høj grad "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle"
- 34 % bruger i høj grad "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)"

- 15 % bruger i høj grad ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”, er statistisk signifikant forskellig mellem bedømmere i engelsk (19 %) og bedømmere i matematik (12 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hhx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”, er statistisk signifikant forskellig mellem bedømmere på HTX (58 %) og bedømmere på HHX (32 %), såvel som mellem bedømmere på HTX (58 %) og bedømmere på HF (29 %), såvel som mellem bedømmere på HTX (58 %) og bedømmere på STX (32 %).

TABEL 3.15

I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	31 % (221)	36 % (258)	19 % (135)	7 % (54)	5 % (35)	2 % (18)	100 % (721)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	5 % (37)	10 % (70)	18 % (128)	24 % (174)	30 % (215)	13 % (97)	100 % (721)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	9 % (62)	25 % (177)	27 % (197)	16 % (114)	19 % (140)	4 % (31)	100 % (721)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	15 % (109)	32 % (232)	31 % (220)	13 % (96)	4 % (32)	4 % (32)	100 % (721)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

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.

Tabel 3.16 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved skriftlige prøver. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 61 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis”
- 57 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen”

- 57 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 35 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 24 % er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser"
- 11 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes", er statistisk signifikant forskellig mellem bedømmere i dansk (37 %) og bedømmere i matematik (17 %), såvel som mellem bedømmere i engelsk (45 %) og bedømmere i matematik (17 %).
- Andelen, der i høj grad er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles", er statistisk signifikant forskellig mellem bedømmere i dansk (15 %) og bedømmere i matematik (5 %), såvel som mellem bedømmere i engelsk (14 %) og bedømmere i matematik (5 %).
- Andelen, der i høj grad er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse", er statistisk signifikant forskellig mellem bedømmere i dansk (58 %) og bedømmere i matematik (43 %), såvel som mellem bedømmere i engelsk (64 %) og bedømmere i matematik (43 %).
- Andelen, der i høj grad er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser", er statistisk signifikant forskellig mellem bedømmere i dansk (20 %) og bedømmere i engelsk (13 %), såvel som mellem bedømmere i dansk (20 %) og bedømmere i matematik (11 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Andelen, der er "Meget enige" eller "Enige" i hvert af udsagnene, er ikke statistisk signifikant mellem de forskellige gymnasiale uddannelser.

TABEL 3.16

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	15 % (108)	42 % (303)	28 % (199)	5 % (35)	2 % (18)	8 % (58)	100 % (721)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	8 % (57)	27 % (192)	35 % (250)	14 % (98)	3 % (24)	14 % (100)	100 % (721)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	2 % (18)	9 % (68)	40 % (290)	26 % (185)	7 % (54)	15 % (106)	100 % (721)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	13 % (96)	48 % (344)	21 % (151)	5 % (38)	2 % (18)	10 % (74)	100 % (721)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse	12 % (84)	45 % (324)	23 % (169)	6 % (42)	1 % (10)	13 % (92)	100 % (721)
Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser	2 % (15)	12 % (88)	21 % (148)	25 % (181)	33 % (240)	7 % (49)	100 % (721)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

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.

3.3.3 Gennemgang af resultater vedr. mundtlige prøver

Tabel 3.17 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved mundtlige prøver. Det følgende oplystes andelen, der har svaret at de bruger værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 80 % bruger i høj grad "Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier"
- 54 % bruger i høj grad "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 54 % bruger i høj grad "Vejledning til fagets læreplan"
- 46 % bruger i høj grad "Undervisningsbeskrivelser opgivet til eksamen"
- 33 % bruger i høj grad "Vejledning til censorer i faget"
- 32 % bruger i høj grad "Karakterbekendtgørelsen - bekendtgørelse om karakterskala og anden bedømmelse"
- 23 % bruger i høj grad "Andre materialer udarbejdet af Undervisningsministeriet, herunder fagkonsulenterne"
- 20 % bruger i høj grad "Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelser"
- 12 % bruger i høj grad "Andre materialer udarbejdet af opgavekommissionen i faget"
- 7 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet internt på skolen"

- 7 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet af faglige foreninger"
- 1 % bruger i høj grad "Materialer til brug i bedømmelsen udarbejdet på andre skoler".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger "Den generelle karakterbeskrivelse af 7-trins-skalaen", er statistisk signifikant forskellig mellem bedømmere i dansk (41 %) og bedømmere i matematik (25 %), såvel som mellem bedømmere i engelsk (45 %) og bedømmere i matematik (25 %).
- Andelen, der i høj grad bruger "Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelser", er statistisk signifikant forskellig mellem bedømmere i dansk (22 %) og bedømmere i matematik (14 %).
- Andelen, der i høj grad bruger "Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier", er statistisk signifikant forskellig mellem bedømmere i dansk (75 %) og bedømmere i matematik (61 %), såvel som mellem bedømmere i engelsk (79 %) og bedømmere i matematik (61 %).
- Andelen, der i høj grad bruger "Vejledning til fagets læreplan", er statistisk signifikant forskellig mellem bedømmere i dansk (55 %) og bedømmere i matematik (43 %), såvel som mellem bedømmere i engelsk (57 %) og bedømmere i matematik (43 %).
- Andelen, der i høj grad bruger "Andre materialer fra Undervisningsministeriet, herunder fagkonsulenterne", er statistisk signifikant forskellig mellem bedømmere i dansk (38 %) og bedømmere i matematik (55 %), såvel som mellem bedømmere i engelsk (45 %) og bedømmere i matematik (55 %).
- Andelen, der i høj grad bruger "Materialer til brug i bedømmelsen udarbejdet internt på skolen", er statistisk signifikant forskellig mellem bedømmere i dansk (15 %) og bedømmere i engelsk (8 %), såvel som mellem bedømmere i dansk (15 %) og bedømmere i matematik (6 %).
- Andelen, der i høj grad bruger "Formelle retteark til den skriftlige eksamen", er statistisk signifikant forskellig mellem bedømmere i dansk (47 %) og bedømmere i engelsk (60 %), såvel som mellem bedømmere i dansk (47 %) og bedømmere i matematik (19 %), såvel som mellem bedømmere i engelsk (60 %) og bedømmere i matematik (19 %).
- Andelen, der i høj grad bruger "Egne retteark til den skriftlige eksamen", er statistisk signifikant forskellig mellem bedømmere i dansk (50 %) og bedømmere i matematik (73 %), såvel som mellem bedømmere i engelsk (43 %) og bedømmere i matematik (73 %).
- Andelen, der i høj grad bruger "Mindstekrav til den skriftlige eksamen", er statistisk signifikant forskellig mellem bedømmere i dansk (13 %) og bedømmere i engelsk (20 %), såvel som mellem bedømmere i dansk (43 %) og bedømmere i matematik (20 %).
- Andelen, der i høj grad bruger "Standardoversættelsesskala fra point til karakter til den skriftlige eksamen", er statistisk signifikant forskellig mellem bedømmere i dansk (2 %) og bedømmere i matematik (93 %), såvel som mellem bedømmere i engelsk (5 %) og bedømmere i matematik (93 %).

TABEL 3.17

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver karakter ved en skriftlig prøve?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	16%	38%	32%	11%	3%	0%	100%
	(137)	(337)	(284)	(97)	(27)	(0)	(882)
Karakterbekendtgørelsen - bekendtgørelse om karakterskala og anden bedømmelse	8%	24%	30%	23%	15%	0%	100%
	(72)	(213)	(262)	(205)	(130)	(0)	(882)
Almen eksamensbekendtgørelse - bekendtgørelse om prøver og eksamen i de almene og studieforberedende ungdoms- og voksenuddannelser	4%	16%	29%	27%	23%	0%	100%
	(38)	(145)	(257)	(241)	(201)	(0)	(882)
Undervisningsbeskrivelser opgivet til eksamen	12%	34%	31%	16%	7%	0%	100%
	(108)	(296)	(274)	(139)	(65)	(0)	(882)
Fagets læreplan med faglige mål, kernestof og bedømmelseskriterier	36%	44%	15%	4%	2%	0%	100%
	(314)	(388)	(128)	(35)	(17)	(0)	(882)
Vejledning til fagets læreplan	19%	35%	26%	12%	8%	0%	100%
	(168)	(308)	(232)	(107)	(67)	(0)	(882)
Vejledning til censorer i faget	10%	23%	27%	20%	20%	0%	100%
	(87)	(205)	(235)	(177)	(178)	(0)	(882)
Andre materialer udarbejdet af opgavekommissionen i faget	3%	9%	24%	26%	38%	0%	100%
	(26)	(79)	(210)	(229)	(338)	(0)	(882)
Andre materialer udarbejdet af Undervisningsministeriet, herunder fagkonsulenterne	6%	17%	29%	23%	25%	0%	100%
	(55)	(148)	(253)	(204)	(222)	(0)	(882)
Materialer til brug i bedømmelsen udarbejdet internt på skolen	1%	6%	16%	28%	50%	0%	100%
	(8)	(50)	(139)	(243)	(442)	(0)	(882)
Materialer til brug i bedømmelsen udarbejdet på andre skoler	0%	1%	10%	25%	64%	0%	100%
	(1)	(11)	(86)	(222)	(562)	(0)	(882)
Materialer til brug i bedømmelsen udarbejdet af faglige foreninger	2%	5%	14%	25%	54%	0%	100%
	(15)	(44)	(127)	(221)	(475)	(0)	(882)
Andet	2%	1%	0%	0%	0%	97%	100%
	(21)	(5)	(3)	(1)	(0)	(852)	(882)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Tabel 3.18 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet ”I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende?”. Det følgende oplystes andelen, der har svaret at de baserer sin bedømmelse på værktøjet i høj grad (enten ”I meget høj grad” eller ”I høj grad”):

- 60 % bruger i høj grad ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”
- 58 % bruger i høj grad ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 15 % bruger i høj grad ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”
- 13 % bruger i høj grad ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”, er statistisk signifikant forskellig mellem bedømmere i dansk (62 %) og bedømmere i matematik (52 %), såvel som mellem bedømmere i engelsk (68 %) og bedømmere fra matematik (52 %).
- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”, er statistisk signifikant forskellig mellem bedømmere i dansk (15 %) og bedømmere i matematik (7 %), såvel som mellem bedømmere i engelsk (17 %) og bedømmere fra matematik (7 %).
- Andelen, der i høj grad bruger ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”, er statistisk signifikant forskellig mellem bedømmere i dansk (18 %) og bedømmere i matematik (9 %), såvel som mellem bedømmere i engelsk (16 %) og bedømmere fra matematik (9 %).
- Andelen, der i høj grad bruger ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”, er statistisk signifikant forskellig mellem bedømmere i engelsk (53 %) og bedømmere i matematik (65 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Besvarelsene på hvert bedømmelsesværktøj er ikke statistisk signifikant forskellig mellem de forskellige gymnasiale uddannelser.

TABEL 3.18

I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	23 % (203)	37 % (330)	21 % (182)	9 % (77)	7 % (58)	4 % (32)	100 % (882)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	3 % (24)	10 % (85)	19 % (170)	24 % (212)	32 % (282)	12 % (109)	100 % (882)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	3 % (27)	12 % (105)	24 % (214)	23 % (200)	32 % (281)	6 % (55)	100 % (882)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	22 % (193)	36 % (318)	24 % (216)	10 % (89)	4 % (32)	4 % (34)	100 % (882)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Tabel 3.19 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved mundtlige prøver. Det følgende oplystes andelen, der har svaret at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 49 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 49 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 49 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 33 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 9 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er meget enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen", er statistisk signifikant forskellig mellem bedømmere fra engelsk (55 %) og bedømmere fra matematik (45 %).
- Andelen, der er meget enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes", er statistisk signifikant forskellig mellem bedømmere i dansk (37 %) og bedømmere i matematik (20 %), såvel som bedømmere i engelsk (42 %) og bedømmere i matematik (20 %).

- Andelen, der er meget enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis”, er statistisk signifikant forskellig mellem bedømmere i dansk (43 %) og bedømmere i engelsk (59 %), såvel som bedømmere i engelsk (59 %) og bedømmere i matematik (43 %).
- Andelen, der er meget enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse”, er statistisk signifikant forskellig mellem bedømmere i dansk (46 %) og bedømmere i engelsk (60 %), såvel som bedømmere i engelsk (60 %) og bedømmere i matematik (40 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er meget enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes”, er statistisk signifikant forskellig mellem bedømmere på stx (32 %) og bedømmere på hf (20 %), såvel som mellem bedømmere på htx (42 %) og bedømmere på hf (20 %).

TABEL 3.19

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	10 % (85)	39 % (344)	32 % (282)	6 % (54)	2 % (22)	11 % (95)	100 % (882)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	6 % (55)	27 % (236)	38 % (334)	11 % (95)	3 % (26)	15 % (136)	100 % (882)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	2 % (15)	7 % (59)	41 % (359)	27 % (239)	6 % (49)	18 % (161)	100 % (882)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	8 % (74)	41 % (360)	28 % (243)	7 % (64)	2 % (15)	14 % (126)	100 % (882)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse	9 % (80)	40 % (355)	29 % (253)	5 % (47)	1 % (11)	15 % (136)	100 % (882)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Gennemgang af resultater vedr. standpunktskarakter/årskarakterer

Tabel 3.20 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer. Det følgende oplystes andelen, der har svaret, at elementet indgår i høj grad i karakterafgivelsen (enten "I meget høj grad" eller "I høj grad"):

- For 100 % indgår i høj grad "Vurdering af elevernes faglige kompetence"
- For 49 % indgår i høj grad "Vurdering af elevernes deltagelse"
- For 3 % indgår i høj grad "Vurdering af elevernes opførsel".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har angivet at "Vurdering af elevernes deltagelse" indgår i høj grad, er statistisk signifikant forskellig mellem bedømmere i dansk (55 %) og bedømmere i matematik (36 %), såvel som mellem bedømmere i engelsk (54 %) og bedømmere i matematik (36 %).

TABEL 3.20

I hvilken grad indgår følgende elementer når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Total
Vurdering af elevernes faglige kompetence	84%	16%	0%	0%	0%	100%
	(482)	(92)	(2)	(0)	(0)	(576)
Vurdering af elevernes deltagelse	11%	38%	34%	13%	3%	100%
	(66)	(221)	(194)	(75)	(20)	(576)
Vurdering af elevernes opførsel	1%	2%	14%	35%	49%	100%
	(4)	(10)	(78)	(199)	(285)	(576)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Tabel 3.21 viser, hvordan bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer. Det følgende oplystes bedømmelsesværktøjerne, som bedømmerne gør brug af i høj grad (enten "I høj grad" eller "I meget høj grad"):

- 72 % af bedømmerne gør i høj grad brug af "Beskrivelserne af fælles mål/faglige mål i faget"
- 64 % af bedømmerne gør i høj grad brug af "De krav der i sidste ende stilles til den afsluttende prøve i faget"
- 52 % af bedømmerne gør i høj grad brug af "Den generelle beskrivelse af karaktertrin i 7-trins-skalaen"
- 6 % af bedømmerne gør i høj grad brug af "Lokalt udarbejdet hjælpemiddel (fx bedømmelses-skema)".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad gør brug af ”Den generelle beskrivelse af karaktertrin i 7-trins-skalaen”, er statistisk signifikant forskellig mellem bedømmere fra engelsk (57 %) og bedømmere fra matematik (44 %).
- Andelen, der i høj grad gør brug af ”Lokalt udarbejdet hjælpemiddel (fx bedømmelsesskema)”, er statistisk signifikant forskellig mellem bedømmere fra engelsk (8 %) og bedømmere fra matematik (3 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Besvarelsene er ikke statistisk signifikant forskellig mellem de forskellige gymnasiale uddannelser.

TABEL 3.21

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle beskrivelse af karaktertrin i 7-trins-skalaen	16% (95)	36% (206)	30% (172)	13% (72)	5% (31)	0% (0)	100% (576)
Beskrivelserne af fælles mål/faglige mål i faget	27% (156)	45% (259)	22% (126)	4% (24)	2% (11)	0% (0)	100% (576)
De krav der i sidste ende stilles til den afsluttende prøve i faget	19% (111)	45% (261)	27% (156)	7% (39)	2% (9)	0% (0)	100% (576)
Lokalt udarbejdet hjælpemiddel (fx bedømmelsesskema)	2% (9)	4% (22)	13% (74)	23% (135)	58% (336)	0% (0)	100% (576)
Andet	2% (11)	1% (4)	1% (7)	0% (1)	0% (0)	96% (553)	100% (576)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Tabel 3.22 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet ”I hvilken grad baserer du afgivelse af standpunktskarakterer/årskarakterer på følgende”. Nedenfor oplistes andelen, der har svaret, at de baserer deres karaktergivning på værktøjet i høj grad (enten ”I meget høj grad” eller ”I høj grad”):

- 59 % bruger i høj grad ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 51 % bruger i høj grad ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”

- 18 % bruger i høj grad "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)"
- 12 % bruger i høj grad "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)", er statistisk signifikant forskellig mellem bedømmere i dansk (55 %) og bedømmere i matematik (38 %), såvel som mellem bedømmere i engelsk (56 %) og bedømmere i matematik (38 %).
- Andelen, der i høj grad bruger "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed", er statistisk signifikant forskellig mellem bedømmere i dansk (13 %) og bedømmere i matematik (5 %), såvel som mellem bedømmere i engelsk (15 %) og bedømmere i matematik (5 %).
- Andelen, der i høj grad bruger "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle", er statistisk signifikant forskellig mellem bedømmere i engelsk (54 %) og bedømmere i matematik (68 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Andelen, der "I meget høj grad" eller "I høj grad" benytter sig af de forskellige bedømmelsesværktøjer, er ikke statistisk signifikant forskellig mellem de forskellige gymnasiale uddannelser.

TABEL 3.22

I hvilken grad baserer du afgivelse af standpunktskarakterer/årskaracterer på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	17 % (100)	34 % (196)	26 % (149)	11 % (63)	9 % (50)	3 % (18)	100 % (576)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	3 % (15)	9 % (50)	18 % (106)	23 % (131)	34 % (198)	13 % (76)	100 % (576)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	4 % (24)	14 % (80)	26 % (147)	22 % (125)	29 % (168)	6 % (32)	100 % (576)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	21 % (119)	38 % (217)	27 % (155)	9 % (54)	2 % (14)	3 % (17)	100 % (576)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Tabel 3.23 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskaracterer. I

det følgende oplystes andelen, der har svaret at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 49 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen”
- 48 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis”
- 49 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse”
- 35 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes”
- 9 % er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes”, er statistisk signifikant forskellig mellem bedømmere i dansk (39 %) og bedømmere i matematik (19 %), såvel som mellem bedømmere i engelsk (42 %) og bedømmere i matematik (19 %).
- Andelen, der er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis”, er statistisk signifikant forskellig mellem bedømmere i dansk (44 %) og bedømmere i engelsk (56 %), såvel som mellem bedømmere i engelsk (56 %) og bedømmere i matematik (37 %).
- Andelen, der er enige i udsagnet ”Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål og min bedømmelse”, er statistisk signifikant forskellig mellem bedømmere i dansk (50 %) og bedømmere i matematik (37 %), såvel som mellem bedømmere i engelsk (55 %) og bedømmere i matematik (37 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Andelen, der har svaret ”Meget enig” eller ”enig” til de forskellige udsagn, er ikke statistisk signifikant forskellig mellem de forskellige gymnasiale uddannelser.

TABEL 3.23

Hvor enig eller uenig er du i følgende udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskarakterer?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	10 % (58)	39 % (223)	31 % (180)	7 % (39)	2 % (13)	11 % (63)	100 % (576)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	6 % (35)	29 % (167)	35 % (202)	11 % (66)	2 % (12)	16 % (94)	100 % (576)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	1 % (5)	8 % (47)	42 % (243)	26 % (151)	5 % (30)	17 % (100)	100 % (576)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	8 % (48)	40 % (228)	30 % (170)	6 % (37)	1 % (8)	15 % (85)	100 % (576)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål og min bedømmelse	8 % (48)	41 % (237)	30 % (173)	5 % (26)	2 % (9)	14 % (83)	100 % (576)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen

Tabel 3.24 viser, hvad bedømmerne på de gymnasiale uddannelser svarer på spørgsmålet "I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?". Nedenfor oplystes andelen, der har svaret at de anvender kilden i høj grad (enten "I meget høj grad" eller "I høj grad"):

- 85 % anvender i høj grad "Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)"
- 63 % anvender i høj grad "Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)"
- 36 % anvender i høj grad "Resultater fra lokalt udviklede test"
- 15 % anvender i høj grad "Resultater fra standardiserede testredskaber (fx nationale test mv.)"

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad anvender "Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)", er statistisk signifikant forskellig mellem bedømmere i engelsk (89 %) og bedømmere i matematik (81 %).

- Andelen, der i høj grad anvender "Resultater fra standardiserede testredskaber (fx nationale test mv.)", er statistisk signifikant forskellig mellem bedømmere i dansk (5 %) og bedømmere i engelsk (19 %), såvel som mellem bedømmere i dansk (5 %) og bedømmere i matematik (24 %).
- Andelen, der i høj grad anvender "Resultater fra lokalt udviklede test", er statistisk signifikant forskellig mellem bedømmere i dansk (16 %) og bedømmere i engelsk (31 %), såvel som mellem bedømmere i dansk (16 %) og bedømmere i matematik (67 %), såvel som mellem bedømmere i engelsk (31 %) og bedømmere i matematik (67 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad anvender "Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)", er statistisk signifikant forskellig mellem bedømmere på stx (93 %) og bedømmere på hf (67 %), såvel som mellem bedømmere på hf (67 %) og bedømmere på htx (91 %), såvel som mellem bedømmere på htx (91 %) og bedømmere på htx (66 %), såvel som mellem bedømmere på stx (93 %) og bedømmere på htx (66 %).
- Andelen, der i høj grad anvender "Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)", er statistisk signifikant forskellig mellem bedømmere på stx (64 %) og bedømmere på hf (39 %), såvel som mellem bedømmere på hf (39 %) og bedømmere på htx (71 %).

TABEL 3.24

I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)	40 % (230)	45 % (257)	14 % (78)	1 % (6)	1 % (5)	0 % (0)	100 % (576)
Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)	25 % (143)	38 % (220)	23 % (130)	11 % (65)	3 % (18)	0 % (0)	100 % (576)
Resultater fra standardiserede testredskaber (fx nationale test mv.)	5 % (30)	10 % (59)	16 % (93)	26 % (147)	43 % (247)	0 % (0)	100 % (576)
Resultater fra lokalt udviklede test	12 % (70)	24 % (136)	26 % (152)	20 % (116)	18 % (102)	0 % (0)	100 % (576)
Andet	1 % (7)	1 % (7)	2 % (10)	0 % (0)	0 % (0)	96 % (552)	100 % (576)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Gennemgang af resultater vedr. sammenhængen mellem læringsmål og bedømmelse

Tabel 3.25 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om sammenhængen mellem læringsmål og bedømmelse. Det følgende oplystes andelen, der har svaret at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 80 % er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig".
- 46 % er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring".
- 10 % er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring" er statistisk signifikant forskellig mellem bedømmere fra engelsk (50 %) og bedømmere fra matematik (41 %).
- Andelen, der er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig", er statistisk signifikant forskellig mellem bedømmere fra dansk (75 %) og bedømmere fra matematik (86 %), såvel som mellem bedømmere i engelsk (77 %) og bedømmere i matematik (86 %).
- Andelen, der er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring", er statistisk signifikant forskellig mellem bedømmere fra hf (37 %) og bedømmere fra hhx (52 %), ifølge en almindelig z-test.

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hhx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring", er statistisk signifikant forskellig mellem bedømmere fra HF (37 %) og bedømmere fra HHX (52 %).
- Andelen, der er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne", er statistisk signifikant forskellig mellem bedømmere fra stx (9 %) og bedømmere fra hhx (17 %), såvel som mellem bedømmere på hf (7 %) og bedømmere på hhx (17 %).
- Andelen, der er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig", er statistisk signifikant forskellig mellem bedømmere fra hf (85 %) og bedømmere fra hhx (74 %).

TABEL 3.25

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring	8 % (95)	38 % (449)	30 % (360)	14 % (166)	5 % (54)	5 % (55)	1 % (8)	100 % (1187)
Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne	1 % (13)	9 % (104)	21 % (253)	43 % (512)	22 % (264)	1 % (15)	2 % (26)	100 % (1187)
Mine medbedømmere lægger typisk det faglige niveau samme sted som mig	14 % (162)	66 % (778)	16 % (184)	3 % (41)	0 % (4)	1 % (12)	1 % (6)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

3.4 Survey blandt bedømmere på erhvervsuddannelserne

3.4.1 Sammenfatning

Surveyen blandt bedømmere på erhvervsuddannelserne viser overordnet set, at:

- Bedømmerne i varierende grad anvender bedømmelsesværktøjerne, som er til rådighed ved mundtlige og skriftlige prøver
- Ca. halvdelen af bedømmerne er meget enig eller enige i, at bedømmelsesværktøjerne, der stilles til rådighed centralt eller af uddannelsen, er med til at sikre en ensartet bedømmelsespraksis, og at de giver et solidt udgangspunkt for bedømmelsen
- Alle bedømmere vurderer elevernes faglige kompetence ved afgivelse af standpunkts/årskarakter, og mere end halvdelen vurderer også elevernes deltagelse
- Størstedelen af bedømmerne er enige i, at deres medbedømmere typisk lægger det faglige niveau samme sted som dem.

3.4.2 Gennemgang af resultater vedr. skriftlige prøver

Tabel 3.26 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved skriftlige prøver. I det følgende oplystes andelen, der har svaret at de bruger værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"), listet fra højeste andel til laveste andel:

- 66 % bruger i høj grad "Bedømmelseskriterier"
- 61 % bruger i høj grad "Målbeskrivelser som redegør for målene i et fag på de forskellige niveauer"
- 58 % bruger i høj grad "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 30 % bruger i høj grad "Den lokale undervisningsplan".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Besvarelsene på hvert udsagn er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.26

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver karakter ved en skriftlig prøve?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	21%	37%	26%	9%	7%	0%	100%
	(12)	(21)	(15)	(5)	(4)	(0)	(57)
Den lokale undervisningsplan	2%	28%	39%	21%	11%	0%	100%
	(1)	(16)	(22)	(12)	(6)	(0)	(57)
Målbeskrivelser som redegør for målene i et fag på de forskellige niveauer	14%	47%	18%	14%	7%	0%	100%
	(8)	(27)	(10)	(8)	(4)	(0)	(57)
Bedømmelseskriterier	12%	54%	26%	4%	4%	0%	100%
	(7)	(31)	(15)	(2)	(2)	(0)	(57)
Andet	9%	5%	0%	0%	0%	86%	100%
	(5)	(3)	(0)	(0)	(0)	(49)	(57)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne

Tabel 3.27 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet "I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende?". I det følgende oplystes andelen, der har svaret at de baserer sin bedømmelse på værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"), listet fra højeste andel til laveste andel:

- 64 % baserer i høj grad bedømmelsen på "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)"
- 46 % baserer i høj grad bedømmelsen på "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed"
- 46 % baserer i høj grad bedømmelsen på "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)"
- 27 % baserer i høj grad bedømmelsen på "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Besvarelsene på hvert udsagn er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.27

I hvilken grad baserer du bedømmelser ved skriftlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	25 % (14)	39 % (22)	18 % (10)	4 % (2)	9 % (5)	7 % (4)	100 % (57)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	9 % (5)	37 % (21)	23 % (13)	11 % (6)	14 % (8)	7 % (4)	100 % (57)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	4 % (2)	42 % (24)	19 % (11)	14 % (8)	14 % (8)	7 % (4)	100 % (57)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	4 % (2)	23 % (13)	33 % (19)	26 % (15)	4 % (2)	11 % (6)	100 % (57)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

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.

Tabel 3.28 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved skriftlige prøver. I det følgende oplyses andelen, der har svaret at de er enige (enten "Meget enig" eller "Enig") i udsagnene, listet fra højeste andel til laveste andel:

- 59 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 55 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 51 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 44 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 22 % er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser"
- 18 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplyses alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere i dansk (33 %) og bedømmere i matematik (72 %).

TABEL 3.28

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	5 % (3)	54 % (31)	25 % (14)	5 % (3)	2 % (1)	9 % (5)	100 % (57)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	9 % (5)	35 % (20)	42 % (24)	4 % (2)	2 % (1)	9 % (5)	100 % (57)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	4 % (2)	14 % (8)	46 % (26)	19 % (11)	4 % (2)	14 % (8)	100 % (57)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	2 % (1)	53 % (30)	30 % (17)	11 % (6)	0 % (0)	5 % (3)	100 % (57)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbemærkninger og min bedømmelse	7 % (4)	44 % (25)	32 % (18)	7 % (4)	4 % (2)	7 % (4)	100 % (57)
Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser	4 % (2)	18 % (10)	23 % (13)	28 % (16)	25 % (14)	4 % (2)	100 % (57)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

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.

3.4.3 Gennemgang af resultater vedr. mundtlige prøver

Tabel 3.29 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om deres brug af bedømmelsesværktøjer ved mundtlige prøver. I det følgende oplyses andelen, der har svaret at de bruger værktøjet i høj grad (enten "I meget høj grad" eller "I høj grad"), listet fra højeste andel til laveste andel:

- 70 % bruger i høj grad "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 66 % bruger i høj grad "Målbemærkninger som redegør for målene i et fag på de forskellige niveauer"
- 60 % bruger i høj grad "Bedømmelseskriterier"
- 35 % bruger i høj grad "Den lokale undervisningsplan".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har angivet at de i høj grad bruger ”Den generelle karakterbeskrivelse af 7-trins-skalaen”, er statistisk signifikant forskellig mellem bedømmere i engelsk (75 %) og bedømmere i matematik (59 %).
- Andelen, der har angivet at de i høj grad bruger ”Bedømmelseskriterier”, er statistisk signifikant forskellig mellem bedømmere fra engelsk (70 %) og bedømmere fra matematik (50 %).

TABEL 3.29

I hvilken grad gør du brug af følgende bedømmelsesværktøjer når du afgiver karakter ved en mundtlig prøve?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	24%	46%	26%	3%	1%	0%	100%
	(69)	(133)	(74)	(10)	(2)	(0)	(288)
Den lokale undervisningsplan	9%	26%	32%	21%	11%	0%	100%
	(26)	(75)	(93)	(61)	(33)	(0)	(288)
Målbeskrivelser som redegør for målene i et fag på de forskellige niveauuer	22%	44%	24%	8%	2%	0%	100%
	(64)	(127)	(68)	(22)	(7)	(0)	(288)
Bedømmelseskriterier	21%	39%	33%	4%	3%	0%	100%
	(60)	(112)	(95)	(12)	(9)	(0)	(288)
Andet	3%	0%	2%	0%	0%	94%	100%
	(8)	(1)	(6)	(1)	(1)	(271)	(288)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Tabel 3.30 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet ”I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende?”. I det følgende oplistes andelen, der har svaret, at de i høj grad (enten ”I meget høj grad” eller ”I høj grad”) baserer deres bedømmelse på værktøjet:

- 64 % baserer i høj grad bedømmelsen på ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”
- 36 % baserer i høj grad bedømmelsen på ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”
- 35 % baserer i høj grad bedømmelsen på ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 22 % baserer i høj grad bedømmelsen på ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Besvarelsene på hvert bedømmelsesværktøj er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.30

I hvilken grad baserer du bedømmelser ved mundtlige prøver på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	21 % (60)	43 % (123)	25 % (71)	6 % (16)	4 % (12)	2 % (6)	100 % (288)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	9 % (26)	27 % (79)	30 % (86)	18 % (52)	11 % (31)	5 % (14)	100 % (288)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	5 % (13)	17 % (48)	34 % (98)	19 % (55)	23 % (66)	3 % (8)	100 % (288)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	8 % (22)	27 % (79)	32 % (93)	17 % (49)	10 % (29)	6 % (16)	100 % (288)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Tabel 3.31 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved mundtlige prøver. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 53 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 51 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 49 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 38 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 22 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Besvarelsene på hvert bedømmelsesværktøj er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.31

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	8 % (24)	45 % (131)	33 % (95)	7 % (19)	2 % (5)	5 % (14)	100 % (288)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	9 % (25)	29 % (84)	40 % (116)	13 % (38)	2 % (6)	7 % (19)	100 % (288)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	4 % (12)	18 % (52)	44 % (127)	19 % (56)	5 % (13)	10 % (28)	100 % (288)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	5 % (14)	44 % (126)	33 % (94)	8 % (24)	2 % (7)	8 % (23)	100 % (288)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbemærkninger og min bedømmelse	7 % (20)	44 % (128)	34 % (99)	4 % (11)	2 % (5)	9 % (25)	100 % (288)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en mundtlig prøve/eksamen i et af fagene dansk, engelsk eller matematik i løbet af 2017 eller 2018.

Gennemgang af resultater vedr. standpunktskarakter/årskarakterer

Tabel 3.32 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer. I det følgende oplystes andelen, der har svaret at elementet indgår i høj grad i karakterafgivelsen (enten ”I meget høj grad” eller ”I høj grad”), listet fra højeste andel til laveste andel:

- For 99 % indgår i høj grad ”Vurdering af elevernes faglige kompetence”
- For 64 % indgår i høj grad ” Vurdering af elevernes deltagelse”
- For 8 % indgår i høj grad ” Vurdering af elevernes opførsel”

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, for hvem ”Vurdering af elevernes deltagelse” indgår i høj grad, er statistisk signifikant forskellig mellem bedømmere i dansk (75 %) og bedømmere i engelsk (52 %), såvel som mellem bedømmere i dansk (75 %) og bedømmere i matematik (57 %).

TABEL 3.32

I hvilken grad indgår følgende elementer når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Total
Vurdering af elevernes faglige kompetence	73%	26%	1%	1%	0%	100%
	(110)	(39)	(1)	(1)	(0)	(151)
Vurdering af elevernes deltagelse	26%	38%	25%	9%	2%	100%
	(40)	(57)	(38)	(13)	(3)	(151)
Vurdering af elevernes opførsel	3%	5%	19%	28%	45%	100%
	(5)	(7)	(28)	(43)	(68)	(151)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne

Tabel 3.33 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om deres tildeling af standpunktskarakter/årskarakterer. I det følgende oplistes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 72 % af bedømmere af standpunktskarakter/årskarakterer svarer at de ”I meget høj grad” eller ”I høj grad” gør brug af ”Beskrivelserne af fælles mål/faglige mål i faget”
- 64 % af bedømmere af standpunktskarakter/årskarakterer svarer at de ”I meget høj grad” eller ”I høj grad” gør brug af ”Den generelle beskrivelse af karaktertrin i 7-trins-skalaen”
- 22 % af bedømmere af standpunktskarakter/årskarakterer svarer at de ”I meget høj grad” eller ”I høj grad” gør brug af ”De krav der i sidste ende stilles til den afsluttende prøve i faget”
- 22 % af bedømmere af standpunktskarakter/årskarakterer svarer at de ”I meget høj grad” eller ”I høj grad” gør brug af ”Lokalt udarbejdet hjælpemiddel (fx bedømmelseskema)”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der ”I meget høj grad” eller ”I høj grad” gør brug af hvert bedømmelsesværktøj, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.33

I hvilken grad gør du brug af følgende bedømmelsesværktøjer, når du afgiver standpunktskarakter/årskarakterer?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Den generelle beskrivelse af karaktertrin i 7-trins-skalaen	22%	42%	27%	8%	1%	0%	100%
	(33)	(63)	(41)	(12)	(2)	(0)	(151)
Beskrivelserne af fælles mål/faglige mål i faget	27%	45%	21%	5%	1%	0%	100%
	(41)	(68)	(32)	(8)	(2)	(0)	(151)
De krav der i sidste ende stilles til den afsluttende prøve i faget	5%	17%	19%	27%	32%	0%	100%
	(7)	(25)	(29)	(41)	(49)	(0)	(151)
Lokalt udarbejdet hjælpemiddel (fx bedømmelsesskema)	5%	17%	19%	27%	32%	0%	100%
	(7)	(25)	(29)	(41)	(49)	(0)	(151)
Andet	2%	0%	0%	0%	0%	98%	100%
	(3)	(0)	(0)	(0)	(0)	(148)	(151)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne

Tabel 3.34 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet ”I hvilken grad baserer du afgivelse af standpunktskarakterer/årskarakterer på følgende”. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 68 % bruger i høj grad ”Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)”
- 37 % bruger i høj grad ”Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle”
- 32 % bruger i høj grad ”Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)”
- 32 % bruger i høj grad ”Bedømmelsesværktøjer, som uddannelsen stiller til rådighed”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der ”I meget høj grad” eller ”I høj grad” gør brug af hvert bedømmelsesværktøj, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.34

I hvilken grad baserer du afgivelse af standpunktskarakterer/årskarakterer på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	23 % (34)	45 % (68)	20 % (30)	7 % (10)	3 % (4)	3 % (5)	100 % (151)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	9 % (13)	23 % (35)	36 % (54)	14 % (21)	13 % (20)	5 % (8)	100 % (151)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	5 % (8)	27 % (41)	24 % (36)	21 % (31)	18 % (27)	5 % (8)	100 % (151)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	9 % (13)	28 % (43)	34 % (51)	16 % (24)	9 % (13)	5 % (7)	100 % (151)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Tabel 3.35 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskarakterer. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 53 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 52 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 48 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 34 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 25 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål og min bedømmelse", er statistisk signifikant forskellig mellem bedømmere i dansk (42 %) og bedømmere i engelsk (53 %).

TABEL 3.35

Hvor enig eller uenig er du i følgende udsagn om bedømmelsesværktøjer ved afgivelse af standpunktskarakter/årskarakterer?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	11 % (17)	42 % (64)	31 % (47)	5 % (8)	2 % (3)	8 % (12)	100 % (151)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	5 % (7)	29 % (44)	42 % (64)	12 % (18)	2 % (3)	10 % (15)	100 % (151)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	4 % (6)	21 % (31)	44 % (66)	17 % (25)	3 % (5)	12 % (18)	100 % (151)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	6 % (9)	42 % (63)	33 % (50)	8 % (12)	3 % (4)	9 % (13)	100 % (151)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål og min bedømmelse	7 % (11)	45 % (68)	32 % (49)	5 % (8)	1 % (2)	9 % (13)	100 % (151)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen

Tabel 3.36 viser, hvad bedømmerne på erhvervsuddannelserne svarer på spørgsmålet "I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?". I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 78 % anvender i høj grad "Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)"
- 68 % anvender i høj grad "Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præsentationer, skriftlige produkter, håndsoprækninger, observation mv.)"
- 25 % anvender i høj grad "Resultater fra lokalt udviklede test"
- 14 % anvender i høj grad "Resultater fra standardiserede testredskaber (fx nationale test mv.)"

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der i høj grad anvender hver kilde, er ikke statistisk signifikant forskellig på tværs af fag.

TABEL 3.36

I hvilken grad anvender du følgende kilder når du skal vurdere elevernes faglige niveau i forbindelse med afgivelse af standpunktskarakter?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ikke relevant	Total
Mine erindringer af hvordan eleven har arbejdet med stoffet (fx erindringer af præstationer, skriftlige produkter, håndsoprækninger, observation mv.)	24 % (36)	44 % (67)	23 % (34)	7 % (11)	2 % (3)	0 % (0)	100 % (151)
Nedskrevne dokumenter om hvordan eleven har arbejdet med stoffet (fx elevplan, egne noter mv.)	37 % (56)	41 % (62)	17 % (25)	2 % (3)	3 % (5)	0 % (0)	100 % (151)
Resultater fra standardiserede testredskaber (fx nationale test mv.)	3 % (5)	11 % (17)	13 % (19)	25 % (37)	48 % (73)	0 % (0)	100 % (151)
Resultater fra lokalt udviklede test	9 % (13)	16 % (24)	24 % (36)	17 % (26)	34 % (52)	0 % (0)	100 % (151)
Andet	1 % (2)	3 % (4)	1 % (1)	0 % (0)	0 % (0)	95 % (144)	100 % (151)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskaracter/standpunktskarakter på baggrund af 7-trins-skalaen.

Gennemgang af resultater vedr. sammenhængen mellem læringsmål og bedømmelse

Tabel 3.37 viser, hvad bedømmerne på de erhvervsuddannelserne svarer i forbindelse med en række udsagn om sammenhængen mellem læringsmål og bedømmelse. I det følgende oplyses andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 79 % er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig"
- 55 % er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring"
- 33 % er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplyses alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring", er statistisk signifikant forskellig mellem bedømmere i dansk (45 %) og bedømmere i matematik (62 %).

TABEL 3.37

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring	12 % (39)	43 % (134)	28 % (87)	10 % (32)	4 % (13)	2 % (6)	1 % (3)	100 % (314)
Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne	8 % (24)	25 % (79)	30 % (93)	25 % (78)	10 % (31)	2 % (5)	1 % (4)	100 % (314)
Mine medbedømmere lægger typisk det faglige niveau samme sted som mig	16 % (50)	63 % (198)	16 % (50)	4 % (14)	0 % (1)	0 % (0)	0 % (1)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

3.5 Survey blandt bedømmere på de videregående uddannelser

3.5.1 Sammenfatning

Surveyen blandt bedømmere på de videregående uddannelser viser overordnet set, at:

- Bedømmerne i varierende grad anvender bedømmelsesværktøjerne, som er til rådighed ved mundtlige og skriftlige prøver
- Under halvdelen af bedømmerne er meget enig eller enige i, at bedømmelsesværktøjerne, der stilles til rådighed centralt eller af uddannelsen, er med til at sikre en ensartet bedømmelsespraksis, og at de giver et solidt udgangspunkt for bedømmelsen
- Størstedelen af bedømmerne er enige i, at deres medbedømmere typisk lægger det faglige niveau samme sted som dem, men lidt over halvdelen at bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring.
- Langt størstedelen af bedømmerne ønsker ikke supplerende bedømmelsesværktøjer til prøverne
- Lidt over halvdelen oplever, at bedømmelse af de afsluttende prøver er en retvisende afspejling af den studerendes kunnen.

3.5.2 Resultater vedr. anvendelse af bedømmelsesværktøjer

Tabel 3.38 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet ”I hvilken grad oplever du at følgende bedømmelsesværktøjer understøtter din afgivelse af bedømmelser efter 7-trins-skalaen?”. I det følgende oplystes andelen, der har svaret, at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 53 % understøttes i høj grad af "Beskrivelsen af de faglige læringsmål i institutionens studieordning"
- 51 % understøttes i høj grad af "Den generelle karakterbeskrivelse af 7-trins-skalaen"
- 46 % understøttes i høj grad af "Beskrivelsen af de faglige læringsmål i uddannelsesbekendtgørelsen"
- 22 % understøttes i høj grad af "Pensumoplistes"
- 21 % understøttes i høj grad af "Andre lokalt udarbejdede vejledninger mv. (fx bedømmelseskriterier)".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der i høj grad understøttes af "Pensumoplistes", er statistisk signifikant forskellig mellem bedømmere på Professionsbacheloruddannelserne (14 %) og bedømmere på erhvervsakademiuddannelserne (25 %), såvel som mellem bedømmere på universitetsuddannelser (24 %) og bedømmere på professionsbacheloruddannelser (14 %).
- Andelen, der i høj grad understøttes af "Andre lokalt udarbejdede vejledninger mv.", er statistisk signifikant forskellig mellem bedømmere på erhvervsakademiuddannelserne (28 %) og bedømmere på universitetsuddannelserne (19 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.38

I hvilken grad oplever du at følgende bedømmelsesværktøjer understøtter din afgivelse af bedømmelser efter 7-trins-skalaen?

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Kender ikke	Ikke relevant	Total
Den generelle karakterbeskrivelse af 7-trins-skalaen	12 % (362)	39 % (1152)	38 % (1127)	8 % (247)	1 % (34)	0 % (2)	1 % (22)	100 % (2946)
Beskrivelsen af de faglige læringsmål i uddannelsesbekendtgørelsen	9 % (86)	37 % (358)	31 % (305)	18 % (171)	3 % (30)	1 % (10)	1 % (12)	100 % (970)
Beskrivelsen af de faglige læringsmål i institutionens studieordning	13 % (388)	40 % (1168)	31 % (919)	11 % (335)	2 % (67)	1 % (19)	2 % (50)	100 % (2946)
Pensumoplistes	4 % (125)	18 % (527)	27 % (797)	27 % (797)	13 % (375)	2 % (71)	9 % (253)	100 % (2946)
Andre lokalt udarbejdede vejledninger mv. (fx bedømmelseskriterier)	5 % (147)	16 % (482)	23 % (689)	18 % (525)	12 % (352)	9 % (275)	16 % (475)	100 % (2946)
Andet	6 % (176)	7 % (203)	4 % (105)	1 % (20)	0 % (14)	1 % (21)	82 % (2407)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.39 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "I hvilken grad baserer du bedømmelser prøver på følgende?". I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 42 % bruger i høj grad "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)"
- 42 % bruger i høj grad "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle"
- 38 % bruger i høj grad "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed"
- 29 % bruger i høj grad "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der i høj grad bruger "Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelserne (37 %) og bedømmere på professionsbacheloruddannelserne (55 %), såvel som mellem bedømmere på universitetsuddannelserne (37 %) og erhvervsakademiuddannelserne (46 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (55 %) og bedømmere på erhvervsakademiuddannelserne (46 %).
- Andelen, der i høj grad bruger "Bedømmelsesværktøjer, som uddannelsen stiller til rådighed", er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelserne (45 %) og bedømmere på Universitetsuddannelserne (33 %), såvel som mellem bedømmere på universitetsuddannelserne (33 %) og erhvervsakademiuddannelserne (40 %).
- Andelen, der i høj grad bruger "Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelserne (29 %) og bedømmere på Professionsbacheloruddannelserne (23 %), såvel som mellem bedømmere på universitetsuddannelserne (29 %) og erhvervsakademiuddannelserne (39 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (23 %) og bedømmere på erhvervsakademiuddannelserne (39 %).
- Andelen, der i høj grad bruger "Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle", er statistisk signifikant forskellig mellem bedømmere på Professionsbacheloruddannelserne (30 %) og bedømmere på Universitetsuddannelserne (48 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (33 %) og erhvervsakademiuddannelserne (40 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.39

I hvilken grad baserer du bedømmelser på følgende:

	I meget høj grad	I høj grad	I nogen grad	I mindre grad	Slet ikke	Ved ikke	Total
Bedømmelsesværktøjer der stilles til rådighed centralt (bekendtgørelser i fag, om skalaen etc.)	11 % (336)	31 % (925)	30 % (889)	15 % (434)	8 % (247)	4 % (115)	100 % (2946)
Bedømmelsesværktøjer, som uddannelsen stiller til rådighed	9 % (251)	29 % (852)	26 % (758)	14 % (416)	15 % (450)	7 % (220)	100 % (2946)
Bedømmelsesværktøjer jeg selv udvikler (eventuelt i samarbejde med kollegaer)	7 % (212)	22 % (638)	23 % (686)	16 % (469)	26 % (753)	6 % (188)	100 % (2946)
Egne erfaringer – mens de forskellige bedømmelsesværktøjer spiller en mindre rolle	12 % (353)	32 % (933)	29 % (859)	15 % (441)	8 % (231)	4 % (129)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.40 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om bedømmelsesværktøjer. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 44 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen"
- 43 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse"
- 42 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis"
- 34 % er enige i udsagnet "Det er en udfordring ved mundtlige prøver, at man skal sætte niveauet ud fra den første bedømmelse alene"
- 34 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes"
- 27 % er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser"
- 15 % er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (50 %) og bedømmere på Universitetsuddannelserne (38 %), såvel som mellem bedømmere på universitetsuddannelserne (38 %) og bedømmere på professionsbacheloruddannelserne (48 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (41 %) og bedømmere på Universitetsuddannelserne (30 %), såvel som mellem bedømmere på universitetsuddannelserne (30 %) og bedømmere på professionsbacheloruddannelserne (43 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (18 %) og bedømmere på Universitetsuddannelserne (12 %), såvel som mellem bedømmere på universitetsuddannelserne (12 %) og bedømmere på professionsbacheloruddannelserne (20 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (51 %) og bedømmere på Universitetsuddannelserne (37 %), såvel som mellem bedømmere på universitetsuddannelserne (37 %) og bedømmere på professionsbacheloruddannelserne (48 %).
- Andelen, der er enige i udsagnet "Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (52 %) og bedømmere på Universitetsuddannelserne (38 %), såvel som mellem bedømmere på universitetsuddannelserne (38 %) og bedømmere på professionsbacheloruddannelserne (51 %).
- Andelen, der er enige i udsagnet "Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser", er statistisk signifikant forskellig mellem bedømmere på Erhvervsakademiuddannelserne (23 %) og bedømmere på Universitetsuddannelserne (30 %), såvel som mellem bedømmere på universitetsuddannelserne (30 %) og bedømmere på professionsbacheloruddannelserne (20 %).
- Andelen, der er enige i udsagnet "Det er en udfordring ved mundtlige prøver, at man skal sætte niveauet ud fra den første bedømmelse alene", er statistisk signifikant forskellig mellem bedømmere på professionsbacheloruddannelserne (30 %) og bedømmere på Erhvervsakademiuddannelserne (37 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.40

Hvor enig eller uenig er du i følgende udsagn om bedømmelsesværktøjerne?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen giver mig et solidt udgangspunkt for bedømmelsen	6 % (164)	37 % (1079)	33 % (982)	10 % (303)	2 % (65)	12 % (353)	100 % (2946)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel uddybes	5 % (160)	29 % (858)	38 % (1119)	11 % (329)	2 % (56)	14 % (422)	100 % (2946)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen kunne med fordel forenkles	2 % (69)	13 % (371)	41 % (1197)	24 % (708)	3 % (87)	17 % (514)	100 % (2946)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen er med til at sikre en ensartet bedømmelsespraksis	5 % (149)	37 % (1092)	29 % (855)	11 % (310)	3 % (88)	15 % (452)	100 % (2946)
Bedømmelsesværktøjerne der stilles til rådighed centralt eller af uddannelsen sikrer sammenhæng mellem fælles mål/faglige mål/målbeskrivelser og min bedømmelse	5 % (156)	38 % (1108)	31 % (914)	8 % (240)	2 % (59)	16 % (469)	100 % (2946)
Jeg fastsætter niveauet i bedømmelserne af skriftlige prøver ud fra et udsnit af besvarelser	4 % (86)	23 % (501)	21 % (455)	31 % (667)	17 % (379)	4 % (81)	100 % (2169)
Det er en udfordring ved mundtlige prøver, at man skal sætte niveauet ud fra den første bedømmelse alene	10 % (249)	24 % (600)	21 % (536)	25 % (635)	14 % (343)	6 % (151)	100 % (2513)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.41 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "Er der behov for supplerende bedømmelsesværktøjer ved følgende prøveformer?".

- 24 % svarer ja ved Afsluttende opgaver (specialer, bachelorprojekter, afsluttende projekter mv.)
- 20 % svarer ja ved Mundtlige prøver
- 18 % svarer ja ved Skriftlige prøver
- 17 % svarer ja ved Projektopgaver.

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. For alle prøvetyper er andelen, der svarer ja lavere mellem bedømmere på Universitetsuddannelserne i forhold til bedømmere på Professionsbacheloruddannelserne og Erhvervsakademiuddannelserne. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der svarede Ja til, at der er behov for supplerende bedømmelsesværktøjer ved "Mundtlige prøver", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (18 %) og bedømmere på professionsbacheloruddannelser (26 %), såvel som mellem bedømmere på universitetsuddannelser (18 %) og bedømmere på erhvervsakademiuddannelser (25 %).
- Andelen, der svarede Ja til, at der er behov for supplerende bedømmelsesværktøjer ved "Skriftlige prøver", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (16 %) og bedømmere på professionsbacheloruddannelser (23 %).
- Andelen, der svarede Ja til, at der er behov for supplerende bedømmelsesværktøjer ved "Projekt-opgaver", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (16 %) og bedømmere på erhvervsakademiuddannelser (26 %), såvel som mellem bedømmere på professionsbacheloruddannelser (18 %) og bedømmere på erhvervsakademiuddannelser (26 %).
- Andelen, der svarede Ja til, at der er behov for supplerende bedømmelsesværktøjer ved "Afsluttende opgaver (specialer, bachelorprojekter, afsluttende projekter mv.)", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (22 %) og bedømmere på professionsbacheloruddannelser (29 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.41

Er der behov for supplerende bedømmelsesværktøjer ved følgende prøveformer:

	Ja	Nej	Ved ikke	Total
Mundtlige prøver	20 % (601)	55 % (1612)	25 % (732)	100 % (2946)
Skriftlige prøver	18 % (532)	53 % (1559)	29 % (855)	100 % (2946)
Projektopgaver	17 % (511)	49 % (1441)	34 % (994)	100 % (2946)
Afsluttende opgaver (specialer, bachelorprojekter, afsluttende projekter mv.)	24 % (707)	53 % (1564)	23 % (676)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.42 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om bedømmelse af afsluttende prøver. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 58 % er enige i udsagnet "Bedømmelse af de afsluttende prøver er en retvisende afspejling af den studerendes kunnen"
- 38 % er enige i udsagnet "Læringsmålene for de afsluttende prøver understøtter en tydelig differentiering i de studerendes præstation"
- 33 % er enige i udsagnet "Det er naturligt, at karaktererne er højere i afsluttende opgaver end ved andre prøver"
- 27 % er enige i udsagnet "Omfattende vejledning i forbindelse med en prøve fx afsluttende opgaver gør det svært at bruge de lave karakterer i bedømmelserne".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet "Det er naturligt, at karaktererne er højere i afsluttende opgaver end ved andre prøver", er statistisk signifikant forskellig mellem bedømmere Universitetsuddannelserne (38 %), og bedømmere på Professionsbacheloruddannelserne (24 %), såvel som mellem bedømmere på universitetsuddannelser (38 %) og bedømmere på erhvervsakademiuddannelser (25 %).
- Andelen, der er enige i udsagnet "Bedømmelse af de afsluttende prøver er en retvisende afspejling af den studerendes kunnen", er statistisk signifikant forskellig mellem bedømmere Universitetsuddannelserne (61 %), og bedømmere på Professionsbacheloruddannelserne (47 %), såvel som mellem bedømmere på professionsbacheloruddannelser (47 %) og bedømmere på erhvervsakademiuddannelser (63 %).
- Andelen, der er enige i udsagnet "Læringsmålene for de afsluttende prøver understøtter en tydelig differentiering i de studerendes præstation", er statistisk signifikant forskellig mellem bedømmere Universitetsuddannelserne (36 %), og bedømmere på Erhvervsakademiuddannelserne (47 %), såvel som mellem bedømmere på professionsbacheloruddannelser (38 %) og bedømmere på erhvervsakademiuddannelser (47 %).
- Andelen, der er enige i udsagnet "Omfattende vejledning i forbindelse med en prøve (fx afsluttende opgaver) gør det svært at bruge de lave karakterer i bedømmelserne", er statistisk signifikant forskellig mellem bedømmere Universitetsuddannelserne (32 %), og bedømmere på Professionsbacheloruddannelserne (19 %), såvel som mellem bedømmere på universitetsuddannelser (32 %) og bedømmere på erhvervsakademiuddannelser (21 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.42

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Det er naturligt, at karaktererne er højere i afsluttende opgaver end ved andre prøver	6 % (162)	27 % (680)	24 % (607)	32 % (823)	8 % (213)	3 % (71)	100 % (2556)
Bedømmelse af de afsluttende prøver er en retvisende afspejling af den studerendes kunnen	8 % (199)	50 % (1275)	26 % (676)	12 % (301)	2 % (51)	2 % (55)	100 % (2556)
Læringsmålene for de afsluttende prøver understøtter en tydelig differentiering i de studerendes præstation	4 % (99)	34 % (873)	34 % (863)	14 % (370)	3 % (85)	10 % (266)	100 % (2556)
Omfattende vejledning i forbindelse med en prøve fx afsluttende opgaver gør det svært at bruge de lave karakterer i bedømmelserne	5 % (122)	22 % (567)	23 % (592)	34 % (865)	8 % (211)	8 % (199)	100 % (2556)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har bedømt en præstation på baggrund af 7-trins-skalaen til en afsluttende prøve i løbet af 2016, 2017 eller 2018.

Resultater vedr. sammenhængen mellem læringsmål og bedømmelse

Tabel 3.43 viser hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om sammenhængen mellem læringsmål og bedømmelse. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 74 % er enige i udsagnet "Mine medbedømmere lægger typisk det faglige niveau samme sted som mig"
- 70 % er enige i udsagnet "Der er på forhånd fastsat faglige mål til alle prøver, jeg bedømmer"
- 54 % er enige i udsagnet "Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring"
- 44 % er enige i udsagnet "Jeg oplever generelt, at de faglige mål er formuleret tydeligt og anvendeligt i forhold til brug i bedømmelsessituationen"
- 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"
- 37 % er enige i udsagnet "Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne"
- 20 % er enige i udsagnet "Jeg savner ofte konkrete bedømmelseskriterier for prøverne".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der er enige i udsagnet ”Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (34 %) og bedømmere fra erhvervsakademiuddannelserne (46 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (36 %) og bedømmere på erhvervsakademiuddannelserne (46 %).
- Andelen, der er enige i udsagnet ”Mine medbedømmere lægger typisk det faglige niveau samme sted som mig”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (74 %) og bedømmere fra professionsbacheloruddannelserne (69 %).
- Andelen, der er enige i udsagnet ”Der er på forhånd fastsat faglige mål til alle prøver, jeg bedømmer”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (68 %) og bedømmere fra professionsbacheloruddannelserne (75 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (75 %) og bedømmere på erhvervsakademiuddannelserne (67 %).
- Andelen, der er enige i udsagnet ”Jeg oplever generelt, at de faglige mål er formuleret tydeligt og anvendeligt i forhold til brug i bedømmelsessituationen”, er statistisk signifikant forskellig mellem bedømmere fra professionsbacheloruddannelserne (40 %) og bedømmere fra erhvervsakademiuddannelserne (49 %).
- Andelen, der er enige i udsagnet ”Jeg savner ofte konkrete bedømmelseskriterier for prøverne”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (18 %) og bedømmere fra professionsbacheloruddannelserne (25 %).
- Andelen, der 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”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (143 %) og bedømmere fra erhvervsakademiuddannelserne (53 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.43

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
Bedømmelser ud fra på forhånd fastsatte læringsmål fremmer fokus på de studerendes læring	11 % (314)	43 % (1275)	26 % (752)	12 % (364)	5 % (135)	3 % (74)	1 % (32)	100 % (2946)
Jeg drøfter altid forståelsen af læringsmål med min eventuelle medbedømmer, inden vi starter eksaminationerne	8 % (222)	29 % (858)	28 % (816)	27 % (786)	7 % (193)	1 % (17)	2 % (54)	100 % (2946)
Mine medbedømmere lægger typisk det faglige niveau samme sted som mig	12 % (341)	62 % (1812)	19 % (547)	6 % (188)	1 % (20)	1 % (15)	1 % (24)	100 % (2946)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
Der er på forhånd fastsat faglige mål til alle prøver, jeg bedømmer	20 % (581)	50 % (1470)	15 % (443)	10 % (305)	2 % (59)	2 % (45)	1 % (43)	100 % (2946)
Jeg oplever generelt, at de faglige mål er formuleret tydeligt og anvendeligt i forhold til brug i bedømmelsessituationen	7 % (212)	37 % (1078)	30 % (884)	19 % (567)	5 % (136)	1 % (41)	1 % (27)	100 % (2946)
Jeg savner ofte konkrete bedømmeskriterier for prøverne	3 % (103)	17 % (495)	27 % (806)	37 % (1081)	13 % (371)	2 % (56)	1 % (34)	100 % (2946)
Konkrete faglige mål med præcise referencer til specifik viden, kompetencer eller indhold giver en bedre understøttelse af bedømmelserne end generiske mål om overordnede kompetencer	10 % (307)	34 % (1013)	29 % (862)	14 % (411)	7 % (197)	4 % (108)	2 % (47)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.44 viser at 59 % af bedømmerne på de videregående uddannelser svarer ja på spørgsmålet "Bidrager du selv til udarbejdelsen af de faglige læringsmål?"

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (66 %) og bedømmere fra professionsbacheloruddannelserne (42 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (68 %) og bedømmere på erhvervsakademiuddannelserne (63).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.44

Bidrager du selv til udarbejdelsen af de faglige læringsmål?

	Procent (antal)
Ja	59 % (1737)
Nej	41 % (1209)
Total	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 3.45 viser, hvad bedømmerne på de videregående uddannelser svarer spørgsmålet ”I hvilken grad føler du dig klædt på til at udarbejde faglige læringsmål?”. Spørgsmålet er stillet til bedømmere, der selv bidrager til udarbejdelsen af de faglige læringsmål. 78 % svarer enten ”I meget høj grad” eller ”I høj grad”.

Variationer i svar mellem bedømmerne fra hhv. universitetsuddannelserne, professionsbacheloruddannelserne og erhvervsakademiuddannelserne, samt de kunstneriske – og maritime uddannelser er at finde i den særskilte tabelrapport. Andelen, der har svaret ”I meget høj grad” eller ”I høj grad”, er *ikke* statistisk signifikant forskellig mellem grupperne. Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.45

I hvilken grad føler du dig klædt på til at udarbejde faglige læringsmål?

	Procent (antal)
I meget høj grad	31 % (530)
I høj grad	47 % (814)
I nogen grad	21 % (356)
I mindre grad	2 % (29)
Slet ikke	0 % (3)
Ved ikke	0 % (4)
Total	100 % (1737)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der selv bidrager til udarbejdelsen af de faglige læringsmål.

Tabel 3.46 viser, at 89 % af bedømmerne på de videregående uddannelser svarer ja på spørgsmålet ”Kan du få hjælp til udarbejdelse af læringsmålene, hvis du har spørgsmål?”

Variationer i svar mellem bedømmerne fra hhv. universitetsuddannelserne, professionsbacheloruddannelserne og erhvervsakademiuddannelserne, samt de kunstneriske – og maritime uddannelser er at finde i den særskilte tabelrapport. Andelen, der har svaret ja, er *ikke* statistisk signifikant forskellig mellem grupperne. Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.46

Kan du få hjælp til udarbejdelse af læringsmålene, hvis du har spørgsmål?

	Procent (antal)
Ja	89 % (1549)
Nej	11 % (188)
Total	100 % (1737)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der selv bidrager til udarbejdelsen af de faglige læringsmål.

Tabel 3.47 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "Hvem godkender dine faglige læringsmål?". Spørgsmålet er stillet til bedømmere, der selv bidrager til udarbejdelsen af de faglige læringsmål.

- 45 % svarer "Administrativ leder fx uddannelseschef eller studiechef"
- 21 % svarer "Leder for et mindre fagligt miljø fx institutleder"
- 8 % svarer "Leder for større fagligt område fx dekan"
- 2 % svarer "Rektor".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Rektor", er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (1 %) og bedømmere fra professionsbacheloruddannelserne (4 %), såvel som mellem bedømmere på universitetsuddannelserne (1 %) og bedømmere på erhvervsakademiuddannelserne (5 %).
- Andelen, der har svaret "Administrativ leder fx uddannelseschef eller studiechef", er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (43 %) og bedømmere fra erhvervsakademiuddannelserne (59 %), såvel som mellem bedømmere på professionsbacheloruddannelserne (46 %) og bedømmere på erhvervsakademiuddannelserne (59 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 3.47

Hvem godkender dine faglige læringsmål? (n=1737)

	Procent (antal)
Rektor	2 % (38)
Leder for større fagligt område fx dekan	8 % (145)
Leder for et mindre fagligt miljø fx institutter	21 % (365)
Administrativ leder fx uddannelseschef eller studiechef	45 % (789)
Anden, angiv hvem:	31 % (534)
Ved ikke	8 % (142)
Total	116 % (1737)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet som et multiple-choice spørgsmål, hvorfor procenterne summer til mere end 100.

Note: Spørgsmålet er stillet til bedømmere, der selv bidrager til udarbejdelsen af de faglige læringsmål.

3.6 Fokusgruppeinterviews med bedømmere

I det følgende beskrives bedømmernes anvendelse af bedømmelsesværktøjer, som det er blevet beskrevet i fokusgruppeinterviewene.

3.6.1 Bedømmere på grundskoleområdet

Bedømmerne beskriver, at de i deres bedømmelse af elever på forskellig vis anvender følgende bedømmelsesværktøjer i de enkelte fag: vejledende karakterbeskrivelser og vurderingskriterier samt vurderings- og evalueringsskemaer fra fagenes prøvevejledninger. I nogle fag anvendes der desuden mere konkrete bedømmelsesværktøjer. Det gælder fx i skriftlig matematik og dele af de skriftlige engelsk- og danskprøver, hvor der anvendes konkrete rettevejledninger med eksempler og om-sætningstabeller til at oversætte antallet af point til karakterer. Indtrykket fra fokusgruppeinterviewene er, at den generelle trinbeskrivelse anvendes meget lidt, da den vurderes at være for overordnet til at kunne bruges som konkret bedømmelsesværktøj. Generelt vurderer bedømmerne, at de nuværende bedømmelsesværktøjer fungerer godt og anvendes aktivt. Eksaminator og censor anvender typisk fagenes vejledende karakterbeskrivelser og vurderingskriterier i deres dialog om de forskellige elementer i elevens præstation og karakterfastsættelsen. Brugen af bedømmelsesværktøjerne kan dog ikke ændre på, at bedømmerne oplever, at karaktergivning indebærer elementer af fortolkning, fx ift. hvornår noget er en væsentlig mangel.

Den generelle trinbeskrivelse er for overordnet i sin ordlyd

Et gennemgående perspektiv i interviewene med bedømmerne er, at ordlyden i den generelle trinbeskrivelse efterlader et for stort fortolkningsrum. Bedømmerne oplever, at de bærende ord i beskrivelsen af de forskellige karakterer som fx "jævnt", "fortrinligt" og "utilstrækkeligt" er for overordnede og derfor ikke tilstrækkelig præcise som bedømmelsesgrundlag. En særlig kritik rettes imod 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. Nogle bedømmere fortæller, hvordan de har den generelle trinbeskrivelse i baghovedet under prøven og bruger beskrivelserne, fx "jævn" eller "tilstrækkelig", men at den ikke er tilstrækkelig som bedømmelsesværktøj, fordi det kræver flere ord og fagspecifikke parametre at bedømme elevens præstation.

Bedømmerne gør stor brug af fagenes vurderingsskemaer og vejledende karakterbeskrivelser

På tværs af fagene peger bedømmerne på, at de forskellige evaluerings- og vurderingsskemaer fra prøvevejledningerne er væsentlige redskaber i bedømmelsesprocessen. Nogle bedømmere bruger aktivt vurderings- og evalueringsskemaer under selve prøven, mens andre primært tager noter og først anvender skemaerne i voteringen med censor. Et perspektiv er, at brugen af afkrydsningsskemaet under selve prøven kan forstyrre kontakten til eleven og gøre eleven unødigt nervøs. Samtidig oplever nogle bedømmere det som mere anvendeligt at tage noter om indholdet af elevens præstation som afsæt for at gennemgå vurderingsskemaet i dialogen med censor. Andre bedømmere oplever derimod, at afkrydsningsskemaet er nemt at anvende under selve prøven og en hurtig måde at få overblik over elevens præstation på, når de skal votere med censor.

Gennemgående peger bedømmerne på, at det er vigtigt, at de vejledende karakterbeskrivelser dækker hvert trin på skalaen og ikke kun top-, midt- og bundpræstationen, som det tidligere har været tilfældet i nogle fag. Samtidig peger de på, at fagenes vejledende karakterbeskrivelser og vurderingsskemaer bruges som afsæt for at drøfte de forskellige elementer i elevens præstation, fx elevens evne til at kunne analysere og perspektivere, med censor. Generelt udtrykker bedømmerne tilfredshed med de nuværende fagspecifikke bedømmelsesværktøjer. For de skriftlige censorer gælder desuden, at der er tilfredshed med de udvidede rettevejledninger, der udsendes af Undervisningsministeriet, fordi de konkrete eksempler på bedømmelser hjælper dem i deres retteproces.

Karaktergivning rummer elementer af fortolkning, og derfor er bedømmelsesværktøjerne vigtige ved tvivl eller uenighed

Bedømmerne fortæller, at bedømmelsesværktøjerne særligt anvendes, hvis der er uklarhed eller uenighed om, hvilken karakter eleven skal have. Erfaringerne er, at fagenes vejledende karakterbeskrivelser giver et godt afsæt for at drøfte elevens karakter, men at det trods vejledningerne kan være vanskeligt at afgøre, om eleven fx skal have 7 eller 10:



Jeg synes stadig, det kan være svært at afgøre, om eleven skal have 7 eller 10. Det er så stort et fortolkningsarbejde, at man kan lægge hvad som helst i det. Og hvad er "utilstrækkeligt" og hvad er "jævnt"? Det handler rigtig meget om at opnå konsensus med censor.

Ligeledes peger en gruppe af bedømmerne på, at selvom bedømmelsesværktøjerne understøtter bedømmelsesprocessen og dialogen mellem censor og eksaminator, så kan det være en udfordring, at censor og eksaminator ikke nødvendigvis har samme tolkning eller vægtning af elementerne i de fagspecifikke vurderingskriterier. Blandt nogle bedømmere er erfaringen, at de ofte vil kunne argumentere for to forskellige karakterer med afsæt i vurderingskriterierne. I sådanne tilfælde afgøres karakteren gennem drøftelse af vægtningen af elementerne i elevens præstation

sammenholdt med tolkning af, hvornår noget fx vurderes til at være en væsentlig mangel. I sådanne tilfælde spiller fagenes vejledende karakterbeskrivelser en væsentlig rolle, fordi de kobler de fagspecifikke kriterier til karakterskalaen og derfor giver bedømmerne et konkret afsæt for deres drøftelser. Ligeledes peger nogle bedømmere på, at de bruger vurderingsskemaerne med afkrydsning til at sikre, at hver enkelt elev får en absolut bedømmelse på baggrund af sin præstation, og at der ikke foretages en relativ bedømmelse mellem eleverne i en klasse. Gennemgående har bedømmerne dog oplevelsen af, at der sjældent er væsentlige uenigheder i bedømmelsen af eleverne. Der, hvor der er uenighed, er der typisk kun en enkelt karakter til forskel i de indledende karakterudspil i dialogen mellem lærer og censor:



Jeg oplever ikke en masse uenigheder, men der er da altid nogle. Så sidder vi og argumenterer frem og tilbage og henviser til bestemte steder i eksamen eller i vurderingsskemaerne. I de situationer er det rart at have nogle mål og krav, som man kan bedømme ud fra.

Ved bedømmelse af de skriftlige prøver er der modsat tidligere kun én bedømmer i form af en beskikket censor. Derfor har censorerne ikke samme mulighed for at indgå i en dialog eller argumentation om bedømmelsen. Censorerne peger på, at de i tvivlstilfælde ofte bruger vurderingsskemaerne, men at de også betragter opgaven ud fra en helhedsvurdering. Generelt peger censorerne på, at de mangler den sparring, der tidligere var med en medbedømmer.

3.6.2 Bedømmere på det gymnasiale område

Bedømmerne beskriver, at de i deres bedømmelse af elever i forskelligt omfang og på forskellig vis anvender den generelle trinbeskrivelse, de fagspecifikke karakterbeskrivelser for de enkelte fag i gymnasiet, de faglige mål fra læreplaner, vejledningen til læreplanerne og de taksonomiske niveauer. I forbindelse med mundtlige prøver indgår eksaminators undervisningsbeskrivelse også, fordi den afgrænser, hvad eleven er blevet undervist i og dermed kan blive bedømt på. I nogle fag anvendes der også særegne bedømmelsesværktøjer, fx i skriftlig matematik, hvor der anvendes en omsætningstabel til at oversætte point til karakterer. Nogle bedømmere anvender også forskellige skemaer, skabeloner eller retteark udviklet af fagkonsulenten, den faglige forening eller dem selv eller kollegaer.

Den generelle trinbeskrivelse og de faglige mål efterlader et fortolkningsrum

Hos en gruppe af bedømmerne går det igen, at de oplever, at formuleringerne i den generelle trinbeskrivelse såvel som i de faglige mål efterlader et fortolkningsrum til bedømmerne. Ift. de faglige mål oplever nogle bedømmere de faglige mål som mere konkrete, end andre bedømmere gør, hvilket også kan hænge sammen med fagforskelle, hvor eksempelvis de faglige mål i matematik kan være mere konkrete end de faglige mål i dansk. Men ift. trinbeskrivelsen er det en generel oplevelse, at nogle formuleringer kan være vanskelige at afkode og skelne fra hinanden. Eksempelvis opleves det ikke nødvendigvis som entydigt, hvad der er ”få uvæsentlige mangler” (som udløser et 12-tal) kontra ”nogle mindre væsentlige mangler” (som udløser et 10-tal). Ligesom begreber som den ”fremragende” og den ”fortrinlige” præstation for nogle bedømmere opleves som overlappende. Derudover peger bedømmerne også på, at det at tale om antal fejl og mangler (”få”, ”nogle”, ”adskillige”) er mere meningsfuldt i nogle fag, såsom matematik, end i andre fag, såsom dansk, hvor fejl og mangler ikke på samme vis lader sig kvantificere.

Nogle bedømmere fortæller, at de er meget bevidste om at bruge formuleringer fra trinbeskrivelsen, når de skal formidle og begrunde deres karakter over for eleven i forbindelse med en mundtlig eksamen. Herunder at de er påpasselige med ikke at bruge ord, der indgår i beskrivelser af karaktertrin højere på skalaen end den karakter, eleven får, blandt andet for at undgå klager.

Bedømmernes hovedanke imod trinbeskrivelsen synes dog ikke at være, at den anvender generelle vendinger og ikke er konkret og tydelig, men som før nævnt at den opleves beskrevet som en mangelskala. Nogle bedømmere nævner også, at når karakterskalaen skal gælde på tværs af uddannelsesområder, fag og prøveformer, så er der også grænser for, hvor konkret beskrevet den kan blive. Derfor gør bedømmerne på de gymnasiale uddannelser stor brug af fagspecifikke karakterbeskrivelser frem for den generelle trinbeskrivelse.

Bedømmerne gør stor brug af fagspecifikke karakterbeskrivelser

Generelt fortæller bedømmerne, at de gør stor brug af de fagspecifikke karakterbeskrivelser, der er i vejledningen til læreplanerne. Der er typisk tale om beskrivelser af, hvad der skal til for at få hhv. 02, 7 og 12 i faget. Disse anvendes i stor udstrækning, fordi de opleves som en brugbar oversættelse og konkretisering af den generelle trinbeskrivelse til det specifikke fag på gymnasialt niveau, der skal bedømmes.

Fordi en del af bedømmerne gør stor brug af de fagspecifikke karakterbeskrivelser, efterlyser nogle bedømmere, at der udarbejdes fagspecifikke karakterbeskrivelser for hvert trin på skalaen frem for kun for top-, midt- og bundpræstationen. Dog anerkender bedømmerne samtidig, at det måske kan være svært at specificere og adskille de fagspecifikke krav for hver enkelt karakter i tilstrækkelig grad til, at det bliver meningsfuldt og anvendeligt.

Taksonomiske niveauer bruges i vid udstrækning i bedømmelsen

Ud over den generelle trinbeskrivelse, fagets faglige mål og de fagspecifikke karakterbeskrivelser fremhæver bedømmerne taksonomiske modeller som et centralt bedømmelsesværktøj i karaktergivningen. Det kan være enten Blooms taksonomi eller SOLO-taksonomien. En del af bedømmerne bruger taksonomiske niveauer til, i kombination med de faglige mål, at indplacere elevernes præstationer karaktermæssigt. Nogle bedømmere fortæller, at de faglige mål ikke nødvendigvis er taksonomiske i sig selv, og idet man kan opfylde de samme faglige mål på forskellige taksonomiske niveauer, opleves de taksonomiske niveauer som en stor hjælp i bedømmelsen. En bedømmer fortæller:



Hvis man kan genkende og redegøre for begreber, så ligger man der. Hvis man også kan forklare, hvor de kommer fra, og vise, man forstår dem, så ligger man her. Hvis man kan kombinere det og lave en analyse, så ligger man der, etc.

Ud over at en del af bedømmerne benytter taksonomiske niveauer i bedømmelsen af eleverne, så gør de også brug af taksonomier i kommunikationen med eleverne om, hvad der skal til for at få hvilke karakterer, og i deres formidling af afgivne karakterer til eleverne. Det hænger også sammen med, at bedømmerne oplever, at det kan være lettere at formidle de taksonomiske niveauer frem for de faglige mål til eleverne. En bedømmer siger: "Når eleverne spørger, hvad der skal til for at få 12, så er det jo den [Blooms taksonomi], jeg tager fat i, og ikke de faglige mål."

Undervisningsbeskrivelser og spørgsmålsformuleringer afgrænser, hvilke faglige mål der kommer i spil i bedømmelsen

Bedømmerne beskriver, at de forud for bedømmelsen af en mundtlig eller skriftlig prøve typisk genopfrisker karakterbeskrivelserne og fagets faglige mål. Men bedømmerne understreger også, at det ved en mundtlig eksamen er centralt at holde fagets faglige mål op imod eksaminators undervisningsbeskrivelse af, hvad klassen har arbejdet med i faget, samt op imod de stillede eksamensspørgsmål, idet disse afgrænser, hvilke faglige mål man kan bedømme eleverne ud fra. En bedømmer siger: "Man kan ikke forvente noget af eleverne, som de ikke er blevet undervist i." Det vil sige, at bedømmelsen af eleverne kun sker ud fra de faglige mål, der er relevante ift. den undervisning,

eleverne har fået, og de spørgsmål, der stilles til eksamen. På den måde udgør undervisningsbeskrivelser og spørgsmålsformuleringer også vigtige bedømmelsesredskaber. Samtidig understreger dette flere bedømmers pointe om, at ansvaret for, at de faglige mål kommer i spil til eksamen, ikke ligger hos eleven, men ved den mundtlige eksamen hos eksaminator i forbindelse med udformningen af eksamensspørgsmål og ved den skriftlige eksamen hos opgavekommissionen i forbindelse med formuleringen af den skriftlige opgave. Elevernes opgave er så at besvare de stillede spørgsmål, som afspejler et eller flere faglige mål.

Bedømmelsesværktøjer kommer særligt i spil ved tvivl eller uenighed og understøtter ensartethed

Bedømmerne beskriver, at bedømmelsesværktøjerne særligt kommer i spil, hvis der opstår tvivl eller uenighed om bedømmelsen af en elev, og at bedømmelsesværktøjerne derfor er med til at understøtte en ensartethed i bedømmelsen. Flere bedømmere peger på, at bedømmelsesværktøjerne er vigtige ift. at skabe ensartethed i bedømmelsen og i tolkningen af læreplanen, så man ikke risikerer, at bedømmernes individuelle kæpheste inden for faget kommer til at blive styrende. En bedømmer siger:



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.

3.6.3 Bedømmere på de videregående uddannelser

I modsætning til på grundskoleområdet og det gymnasiale område findes der ikke mange centralt udviklede bedømmelsesredskaber på det videregående område. De primære bedømmelsesredskaber er læringsmålene og trinbeskrivelsen.

Enkelte prøver har rettevejledninger og rettes efter procentsatser

I fokusgruppeinterviewene beskriver bedømmere, hvordan de til nogle skriftlige prøver har rettevejledninger, og at prøverne rettes efter procentsatser. Det er primært i de tekniske og naturvidenskabelige fag. Tre bedømmere fortæller nedenfor, hvordan det fungerer:



Det afhænger af, hvilken type opgave det er. Ved hovedopgaver er det meget intuitivt. Fordi jeg har meget erfaring. Så ved jeg, hvad der trækker op og ned. Hvis der ved skriftlige opgaver er point og rettevejledning, så er det fælles for censor og eksaminator – der har vi et fælles grundlag. Der sidder vi egentlig og lægger til og trækker fra. Det er typisk de mundtlige, som er lidt sværere at bedømme: Der er vi der eller der. De mundtlige eksaminer foregår jo som regel med udgangspunkt i et skriftligt produkt. Der har jeg vurderet produktet først og forventningsafstemmer med censor.



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



Måske er det den, hvor man tæller op nedefra. For til de skriftlige opgaver følger der en rettevejledning med, hvilket gør vurderingen lidt kvantitativ. For hver opgave tæller man op, hvor mange procent rigtige svar der er. Og selvfølgelig: Hvis der er noget, som er helt forkert, kan det tælle mere eller mindre end andre fejl. Men der tilsendes en skala, som gør det muligt at gå fra procent til karakter. Så på den måde lægger jeg jo egentlig sammen nedefra og op.

Trinbeskrivelsen og læringsmålbeskrivelserne anvendes aktivt

Indtrykket fra fokusgrupperne er, at bedømmerne i høj grad anvender trinbeskrivelsen og læringsmålbeskrivelserne i bedømmelsessituationer. De kommer særligt i spil ved tvivl eller uenighed og understøtter ensartethed i bedømmelsen, men generelt er indtrykket fra fokusgrupperne, at både læringsmålbeskrivelserne og trinbeskrivelsen anvendes meget aktivt. Fx fortæller en bedømmer, at alle eksamensspørgsmål til den mundtlige eksamen er formuleret med udgangspunkt i læringsmålene.

Læringsmålene kan være svære at anvende til bedømmelse

Bedømmerne beskriver i fokusgrupperne, at læringsmålene er deres primære bedømmelsesredskab. Et gennemgående tema er dog, at det ikke altid er lige let at anvende læringsmålene i en bedømmelsessituation, hvilket illustreres i citaterne nedenfor:



Bedømmelsen skulle jo gerne ske på baggrund af målbeskrivelserne. Dilemmaet er, at målbeskrivelserne ikke altid er så præcist beskrevet. 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 vi kigger på de overordnede mål og de delmål, vi selv har formuleret.



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.



Jeg oplevede i starten, at læringsmålene var for abstrakte. Både ift. karaktergivning og ift., 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 ultraspæcifikke.



Der var nogle børnesygdomme [ved læringsmålene]. Vi lavede nogle mere runde formuleringer, og nu er de blevet vores bedste ven ift. eksamensklager. De gør det faktisk ekstremt nemt at arbejde med skalaen. Hvis man har lavet gode beskrivelser.

Men med til historien hører også, at bedømmerne på nogle uddannelser oplever, at læringsmålene revideres meget ofte:



Hos os bliver læringsmålene revideret ofte. Der er stærk opmærksomhed om det. Jeg kan godt få kommentarer fra studieledelsen om, at noget skal opdateres. Og så er det jo nærmest, som om der skulle afholdes en ny prøve hvert år.

Andre bedømmere oplever ikke, at læringsmålene revideres særligt ofte:



Aktuelt er det ti år siden, vi reviderede bacheloruddannelserne, men nu gør vi det igen. Der har været smårettelser undervejs.

Trinbeskrivelse efterlader et fortolkningsrum

Selvom bedømmerne anvender trinbeskrivelsen aktivt, så er der delte meninger om, hvorvidt beskrivelserne af de enkelte karakterer er fyldestgørende. Nogle bedømmere oplever, at beskrivelserne efterlader et stort fortolkningsrum:



Ord som ”fortrinlig” og ”fremragende” er jo tilnærmelsesvis synonyme. Det er altså en tynd brug af tillægsord. Vi ved jo bare, at 12 er bedre end 10. Det er problematisk. Men jeg ved ikke, hvordan man kunne gøre det tydeligere.



Jeg er glad for, at de er der [beskrivelserne], for jeg bruger dem aktivt. Men jeg kan godt være i tvivl om, hvordan de skal bruges, og hvordan jeg vurderer ”en del mangler”.

3.7 Enkeltinterviews med bedømmere

I det følgende beskrives bedømmernes anvendelse af bedømmelsesværktøjer, som det er blevet beskrevet i de individuelle interviews.

3.7.1 Bedømmere på grundskoleområdet

Hovedpointerne ift. bedømmelsesredskaber er i stor udstrækning overlappende for fokusgrupper og individuelle interviews med bedømmere fra grundskolen.

Bedømmerne beskriver, at de i deres bedømmelse af elever på forskellig vis anvender følgende bedømmelsesværktøjer for de enkelte fag: vejledende karakterbeskrivelser, vurderingskriterier samt vurderings- og evalueringsskemaer fra fagenes prøvevejledninger. For nogle fag anvendes der desuden mere konkrete bedømmelsesværktøjer. Det gælder fx i skriftlig matematik og dele af de skriftlige engelsk- og danskprøver, hvor der anvendes konkrete rettevejledninger med eksempler og om-sætningstabeller for at oversætte antal af point til karakterer. Kun få anvender den generelle trinbeskrivelse, der vurderes at være for overordnet til at kunne bruges som konkret bedømmelsesværktøj.

Den generelle trinbeskrivelse er for overordnet i sin ordlyd

Et gennemgående perspektiv i interviewene med bedømmerne er, at ordlyden i den generelle trinbeskrivelse efterlader et for stort fortolkningsrum. Bedømmerne oplever, at de bærende ord i beskrivelsen af de forskellige karakterer, fx ”jævnt”, ”fortrinligt” og ”utilstrækkeligt”, er for overordnede og derfor ikke tilstrækkelig præcise som bedømmelsesgrundlag. Nogle bedømmere fortæller, hvordan de har den generelle trinbeskrivelse i baghovedet under prøven, men at den ikke er tilstrækkelig som bedømmelsesværktøj, fordi det kræver flere ord og fagspecifikke parametre at bedømme elevens præstation.

Bedømmerne gør stor brug af fagenes vurderingsskemaer og vejledende karakterbeskrivelser

På tværs af fagene peger bedømmerne på, at de forskellige evaluerings- og vurderingsskemaer fra prøvevejledningerne er væsentlige redskaber i bedømmelsesprocessen. Nogle bedømmere bruger aktivt vurderings- og evalueringsskemaer under selve prøven, mens andre primært tager noter og først anvender skemaerne i voteringen med censor.

Generelt udtrykker bedømmerne tilfredshed med de nuværende fagspecifikke bedømmelsesværktøjer. For de skriftlige censorer gælder desuden, at der er tilfredshed med de udvidede rettevejledninger, der udsendes af Undervisningsministeriet, fordi de konkrete eksempler på bedømmelser hjælper dem i deres retteproces.

Karaktergivning rummer elementer af fortolkning, derfor er bedømmelsesværktøjerne vigtige ved tvivl eller uenighed

Bedømmerne fortæller, at bedømmelsesværktøjerne særligt anvendes, hvis der er uklarhed eller uenighed om, hvilken karakter eleven skal have. Erfaringerne er, at fagenes vejledende karakterbeskrivelser giver et godt afsæt for at drøfte elevens karakter, men at det trods vejledningerne kan være vanskeligt at afgøre, om eleven fx skal have 7 eller 10.

Ligeledes peger flere bedømmere på, at selvom bedømmelsesværktøjerne understøtter bedømmelsesprocessen og dialogen mellem censor og eksaminator, så kan det være en udfordring, at censor og eksaminator ikke nødvendigvis har samme tolkning eller vægtning af elementerne i de fagspecifikke vurderingskriterier. Gennemgående har bedømmerne dog oplevelsen af, at der sjældent er væsentlige uenigheder i bedømmelsen af eleverne. Der, hvor der er uenighed, er der typisk kun en enkelt karakter til forskel på de indledende karakterudspil i dialogen mellem lærer og censor.

Ved bedømmelse af de skriftlige prøver er der modsat tidligere kun én bedømmer i form af en beskikket censor. Derfor har censorerne ikke samme mulighed for at indgå i en dialog eller argumentation om bedømmelsen. Censorerne peger på, at de i tvivlstilfælde ofte bruger vurderingsskemaerne, men at de også betragter opgaven ud fra en helhedsvurdering. Generelt peger censorerne på, at de mangler den sparring, der tidligere var med en medbedømmer.

3.7.2 Bedømmere på det gymnasiale område

Hovedpointerne ift. bedømmelsesredskaber er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere fra de gymnasiale uddannelser. I dette afsnit går en del af overskrifterne fra afsnittet om fokusgrupperne således igen, men pointerne illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Bedømmerne beskriver, at de i deres bedømmelse af elever i forskelligt omfang og på forskellig vis anvender den generelle trinbeskrivelse, de fagspecifikke karakterbeskrivelser for de enkelte fag i gymnasiet, de faglige mål fra læreplaner, vejledningen til læreplanerne og de taksonomiske niveauer. I forbindelse med mundtlige prøver indgår eksaminators undervisningsbeskrivelse også, fordi den afgrænser, hvad eleven er blevet undervist i og dermed kan blive bedømt på. I nogle fag anvendes der også særegne bedømmelsesværktøjer, fx i skriftlig matematik, hvor der anvendes en omsætningstabel til at oversætte point til karakterer. Nogle bedømmere anvender også forskellige skemaer, skabeloner eller retteark udviklet af fagkonsulenten, den faglige forening eller dem selv eller kollegaer. Generelt beskriver bedømmerne, at de gør begrænset brug af bedømmelsesværktøjerne undervejs i selve bedømmelsen, men at de læser dem forud for bedømmelsen og tager dem frem ved uenighed eller tvivl.

Den generelle trinbeskrivelse og de faglige mål efterlader et fortolkningsrum

Hos en del af bedømmerne går det igen, at de oplever, at formuleringerne i den generelle trinbeskrivelse såvel som i de faglige mål efterlader et fortolkningsrum til bedømmerne. Ift. de faglige mål oplever nogle bedømmere de faglige mål som mere konkrete, end andre bedømmere gør, hvilket også kan hænge sammen med fagforskelle, hvor eksempelvis de faglige mål i matematik kan være mere konkrete end de faglige mål i dansk. Men ift. trinbeskrivelsen er det en generel oplevelse, at nogle formuleringer kan være vanskelige at afkode og skelne fra hinanden. En bedømmer siger:



Alle de ord, der står her. Hvad betyder ”fortrinlig”? Hvornår er fagets mål omfattende opfyldt? Hvad er en mindre væsentlig mangel? Trinbeskrivelsen er lige så elastikagtig som fagets mål.

Samtidig peger nogle af bedømmerne dog også på, at det ikke nødvendigvis er muligt eller ønskeligt at gøre trinbeskrivelsen og fagets mål meget konkrete. En bedømmer beskriver, at det er nødt til at være sådan, hvis man ønsker metodefrihed i undervisningen og en faglig helhedsvurdering i bedømmelsen:



Hvis man gør trinbeskrivelsen og fagets mål meget konkrete, så bliver det samlebandsundervisning og -bedømmelse. Så kunne man lige så godt have multiple choice-prøver.

Andre bedømmere mener, at det er acceptabelt, at den generelle trinbeskrivelse ikke er mere konkret, fordi den skal favne forskellige fag, uddannelsesområder og prøveformer – så længe trinbeskrivelsen suppleres af mere konkrete bedømmelsesværktøjer på de enkelte uddannelsesområder og inden for de enkelte fag, som fx de fagspecifikke karakterbeskrivelser på det gymnasiale område.

Bedømmerne gør stor brug af fagspecifikke karakterbeskrivelser

En del af bedømmerne fortæller, at de meget bevidst anvender formuleringer fra den generelle trinbeskrivelse, når de skal formidle karakterer til eleverne, men at de i deres bedømmelse af eleverne i højere grad anvender de fagspecifikke karakterbeskrivelser, der typisk beskriver, hvad der skal til for at få 02, 7 eller 12 i faget. Bedømmerne beskriver, at de kun i begrænset udstrækning bruger den generelle trinbeskrivelse direkte i bedømmelsen af elever, men at den anvendes indirekte, idet de i stor udstrækning anvender den fagspecifikke oversættelse og fortolkning af skalaen, som de fagspecifikke karakterbeskrivelser udgør. Fordi de fagspecifikke karakterbeskrivelser anvendes hyppigt og findes anvendelige, efterspørger nogle bedømmere, at der bliver udarbejdet fagspecifikke karakterbeskrivelser for samtlige karaktertrin frem for kun for top, midt og bund.

Taksonomiske niveauer bruges i vid udstrækning i bedømmelsen

Ud over den generelle trinbeskrivelse, fagets faglige mål og de fagspecifikke karakterbeskrivelser fremhæver bedømmerne taksonomiske modeller som et centralt bedømmelsesværktøj i karaktergivningen. Det kan være enten Blooms taksonomi eller SOLO-taksonomien. Nogle bedømmere beskriver, at brugen af taksonomiske niveauer sammen med de faglige mål kan gøre bedømmelsen mere håndgribelig. En bedømmer siger:



De faglige mål i læreplaner er så elastikagtige. De er lige så elastikagtige som de beskrivelser, der er til karaktererne. Altså, hvornår har de opfyldt det, og hvornår har de ikke opfyldt det? Der bliver vi nødt til at tage fat i noget konkret, og de er ikke konkrete. Men vi får det til at være konkret, hvis vi tager udgangspunkt i taksonomi.

Bedømmelsesværktøjer kommer særligt i spil ved tvivl eller uenighed og understøtter ensartethed

Bedømmerne beskriver, at bedømmelsesværktøjerne særligt kommer i spil, hvis der opstår tvivl eller uenighed om bedømmelsen af en elev, og at bedømmelsesværktøjerne derfor er med til at understøtte en ensartethed i bedømmelsen. Flere bedømmere peger på, at bedømmelsesværktøjerne er vigtige ift. at skabe ensartethed i bedømmelsen og i tolkningen af læreplanen, så man ikke risikerer, at bedømmernes individuelle kæpheste inden for faget kommer til at blive styrende.

3.7.3 Bedømmere på de videregående uddannelser

Hovedpointerne ift. bedømmelsesværktøjer er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere på de videregående uddannelser. I dette afsnit vil der derfor være pointer, der går igen, men de illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Trinbeskrivelse og læringsmål anvendes aktivt

De interviewede bedømmere har alle trinbeskrivelsen med ved votering med censor. Trinbeskrivelsen tages særligt frem ved usikkerhed eller uenighed om, hvilken karakter en given studerende skal

have. En bedømmer fortæller, at han tager trinbeskrivelsen frem og læser de forskellige beskrivelser 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. Tre 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 karakter derudfra.



Jeg bruger trinbeskrivelsen meget, hvis jeg kommer i tvivl om, hvor karakteren skal ligge henne. Mange af prøverne her er skriftlige grupperapporter, hvor den studerende kommer ind til en mundtlig eksamen med en praktisk og en teoretisk del, og der skal gives en samlet karakter. Her har censor og jeg brugt den til at se, hvad karakteren egentlig står for.



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.

Derudover har bedømmerne læringsmålene for den pågældende prøve med. En af de interviewede bedømmere har for nylig bedømt en mundtlig prøve med et skriftligt element. Bedømmeren havde et skema, der blev udfyldt ved læsning af det skriftlige produkt. Der er ikke tale om et officielt dokument, men et dokument, som er udarbejdet af bedømmeren selv, og som også anvendes af andre bedømmere på uddannelsen. Dokumentet indeholder alle læringsmål, som gennemgås ved læsning af det skriftlige produkt, og derudover er der et felt til at skrive spørgsmål, der skal besvares ved den mundtlige prøve for at afgøre, om den studerende har nået det pågældende læringsmål. En anden bedømmer fortæller ligeledes, at læringsmålene er meget styrende for prøven:



Jeg har kørt meget strukturerede eksaminer. Mere, end jeg plejer at gøre, netop fordi vi er blevet pålagt de her læringsmål. Det gør, at jeg simpelthen har et skema med de fire læringsmål. Så stiller jeg spørgsmål til det første læringsmål, så til det andet, etc. Og jeg spørger aldrig om noget, der ligger uden for de her læringsmål.



Jeg holder mig til den i forskellige situationer. Der fungerer det meget godt. Netop fordi der har været så meget polemik om fejl og mangler, så synes jeg, at det har været en markant forbedring ift. 13-skalaen – ikke mindst ift. de studerende. Især ift. ensartetheden. Når man sidder og diskuterer med censor, kan man bruge dem her til at argumentere for eller imod. Det har disciplineret karaktergivningen, og det har gjort det væsentligt lettere at forklare de studerende, hvorfor de har fået den karakter, de har fået. Flertallet er ikke specielt overraskede over den karakter, de får. Nogle er glade, fordi de troede, at det var gået værre, end det var. For dem, der er skuffede, der er det her en stor hjælp, fordi man kan sige ”sådan og sådan...”. Som eksaminator bliver man tvunget til at kunne levere den forklaring, og det synes jeg, styrker de studerendes stilling ift. eksaminationssystemets magt over dem. De studerende har krav på at vide, hvorfor den beslutning er blevet truffet. Det hjælper det her til.

Bedømmerne oplever stor enighed

Bedømmerne fortæller alle i interviewene, at de oplever enighed mellem eksaminator og censor i forbindelse med karaktergivning. Det til trods er det stadig historier om uenigheder i forbindelse med karaktergivningen, som fylder i interviewene:



Jeg oplever en høj grad af enighed. Det er forholdsvis sjældent, at vi er uenige. Det er én karakter til og fra. Heldigvis kan man tale sig til rette. Jeg synes, at det er sjældent, at jeg har oplevet, at vi har været rygende uenige – og det gælder for alle slags eksaminer. Det er sket, at jeg har brugt lang tid på at argumentere frem og tilbage. Vi havde et speciale, som censor gerne ville dumpe. Det brugte jeg en del tid på at tale med ham om. Han havde gode argumenter, og jeg endte med at give mig. Det var måske også, fordi det er ufedt at dumpe specialer. Men jeg måtte indrømme, at der var alvorlige mangler. Vi havde dog en, hvor jeg var overrasket over, at han ville lade ham bestå. Men det er én ud af hundrede.



Det er jo nemmere at være mere objektiv til skriftlige prøver. Ved mundtlige eksaminer kommer det subjektive element ind. Det er også sjældent, at det mundtlige gør et projekt ringere, end det er. Skriftlige eksaminer er mere faste, der kan du ikke lave noget om. Men vi [eksaminator og censor] er samtidig mere enige om de mundtlige end om de skriftlige. Hvis vi har ti oppe til en skriftlig, vil vi typisk kun være enige om fire-fem af dem.

4 Understøttelse af bedømmerne

4.1 Baggrund og undersøgelsesspørgsmål

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 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.

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.

I dette kapitel sætter vi fokus på følgende undersøgelsesspørgsmål:

- Hvordan klædes bedømmere på til at anvende bedømmelsesværktøjerne?

4.2 Survey blandt bedømmere på grundskoleområdet

Tabel 4.1 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn om, hvad de inden for de sidste to år har gjort/deltaget i som forberedelse til at foretage bedømmelser. I det følgende oplystes andelen, der har svaret, at de har gjort/deltaget i en aktivitet:

- 86 % har "Læst vejledninger om bedømmelsespraksis".
- 83 % har "Læst bekendtgørelser/lovtekster om bedømmelsespraksis".

- 51 % har "Deltaget i kursus om bedømmelsespraksis".
- 49 % har "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning".
- 27 % har "Diskuteret karakterafgivelse med en fra ledelsen".
- 25 % har "Overværet kollegaers prøveafholdelse".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja ved aktiviteten "Deltaget i kursus om bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere i dansk (41 %) og bedømmere i engelsk (59 %), såvel som mellem bedømmere fra dansk (41 %) og bedømmere i matematik (54 %).
- Andelen, der har svaret ja ved aktiviteten "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning", er statistisk signifikant forskellig mellem bedømmere i dansk (47 %) og bedømmere i engelsk (63 %).

TABEL 4.1

Har du inden for de sidste 2 år:

	Ja	Nej	Total
Deltaget i kursus om bedømmelsespraksis?	51 % (341)	49 % (324)	100 % (665)
Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning?	49 % (328)	51 % (337)	100 % (665)
Læst vejledninger om bedømmelsespraksis?	86 % (570)	14 % (95)	100 % (665)
Læst bekendtgørelser/lovtekster om bedømmelsespraksis?	83 % (549)	17 % (116)	100 % (665)
Diskuteret karakterafgivelse med en fra ledelsen?	27 % (178)	73 % (487)	100 % (665)
Overværet kollegaers prøveafholdelse?	25 % (164)	75 % (501)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 4.2 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der er enige (enten "Meget enig" eller "Enig") for hvert udsagn:

- 89 % er enige i udsagnet "Jeg føler mig godt klædt på til at bedømme til prøverne".
- 40 % er enige i udsagnet "Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution".
- 33 % er enige i udsagnet "Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Jeg føler mig godt klædt på til at bedømme til prøverne”, er statistisk signifikant forskellig mellem bedømmere fra dansk (84 %) og matematik (91 %).

TABEL 4.2

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg føler mig godt klædt på til at bedømme til prøverne	46 % (306)	43 % (289)	7 % (48)	3 % (19)	0 % (2)	0 % (1)	100 % (665)
Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution	10 % (69)	30 % (202)	31 % (205)	18 % (117)	7 % (45)	4 % (27)	100 % (665)
Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne	11 % (74)	22 % (148)	33 % (217)	19 % (128)	11 % (75)	3 % (23)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskaracter/standpunktskaracter på baggrund af 7-trins-skalaen.

4.3 Survey blandt bedømmere på de gymnasiale uddannelser

Tabel 4.3 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om, hvad de inden for de sidste to år har gjort/deltaget i som forberedelse til at foretage bedømmelser. I det følgende oplistes andelen, der har svaret, at de har gjort/deltaget i en aktivitet:

- 73 % har ”Læst bekendtgørelser/lovtekster om bedømmelsespraksis”.
- 72 % har ”Læst vejledninger om bedømmelsespraksis”.
- 32 % har ”Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning”.
- 31 % har ”Diskuteret karakterafgivelse med en fra ledelsen”.
- 30 % har ”Overværet kollegaers prøveafholdelse”.
- 14 % har ”Deltaget i kursus om bedømmelsespraksis”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja ved aktiviteten "Deltaget i kursus om bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere i dansk (17 %) og bedømmere i matematik (8 %), såvel som mellem bedømmere i engelsk (17 %) og bedømmere i matematik (8 %).
- Andelen, der har svaret ja ved aktiviteten "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning", er statistisk signifikant forskellig mellem bedømmere i dansk (36 %) og bedømmere i matematik (28 %).
- Andelen, der har svaret ja ved aktiviteten "Læst vejledninger om bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere i dansk (76 %) og bedømmere i matematik (65 %), såvel som mellem bedømmere i engelsk (72 %) og bedømmere i matematik (65 %).
- Andelen, der har svaret ja ved aktiviteten "Diskuteret karakterafgivelse med en fra ledelsen", er statistisk signifikant forskellig mellem bedømmere i dansk (34 %) og bedømmere i matematik (22 %), såvel som mellem bedømmere i engelsk (32 %) og bedømmere i matematik (22 %).
- Andelen, der har svaret ja ved aktiviteten "Overværet kollegaers prøveafholdelse", er statistisk signifikant forskellig mellem bedømmere fra dansk (25 %) og matematik (39 %), såvel som mellem bedømmere i engelsk (26 %) og bedømmere i matematik (39 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hhx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja ved at have "Deltaget i kursus om bedømmerpraksis", er statistisk signifikant forskellig mellem bedømmere fra stx (12 %) og bedømmere på hhx (19 %).
- Andelen, der har svaret ja ved at have "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning", er statistisk signifikant forskellig mellem bedømmere fra stx (35 %) og bedømmere på hf (22 %), såvel som mellem bedømmere på hf (22 %) og bedømmere på hhx (35 %).
- Andelen, der har svaret ja ved at have "Diskuteret karakterafgivelse med en fra ledelsen", er statistisk signifikant forskellig mellem bedømmere fra stx (31 %) og bedømmere på hf (20 %), såvel som mellem bedømmere på hf (20 %) og bedømmere på hhx (38 %), såvel som mellem bedømmere på hf (20 %) og bedømmere på htx (38 %).
- Andelen, der har svaret ja ved at have "Overværet kollegaers prøveafholdelse", er statistisk signifikant forskellig mellem bedømmere fra stx (33 %) og bedømmere på hhx (22 %).

TABEL 4.3

Har du inden for de sidste to år:

	Ja	Nej	Total
Deltaget i kursus om bedømmelsespraksis?	14 % (164)	86 % (1023)	100 % (1187)
Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning?	32 % (379)	68 % (808)	100 % (1187)
Læst vejledninger om bedømmelsespraksis?	72 % (855)	28 % (332)	100 % (1187)
Læst bekendtgørelser/lovtekster om bedømmelsespraksis?	73 % (870)	27 % (317)	100 % (1187)

	Ja	Nej	Total
Diskuteret karakterafgivelse med en fra ledelsen?	31 % (372)	69 % (815)	100 % (1187)
Overværet kollegaers prøveafholdelse?	30 % (355)	70 % (832)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskaracter/standpunktskaracter på baggrund af 7-trins-skalaen.

Tabel 4.4 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der er enige (enten "Meget enig" eller "Enig") for hvert udsagn:

- 90 % er enige i udsagnet "Jeg føler mig godt klædt på til at bedømme til prøverne".
- 32 % er enige i udsagnet "Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution".
- 18 % er enige i udsagnet "Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution", er statistisk signifikant forskellig mellem bedømmere i dansk (40 %) og bedømmere i engelsk (32 %), såvel som mellem bedømmere i dansk (40 %) og bedømmere i matematik (19 %), såvel som mellem bedømmere i engelsk (32 %) og bedømmere i matematik (19 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne", er statistisk signifikant forskellig mellem bedømmere i dansk (23 %) og bedømmere i matematik (12 %), såvel som mellem bedømmere i engelsk (18 %) og bedømmere i matematik (12 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Jeg føler mig godt klædt på til at bedømme til prøverne", er statistisk signifikant forskellig mellem bedømmere fra htx (73 %) og bedømmere fra hhx (92 %), såvel som mellem bedømmere på htx (73 %) og bedømmere på hf (89 %), såvel som mellem bedømmere på htx (73 %) og bedømmere på stx (91 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution", er statistisk signifikant forskellig mellem bedømmere fra stx (29 %) og bedømmere fra hhx (49 %), såvel som mellem bedømmere på hf (24 %) og bedømmere på hhx (49 %), såvel som mellem bedømmere på hf (24 %) og bedømmere på htx (44 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne", er statistisk signifikant forskellig mellem bedømmere fra stx (17 %) og bedømmere fra hf (11 %), såvel

som mellem bedømmere på hf (11 %) og bedømmere på hhx (30 %), såvel som mellem bedømmere på hf (11 %) og bedømmere på htx (44 %), såvel som mellem bedømmere på stx (17 %) og bedømmere på htx (44 %).

TABEL 4.4

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg føler mig godt klædt på til at bedømme til prøverne	39 % (468)	51 % (603)	7 % (88)	2 % (21)	0 % (4)	0 % (3)	100 % (1187)
Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution	6 % (68)	26 % (306)	33 % (396)	21 % (254)	10 % (116)	4 % (47)	100 % (1187)
Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne	4 % (49)	14 % (169)	31 % (366)	29 % (340)	18 % (208)	5 % (55)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

4.4 Survey blandt bedømmere på erhvervsuddannelserne

Tabel 4.5 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn om, hvad de inden for de sidste to år har gjort/deltaget i som forberedelse til at foretage bedømmelser. I det følgende oplystes andelen, der har svaret, at de har gjort/deltaget i hver aktivitet:

- 82 % har "Læst bekendtgørelser/lovtekster om bedømmelsespraksis".
- 75 % har "Læst vejledninger om bedømmelsespraksis".
- 35 % har "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning".
- 33 % har "Diskuteret karakterafgivelse med en fra ledelsen".
- 27 % har "Overværet kollegaers prøveafholdelse".
- 16 % har "Deltaget i kursus om bedømmelsespraksis".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret ja til hver aktivitet, er ikke statistisk signifikant forskellig på tværs af fag, ifølge en almindelig z-test.

TABEL 4.5

Har du inden for de sidste to år:

	Ja	Nej	Total
Deltaget i kursus om bedømmelsespraksis?	16 % (50)	84 % (264)	100 % (314)
Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning?	35 % (109)	65 % (205)	100 % (314)
Læst vejledninger om bedømmelsespraksis?	75 % (235)	25 % (79)	100 % (314)
Læst bekendtgørelser/lovtekster om bedømmelsespraksis?	82 % (258)	18 % (56)	100 % (314)
Diskuteret karakterafgivelse med en fra ledelsen?	33 % (104)	67 % (210)	100 % (314)
Overværet kollegaers prøveafholdelse?	27 % (84)	73 % (230)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

Tabel 4.6 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn. I det følgende oplyses andelen, der er enige (enten "Meget enig" eller "Enig") for hvert udsagn:

- 84 % er enige i udsagnet "Jeg føler mig godt klædt på til at bedømme til prøverne".
- 44 % er enige i udsagnet "Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution".
- 32 % er enige i udsagnet "Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret "Meget enig" eller "Enig" til hvert udsagn, er ikke statistisk signifikant forskellig på tværs af fag, ifølge en almindelig z-test.

TABEL 4.6

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg føler mig godt klædt på til at bedømme til prøverne	33 % (103)	51 % (159)	13 % (40)	3 % (10)	0 % (1)	0 % (1)	100 % (314)
Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution	13 % (40)	31 % (98)	37 % (116)	12 % (39)	3 % (8)	4 % (13)	100 % (314)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne	8 % (26)	24 % (76)	39 % (124)	18 % (57)	7 % (22)	3 % (9)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet til bedømmere, der har afgivet en årskarakter/standpunktskarakter på baggrund af 7-trins-skalaen.

4.5 Survey blandt bedømmere på de videregående uddannelser

Tabel 4.7 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn om, hvad de inden for de sidste to år har gjort/deltaget i som forberedelse til at foretage bedømmelser. I det følgende oplystes andelen, der har svaret, at de har gjort/deltaget i hver enkelt aktivitet:

- 66 % har "Læst vejledninger om bedømmelsespraksis".
- 57 % har "Læst bekendtgørelser/lovtekster om bedømmelsespraksis".
- 44 % har "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning".
- 36 % har "Overværet kollegaers prøveafholdelse".
- 34 % har: "Diskuteret karakterafgivelse med en fra ledelsen".
- 11 % har "Deltaget i kursus om bedømmelsespraksis".

Variationer i svar mellem bedømmerne fra hhv. universitetsuddannelserne, professionsbacheloruddannelserne og erhvervsakademiuddannelserne samt de kunstneriske og maritime uddannelser er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja ved aktiviteten "Deltaget i kursus om bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (12 %) og bedømmere fra erhvervsakademiuddannelserne (7 %).
- Andelen, der har svaret ja ved aktiviteten "Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning", er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (46 %) og bedømmere fra professionsbacheloruddannelserne (39 %) såvel som mellem bedømmere fra professionsbacheloruddannelserne (39 %) og bedømmere fra erhvervsakademiuddannelserne (49 %).
- Andelen, der har svaret ja ved aktiviteten "Læst bekendtgørelser/lovtekster om bedømmelsespraksis", er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (52 %) og bedømmere fra professionsbacheloruddannelserne (68 %) såvel som mellem bedømmere fra professionsbacheloruddannelserne (68 %) og bedømmere fra erhvervsakademiuddannelserne (58 %).

- Andelen, der har svaret ja ved aktiviteten ”Diskuteret karakterafgivelse med en fra ledelsen”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (33 %) og bedømmere fra erhvervsakademiuddannelserne (48 %) såvel som mellem bedømmere fra professionsbacheloruddannelserne (29 %) og bedømmere fra erhvervsakademiuddannelserne (48 %).
- Andelen, der har svaret ja ved aktiviteten ”Overværet kollegaers prøveafholdelse”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (36 %) og bedømmere fra professionsbacheloruddannelserne (31 %), mellem bedømmere fra universitetsuddannelserne (36 %) og bedømmere fra erhvervsakademiuddannelserne (45 %) samt mellem bedømmere fra professionsbacheloruddannelserne (31 %) og bedømmere fra erhvervsakademiuddannelserne (45 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 4.7

Har du inden for de sidste to år:

	Ja	Nej	Total
Deltaget i kursus om bedømmelsespraksis?	11 % (328)	89 % (2.618)	100 % (2.946)
Deltaget i en fast praksis, hvor karaktergivning diskuteres mellem kollegaer som led i forberedelse af karaktergivning?	44 % (1.294)	56 % (1.652)	100 % (2.946)
Læst vejledninger om bedømmelsespraksis?	66 % (1.933)	34 % (1.013)	100 % (2.946)
Læst bekendtgørelser/lovtekster om bedømmelsespraksis?	57 % (1.673)	43 % (1.273)	100 % (2.946)
Diskuteret karakterafgivelse med en fra ledelsen?	34 % (988)	66 % (1.958)	100 % (2.946)
Overværet kollegaers prøveafholdelse?	36 % (1.048)	64 % (1.898)	100 % (2.946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 4.8 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der er enige (enten ”Meget enig” eller ”Enig”), for hvert udsagn:

- 87 % er enige i udsagnet ”Jeg føler mig godt klædt på til at bedømme til prøverne”.
- 35 % er enige i udsagnet ”Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution”.
- 27 % er enige i udsagnet ”Der er behov for at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne”.

Variationer i svar mellem bedømmerne fra hhv. universitetsuddannelserne, professionsbacheloruddannelserne og erhvervsakademiuddannelserne samt de kunstneriske og maritime uddannelser er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ””””Meget enig” eller ””””Enig” ved udsagnet ”Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (32 %) og bedømmere fra professionsbacheloruddannelserne (42 %).
- Andelen, der har svaret ””””Meget enig” eller ””””Enig” ved udsagnet ”Der er behov for, at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis til prøverne”, er statistisk signifikant forskellig mellem bedømmere fra universitetsuddannelserne (23 %) og bedømmere fra professionsbacheloruddannelserne (35 %) såvel som mellem bedømmere fra universitetsuddannelserne (23 %) og bedømmere fra erhvervsakademiuddannelserne (33 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 4.8

Hvor enig eller uenig er du i følgende udsagn?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Jeg føler mig godt klædt på til at bedømme til prøverne	34 % (991)	53 % (1.564)	10 % (305)	2 % (72)	0 % (6)	0 % (8)	100 % (2.946)
Det vil være nyttigt at have mere fokus på bedømmelsespraksis på min skole/institution	6 % (189)	29 % (846)	33 % (958)	19 % (549)	5 % (140)	9 % (265)	100 % (2.946)
Der er behov for, at ledelsen i højere grad sætter fokus på arbejdet med bedømmelsespraksis i forbindelse med prøverne	5 % (157)	22 % (648)	32 % (944)	23 % (690)	8 % (239)	9 % (268)	100 % (2.946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

4.6 Fokusgruppeinterviews med bedømmere

I det følgende beskrives bedømmernes erfaringer, og hvordan de som bedømmere bliver understøttet og klædt på til karaktergivning, som det er kommet til udtryk i fokusgruppeinterviewene.

4.6.1 Bedømmere på grundskoleområdet

Beskikkede censorer føler sig godt klædt på til karaktergivning, mens der efterspørges en højere grad af systematik ift. at klæde nye bedømmere på

De bedømmere, der er beskikkede censorer, føler sig generelt godt klædt på til karaktergivning. For beskikkede censorer er der en systematik ift. løbende at blive klædt på til bedømmelse og karaktergivning, fordi censorerne hvert år er på webinar eller kursus.

For de bedømmere, der ikke er beskikkede censorer, herunder nye bedømmere, er det mere forskelligt, hvor meget de enkelte skoler har systematiseret og prioriteret introduktion til skalaen og til karaktergivning. Bedømmerne fortæller, at der primært har været fokus på det og en systematik med hensyn til det, efter at 7-trins-skalaen blev indført. En bedømmer fortæller:



Jeg føler mig rimeligt godt klædt på. Der var masser af møder om det på min gamle arbejdsplads, da vi overgik til den nye skala. Dengang blev elever filmet til prøveeksaminer, og det blev brugt som materiale til seminarer om karaktergivning.

Men for de nye bedømmere, der ikke var med, dengang 7-trins-skalaen blev indført, beskrives introduktion til skalaen og til karaktergivning generelt som noget, man som ny bedømmer selv skal op-
søge, og som sker ad hoc frem for systematisk. På den baggrund oplever en del af de nye bedømmere, at de ikke er tilstrækkelig klædt på til karaktergivning, og efterlyser en højere grad af systematik med hensyn til det. De nye bedømmere efterlyser dels, at bedømmelse og karaktergivning får lov at fylde mere i løbet af læreruddannelsen, dels at det bliver sat mere systematisk på dagsordenen ude på skolerne. En bedømmer siger:



Jeg har ikke haft noget om det på seminariet. Jeg tænker, at det er fjollet, at fire års seminarieuddannelse har så lidt fokus på det.

En anden bedømmer siger:



Jeg er ikke blevet præsenteret for hverken kurser eller vurderingsværktøjer ud over det, jeg har lært ved at snakke med kollegaer. Jeg er ikke blevet klædt på – hverken på uddannelsen eller på arbejdet. Ingen har sat sig ned og gennemgået kravene med mig.

Nogle bedømmere mener, at der generelt er behov for en højere grad af systematik med hensyn til at blive introduceret til bedømmelse og også behov for, at erfarne bedømmere løbende bliver opdateret ift. karaktergivning. En bedømmer udtaler:



Det mener jeg, er ledelsens ansvar at få iværksat. Jeg mener ikke, at vi er gode nok til det. For selvom man på et tidspunkt er en garvet udskolingslærer, så starter man jo ikke sådan. Derudover er der jo også kollegaer, der ikke følger med i opdateringer ift. prøver og krav.

Mulige initiativer ift. at klæde særligt nye bedømmere bedre på

Bedømmerne peger på en række ting, som de oplever, klæder dem på til bedømmelse, og som der derfor kunne skabes en højere grad af systematik med hensyn til. Det drejer sig dels om kollegial sparring og diskussion med kollegaer, dels om fælles bedømmelsessituationer, hvor man sammen med kollegaer bedømmer og diskuterer konkrete opgaver eller mundtlige præstationer, fx ved at sidde med til en prøve eller optage en prøve på video. Derudover foreslår nogle bedømmere, at der sikres en systematik med hensyn til, at alle bedømmere kommer på kursus, ligesom det er tilfældet for de beskikkede censorer. En bedømmer siger:



Det giver en ro i maven, at man ved, hvordan man skal bedømme. Jeg synes, at alle burde komme afsted [på kursus].

Endelig oplever bedømmerne også, at de drøftelser om karaktergivning, de har med censor til mundtlige prøver, gør dem til bedre bedømmere, også med hensyn til standpunktskarakterer. Konkrete censor- og eksaminatorerfaringer udvikler således også bedømmernes karaktergivningskompetencer på tværs af skoler. Derfor peger flere bedømmere også på, at de oplever det som problematisk, at man i dag kun er én bedømmer om at bedømme skriftlige eksaminer. Det opleves som

problematiske ift. at sikre ensartethed i karaktergivning, og fordi en sådan fælles bedømmelsessituation opleves som en af de lejligheder, hvor bedømmerne udvikler og styrker deres bedømmelsespraksis.

4.6.2 Bedømmere på det gymnasiale område

Erfarne bedømmere føler sig godt klædt på til karaktergivning

Bedømmerne i fokusgrupperne føler sig generelt godt klædt på til at bedømme eleverne og give karakterer ud fra 7-trins-skalaen. Værd at bemærke er det dog, at deltagerne i fokusgrupperne overvejende har været bedømmere med mange års erfaring, og denne erfaring har væsentlig betydning for, at de føler sig godt klædt på. De mindre erfarne bedømmere såvel som de erfarne bedømmere peger på, at man som ny bedømmer ikke nødvendigvis oplever at være tilstrækkelig klædt på til bedømmelse og karaktergivning.

Det er først og fremmest erfaringen, der klæder bedømmerne på

Bedømmerne beskriver, at det at give karakterer ikke er en teoretisk disciplin, og selvom det er afgørende at kende og anvende de forskellige bedømmelsesværktøjer i karaktergivningen, så er selve det at bedømme eleverne ikke noget, man kan læse sig til, men noget, der kræver praktisk erfaring at blive klædt på til. Bedømmerne beskriver, at de særligt taler med deres kollegaer i faggruppen om bedømmelse. Men der, hvor bedømmerne særligt oplever at udvikle, forfine og afstemme deres karaktergivningspraksis, er, når de bedømmer elever i fællesskab med andre bedømmere. Sådanne fælles bedømmelsessituationer kan være en egentlig prøvesituation i rollen som censor eller eksaminator, men det kan også være konstruerede fælles bedømmelsessituationer, fx i faggrupperne på skolen eller i forbindelse med pædagogikum. En bedømmer beskriver det på følgende måde: "jo mere erfaring, jo større referenceramme ift. at bedømme og finde niveauet".

Bedømmerne peger på, at det ideelle ift. at blive klædt på som bedømmer er en kombination af sparring med faggruppekollegaer og fælles bedømmelsessituationer med andre bedømmere i kombination med oplæg eller kurser, hvilket kan sikre, at den korrekte praksis videreføres. En bedømmer siger:



Man snakker med hinanden på lærerværelset og sparrer med faggruppekollegaer. Men man kan risikere at videreføre andres ikke helt korrekte praksis. Derfor er det også godt med kurser, og at det indgår i pædagogikum.

Selvom de erfarne bedømmere generelt føler sig godt klædt på til karaktergivning, så påpeger nogle af dem, at det også som erfarne bedømmere er vigtigt løbende at genbesøge bedømmelsesværktøjerne, sådan at de står klart i forbindelse med bedømmelserne, og at man hele tiden sikrer sig, at bedømmelsen sker med udgangspunkt heri. En bedømmer siger:



Man kunne som erfaren censor godt blive bedre til også løbende at kvalificere bedømmelsen i bekendtgørelsen. Det er vigtigt som erfaren bedømmer også lige at vende tilbage til bedømmelsesværktøjerne. Og ikke bare have det i baghovedet, men præsent.

Efterspørgsel efter en højere grad af systematik ift. at klæde nye bedømmere på

Bedømmernes erfaringer peger på, at det er forskelligt, hvor systematisk skolerne arbejder med at klæde nye bedømmere på til karaktergivning, men at der generelt godt kunne være en højere grad af systematik med hensyn til dette. De nye bedømmere oplever at møde introduktion til karaktergivning i forbindelse med pædagogikum og på nogle skoler, ved at lærerne i fællesskab øver og diskuterer bedømmelse. Men generelt beskrives det som værende meget op til den enkelte be-

dømmer eller faggruppe at tage initiativ til at drøfte, hvordan man bedømmer i faget, og der efterlyses på den baggrund en højere grad af systematik med hensyn til at klæde nye bedømmere på til karaktergivning ved hjælp af drøftelser og fælles bedømmelsessituationer med fagfæller. En bedømmer fortæller, hvordan han selv har taget initiativ til at drøfte bedømmelse med kollegaer:



Som ny lærer er jeg da overhovedet ikke blevet klædt på. Der var jo ikke nogen, der sagde ””du skal arbejde med det sådan her””. Man bliver bare kastet ind i det. Men jeg har selv taget initiativ til at snakke med kollegaer. Men til at starte med rettede jeg efter universitetsniveau.

Ved prøver sker bedømmelsen i fællesskab med en censor eller eksaminator, hvorfor den nye bedømmer har mulighed for at lære af mere erfarne bedømmere, som beskrevet i det følgende citat. Men ved standpunktskarakterer og årskarakterer er læreren alene om at give karakterer, og der er derfor ikke samme mulighed for nye bedømmere for at afstemme bedømmelsen med mere erfarne bedømmere, medmindre de selv opsøger kollegaer.



Jeg husker tydeligt, den første gang jeg skulle have elever op mundtligt. Jeg havde ingen erfaring og måtte blankt erkende over for censor, at jeg ikke vidste det. Så brugte vi de to dage på, at jeg meldte ud først, og han så korrigerede mig. Så vi brugte de to dage på, at han lærte mig op. Men på den måde kan det godt blive lidt fejlbehæftet i starten.

4.6.3 Bedømmere på de videregående uddannelser

Erfarne bedømmere føler sig godt klædt på til karaktergivning

Bedømmerne i fokusgrupperne føler sig generelt godt klædt på til at bedømme eleverne og give karakterer ud fra 7-trins-skalaen. Værd at bemærke er det dog, at deltagerne i fokusgrupperne overvejende har været bedømmere med mange års erfaring, og denne erfaring har væsentlig betydning for, at de føler sig godt klædt på. De mindre erfarne bedømmere såvel som de erfarne bedømmere peger på, at man som ny bedømmer ikke nødvendigvis oplever at være tilstrækkelig klædt på til bedømmelse og karaktergivning.

Størst behov hos nyansatte, udenlandske undervisere og aftagercensorer

Ud over at nyansatte som nye bedømmere har et særligt behov for at blive klædt på til bedømmelse og karaktergivning, så vurderer bedømmerne, at de, som derudover har størst behov for understøttelse, er udenlandske undervisere og aftagercensorer.

Som beskrevet i kapitel 6 fortæller bedømmerne, at når der er uenighed om en bedømmelse, så kan det ofte være med aftagercensorer, der ikke har deres daglige gang på uddannelsen. Bedømmerne understreger, at aftagercensorer naturligvis bruges, fordi de har et andet blik på de studerende, end deres undervisere har, men at det samtidig er afgørende, at censorerne bedømmer de studerende korrekt ud fra skala, læringsmål og niveau. Nogle bedømmere fortæller, at aftagercensorer ikke nødvendigvis er klædt tilstrækkelig på til dette. En bedømmer siger:



Eksterne censorer, de har de erhvervsmæssige kompetencer, men har de kompetencerne til karaktergivning? Der tror jeg godt, man kunne gøre noget.

Derudover peger flere bedømmere også på, at der er et stort behov for, at udenlandske undervisere bliver klædt på til at bedømme ud fra den danske karakterskala. En bedømmer siger: ”Det største problem er udenlandske undervisere, der ikke kender den danske skala og ikke har mavefornemmelsen for karaktererne.”

Forskelligt, hvor systematisk bedømmere klædes på til karaktergivning

Af fokusgruppeinterviewene med bedømmere fremgår det, at der er stor forskel på, hvor systematisk der bliver arbejdet med at klæde bedømmere på til karaktergivning på de forskellige uddannelsesinstitutioner. Nogle bedømmere forklarer, at der fra institutionens side arbejdes systematisk med at udvikle bedømmernes kompetencer i relation til bedømmelse og karaktergivning. Aktiviteterne til udvikling og understøttelse af bedømmelsespraksis kan tage forskellige former. De kan tage form af en formaliseret praksis som en del af adjunktuddannelsen eller en mere uformel, dog systematiseret praksis med sparring og supervision, hvor uerfarne bedømmere får sidemandsoplæring af en erfaren bedømmer. En bedømmer fortæller:



Vi har fået en kompetenceudviklingsplan på adjunktuddannelsen, og der er en masse tilbud ift., hvordan man vurderer. Der er blevet gjort meget for at ensrette forståelsen af skalaen og de didaktiske undervisningsfaglige mål for både adjunkter og lektorer.

På tværs af bedømmere går det generelt igen, at noget af det, der giver rigtig meget ift. at udvikle og skærpe sin bedømmelsespraksis, er fælles bedømmelsessituationer med andre bedømmere, hvor man sammen drøfter bedømmelse af helt konkrete opgaver eller præstationer.

På andre institutioner fortæller bedømmerne, at der ikke på samme vis er en systematik og et tilbud i relation til at blive klædt på til bedømmelse, men at det er noget, man i højere grad selv skal opsøge. Nogle steder eksisterer der tilbud, som man selv skal opsøge, mens bedømmerne andre steder ikke oplever, at der eksisterer tilbud. To bedømmere fortæller:



Jeg syntes i begyndelsen ikke, at jeg var klar. Der bad jeg om at følge med en vejleder til eksamen og bedømmelsen – uden at sige noget. Det var på eget initiativ. Jeg sørgede selv for det. Det var ikke sat i system.



Jeg tror, at man med tiden er blevet bedre til det. I dag har man denne her mentorordning. Det kan godt være, at man skal være lidt opsøgende, men den findes.

En del af bedømmerne efterlyser derfor en højere grad af systematik med hensyn til, at bedømmerne bliver klædt på på deres institution, for når det i høj grad er op til den enkelte bedømmer selv at opsøge, så indebærer det en risiko for, at bedømmerne ikke alle har lige veludviklede kompetencer i relation til bedømmelse. En bedømmer siger:



Mine kompetencer ift. karaktergivning er erfaringsbaserede. Jeg synes ikke, at min institution understøtter min karaktergivning. Man klædes ikke systematisk på hos os. Det er noget, man lærer med tiden. Jeg ville gerne have det som en del af en kompetenceudviklingsplan, som vi drøftede hvert år. Vi drøfter det indbyrdes underviserne imellem, men jeg vil gerne have det mere systematisk.

5 Kvalitetssikring

5.1 Baggrund og undersøgelsesspørgsmål

Karakterkommissionen beskrev, at der skulle ske en kvalitetssikring af karaktergivningingen ved at koordinere målene for faget, prøvens tilrettelæggelse og den population, der skal bestå prøven. Det skulle understøtte, at de beståede 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.

Selvom 7-trins-skalaen er implementeret ens på hhv. Undervisningsministeriets og Uddannelses- og Forskningsministeriets område, er der stor forskel på, hvordan der arbejdes med kvalitetssikring af karaktergivningingen 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 svendepøver 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.

I dette kapitel sætter vi fokus på følgende undersøgelsesspørgsmål:

- Bliver der gennemført særlige tiltag for at sikre, at karaktererne følger den forventede fordeling?
- Hvordan sikres en ensartet bedømmelse på tværs af fag/uddannelser/udbud?

5.2 Fokusgruppeinterviews med bedømmere

I det følgende beskrives bedømmernes erfaringer med ensartethed i og kvalitetssikring af karaktergivningingen, som de er kommet til udtryk i fokusgruppeinterviewene.

5.2.1 Bedømmere på grundskoleområdet

Enighed om bedømmelsen

Bedømmerne henviser i store træk til, at de oplever enighed om bedømmelsen og karaktergivningen, når de skal bedømme eleverne sammen med en ekstern eksaminator eller censor i forbindelse med prøver. Og når to bedømmere starter forskellige steder, oplever de i langt de fleste tilfælde at få talt sig frem til enighed om en karakter. En bedømmer fortæller:



Jeg synes, at selvom vi ofte starter forskellige steder, så når vi oftest frem til en fælles karakter – uden store uenigheder. Så dialogen er god og problemfri. Så selvom udgangspunktet er forskelligt, ender vi samme sted.

Bedømmerne beskriver, at når der en gang imellem opstår uenighed om bedømmelse, kan det handle om, hvordan bedømmerne vægter eller fortolker forskellige mål og krav. En bedømmer fortæller eksempelvis, at ”nogle synes, perspektivering er det vigtigste, andre kan være uenige”. En anden bedømmer beskriver, at i og med at bedømmelseskriterierne efterlader et fortolkningsrum til bedømmerne, så kan der også være forskelligartede fortolkninger i spil:



Man kan ud fra kriterierne altid argumentere for både 4 og 7. Så der er noget vurdering og tolkning, og der kan vi være forskellige. Det, vi oplever, når vi er på kursus som beskikkede censorer, er, at folk har forskellige udgangspunkter. Nogle tager udgangspunkt i fejl og mangler, nogle prøver at finde det gode.

Nogle bedømmere peger på, at der kan være uenighed oftere i nogle fag end i andre, afhængigt af hvor mange forskellige parametre eleverne skal bedømmes på i faget, og hvor konkrete parametrene er. En bedømmer fortæller eksempelvis om flere uenigheder i dansk end i engelsk, fordi eleverne skal bedømmes på færre ting i engelsk.

Bedømmerne anser det at have to bedømmere ved prøver som vigtigt

Til de skriftlige prøver er der kun én censor, der retter og bedømmer prøven. Netop fordi der ligger et element af fortolkning af mål og krav i bedømmelsen af eleverne, og fordi forskellige bedømmere kan vægte og fortolke krav og mål forskelligartet, ville bedømmerne finde det mere hensigtsmæssigt for både sig selv og eleverne, hvis der var to bedømmere til alle prøver. En bedømmer siger: ”For elevens skyld bør man være flere bedømmere.” En anden siger: ”Der er jo kun én, der skal rette de skriftlige prøver nu. Det betyder virkelig meget for, hvordan vurderingen er.” Mens en tredje bedømmer fortæller, at muligheden for at få to øjne på en præstation og muligheden for drøftelser med en anden bedømmer er vigtige:



Hvis jeg er meget i tvivl, kan jeg godt tage bedømmelseskemaerne frem. Men jeg synes, at det ville være rart at kunne få to øjne mere på og have muligheden for at snakke med læreren. Man retter jo alligevel virkelig mange.

Det kan være udfordrende ikke at tage forhåndskendskab til eleven med ind i eksamensrummet

Nogle bedømmere forklarer, at eksaminators forhåndskendskab til eleverne også nogle gange kan spille ind i bedømmelsesdiskussionen. En bedømmer siger:



Vi er mennesker, så vi tager selvfølgelig vores egne oplevelser og følelser med ind i eksamenssituationen. Men det er også derfor, det er så vigtigt at have en censor med, som i højere grad kan bedømme udelukkende på baggrund af præstationen i eksamenssituationen. Ja, det kan føles uretfærdigt at give en god elev en dårlig karakter til eksamen. Men det er også derfor, der findes årskarakterer.

En bedømmer fortæller, at man nogle gange skal sørge for at holde fokus på at bedømme eleven ud fra det, eleven præsterer i den konkrete situation:



En sætning, jeg ofte bruger i bedømmelsen er: ”Jamen jeg skal jo kun vurdere det, jeg ser i dag.” Det kan være svært at lægge følelserne væk, når man har med egne elever at gøre. Det er svært, men man når altid til enighed.

Ved uenigheder eller uklarheder kommer bedømmelsesværktøjerne i brug og understøtter ensartethed

Ved uenighed tager bedømmerne ofte fat i bedømmelsesværktøjernes beskrivelser af mål og krav. Enten for gennem forskellige fortolkninger af de faglige mål og krav at søge afklaring på, hvordan de nøjagtige formuleringer i de gældende bedømmelsesværktøjer lyder. Eller for at sikre, at begge bedømmere har den faglige bedømmelse for øje, og at bedømmelsen sker absolut frem for ved relativ sammenligning af elever. En bedømmer fortæller:



Jeg oplever ikke masser af uenigheder, men der er altid nogle. Så sidder man og argumenterer frem og tilbage og henviser til bestemte steder i eksamen eller til evalueringsskemaer. I de situationer er det rart at have nogle mål og krav, som man kan bedømme ud fra.

5.2.2 Bedømmere på det gymnasiale område

Enighed om bedømmelsen

Bedømmerne henviser i store træk til, at de oplever stor enighed om bedømmelsen og karaktergivningen, når de skal bedømme eleverne sammen med en ekstern eksaminator eller censor i forbindelse med prøver. Det gælder både til mundtlig og til skriftlig eksamen. Er der uenighed, er der i langt de fleste tilfælde tale om et karaktertrins forskel, og bedømmerne fortæller, at bedømmelsen er et konstruktivt samarbejde mellem de to bedømmere om at finde frem til den rette bedømmelse.

Bedømmerne beskriver, at når der opstår uenighed om bedømmelsen, handler det som oftest om, at de to bedømmere vægter forskellige elementer af faget forskelligt i den samlede helhedsvurdering af eleven. Det kan handle om, at bedømmerne vægter forskellige faglige mål, kompetencer eller genstandsfelter i faget forskelligt. Det kan i et sprogfag som engelsk eksempelvis handle om, hvor meget det sproglige skal vægte ift. det analytiske.

Tværfaglig bedømmelse, eksempelvis i forbindelse med større skriftlige opgaver, fremhæves som noget, der kan være en udfordrende bedømmelsessituation, hvor der kan opstå uenighed bedømmerne imellem. Det kan igen dreje sig om vægtning, eksempelvis hvor meget hhv. det ene og det andet fag, eleven skriver i, samt det tværfaglige element skal vægtes i bedømmelsen. Men det kan også dreje sig om, at bedømmerne reelt kun kan vurdere den del af opgaven, der omhandler eget fag, og at det derfor kan være svært at have en fælles samtale om opgaven.

Hvor konkrete bedømmelseskriterierne er, afhænger af fag og opgavetyper

Bedømmerne beskriver, at det har betydning for ensartetheden i bedømmelsen, hvor konkrete bedømmelseskriterierne er i faget eller ift. den specifikke opgave. Fagenes forskellige karakter betyder, at bedømmelseskriterierne er lettere at konkretisere og sætte på formel i nogle fag end i andre fag. Som to modpoler i den henseende fremhæves skriftlig matematik og dansk. Hvor bedømmelsen i skriftlig matematik sker ved pointgivning i forbindelse med hver enkelt delopgave og en fælles omsætningstabel fra point til karakter, så beskriver bedømmerne, at der i dansk er en bred vifte af parametre i spil i bedømmelsen, som ikke på samme måde lader sig sætte på formel. Men hvor konkrete bedømmelseskriterierne er, afhænger også af den konkrete opgavetype eller, i tilfældet dansk, genre, hvor bedømmerne fremhæver, at særligt essayet er en genre, hvor bedømmelseskriterierne er uklare, og der derfor kan være væsentlig uenighed om bedømmelsen, mens bedømmelseskriterierne for de øvrige genrer i dansk er klarere, og ensartetheden i bedømmelsen langt større.

Censorerfaringer og bedømmelsesværktøjer understøtter ensartet bedømmelse

Bedømmerne peger på, at konkrete erfaringer med fælles bedømmelse af en elev sammen med en anden bedømmer spiller en stor rolle ift. at sikre ensartethed i bedømmelsen. Det gælder i rollen som eksaminator for egne elever ved prøver med ekstern censur eller ved årsprøver med intern censur, men bedømmerne peger særligt på censorerfaringer både som mundtlig og som skriftlig censor som givende ift. at udvikle en fælles forståelse og ensartethed i bedømmelsen – også på tværs af skoler. Bedømmerne beskriver det som centralt at komme ud som censor jævnligt og herunder at komme rundt i landet som censor på skoler og for elevgrupper, der ikke ligner ens egen, for at sikre ensartethed på tværs. En bedømmer siger: ”Det med at sidde med nogle opgaver fra hele landet som skriftlig censor, det er en slags efteruddannelse. Det samme gælder, når man er mundtlig censor.” Men en del af bedømmerne fortæller om en tendens til, at de i mindre grad end tidligere kommer rundt i landet som censor, men ofte er censor i deres geografiske nærhed og dermed ofte på skoler, der ligner den, de selv kommer fra. En bedømmer siger:



Jeg synes, at der er en tendens til, at man oftest er censor på steder, som ligger tæt på, hvor man selv bor. Jeg tror, det kan betyde noget på sigt, at man ikke kommer lidt mere rundt omkring i landet. Så er man på nogle skoler, der minder meget om ens egen ift. elevtyper og miljø. Jeg ved ikke, om det er besparelser, men jeg tror, at det kunne sikre en større ensartethed, hvis man kom på tværs af landet.

Derudover peger bedømmerne også på, at de fælles bedømmelsesværktøjer såsom trinbeskrivelsen, de fagspecifikke karakterbeskrivelser, de faglige mål og bedømmelseskriterierne spiller en central rolle ift. ensartethed i bedømmelsen, fordi de fungerer som en fælles rettesnor, man kan referere til ved uenighed eller tvivlsspørgsmål.

Endelig nævner bedømmerne også, at fagkonsulentens oplæg forud for den skriftlige censur understøtter en ensartethed i den skriftlige bedømmelse. Ligesom den forcensur, der finder sted i matematik, også udstikker nogle retningslinjer for, hvordan opgaverne skal vurderes.

Bedømmerne beskriver, at der er størst fokus på at understøtte ensartethed i bedømmelsen ved prøver, mens der er mindre fokus på at understøtte ensartethed i bedømmelsen ved løbende standpunkts- og årskarakterer.

Ved uenigheder eller uklarheder kommer bedømmelsesværktøjerne i brug og understøtter ensartethed

Ved uenighed tager bedømmerne ofte fat i bedømmelsesværktøjernes beskrivelser af mål og krav. Enten ved gennem forskellig fortolkning af de faglige mål og krav at søge afklaring på, hvordan de nøjagtige ord i de gældende bedømmelsesværktøjer lyder. Eller for at sikre, at begge bedømmere

har den faglige bedømmelse for øje, og at bedømmelsen sker absolut frem for ved relativ sammenligning af elever.

5.2.3 Bedømmere på de videregående uddannelser

Bedømmerne oplever generelt enighed om bedømmelsen

Bedømmerne oplever generelt enighed om bedømmelsen og karaktergivningen, når de skal bedømme eleverne sammen med en ekstern eksaminator eller censor i forbindelse med prøver. Når der er uenighed, er det sjældent med mere end én karakters forskel. Bedømmerne beskriver, at det at starte ud forskelligt, men tale sig til enighed, ikke nødvendigvis skal opfattes som en konflikt, men derimod som en sund diskussion, hvor man i fællesskab finder frem til den mest korrekte bedømmelse.

Når to bedømmere starter forskellige steder på skalaen, oplever de i langt de fleste tilfælde, at uenigheden skyldes forskellig fortolkning af læringsmålene eller en forskellig vægtning af eksamens forskellige dele. Oftest lykkes det at få talt sig frem til enighed om en karakter. En bedømmer fortæller:



Nogle gange sker der det, at man skal have skudt sig ind på hinanden. Især hvad angår folk, der er væk fra den daglige gang på uddannelserne. Der sidder folk nogle gange med gamle forestillinger. Med de todelte prøver kan der også være forskel i vægtningen af det skriftlige ift. det mundtlige.

Censors baggrund har betydning for bedømmelsen

Bedømmerne vurderer, at censors baggrund har betydning for graden af enighed i bedømmelsen. Bedømmerne fortæller, at der er en tendens til større uenighed, når der er tale om en aftagercensor eller en censor fra et andet fagområde end eksaminators.

Når der er uenigheder med aftagercensorer, kan det, ifølge bedømmerne, blandt andet skyldes, at de ikke har samme indblik i uddannelsen, eller at de har et helt andet fokus i deres bedømmelse. Bedømmerne fortæller blandt andet, at de sommetider oplever, at aftagercensorerne i højere grad har fokus på projekternes praksisanvendelighed eller den studerendes person end på de konkrete læringsmål for faget. I de tilfælde er bedømmelsesværktøjer som trinbeskrivelsen og læringsmålbeskrivelsen særligt anvendelige.



Man skal også zoome ind på grundlaget for bedømmelsen i fællesskab. Der kan ligge en del arbejde i at forstå, hvad læringsmålene egentlig går ud på. At spore sig ind på hinanden ift., hvilke udgangspunkter man har. Jeg har en del praktikereksaminer, og der kommer praktikerne nogle gange til at bedømme ud fra, om de kunne tænke sig at få den studerende som kollega. Så må vi trække målbeskrivelserne ind. Det varierer meget, afhængigt af eksamenstype og censorer.

Bedømmerne fortæller også, at der kan opstå usikkerhed eller uenighed, når eksaminator og censor har meget forskelligartede fagligheder. Det kan enten handle om, at man som censor bliver sat til at bedømme en prøve, der ligger i yderkanten af ens egen faglighed, og derfor ikke er på hjemmebane i bedømmelsen. Men det kan også handle om, at der skal foretages en tværfaglig bedømmelse, hvor det kan være udfordrende for to bedømmere med to forskellige fagligheder at finde fælles fodslag, fordi deres perspektiver på elevens præstation tager udgangspunkt i forskellige fag.



Jeg oplever, at vi får censorer, som tænker meget anderledes end mig qua deres faglige baggrund. Der er meget stor forskel på naturvidenskabere og sociologer. Vores fag har meget stor spændvidde, og derfor kan det være svært at finde fælles fodslag blandt bedømmere.

Ved uenigheder eller uklarheder kommer bedømmelsesværktøjerne i brug og understøtter ensartethed

Ved uenighed tager bedømmerne ofte fat i bedømmelsesværktøjernes beskrivelser af mål og krav. Enten ved gennem forskellig fortolkning af læringsmål og krav at søge afklaring på, hvordan de nøjagtige formuleringer i de gældende bedømmelsesværktøjer lyder. Eller for at sikre, at begge bedømmere har den faglige bedømmelse for øje, og at bedømmelsen sker absolut frem for ved relativ sammenligning af studerende.

Forcensur eller fælles drøftelse af opgaver sikrer ensartethed i forbindelse med de skriftlige opgaver

Flere bedømmere anvender forcensur ved skriftlige opgaver. De vurderer, at dette redskab er meget brugbart ift. at sikre ensartet bedømmelse på tværs af bedømmere. Det giver desuden bedømmerne mulighed for at afklare eventuelle uenigheder eller fastlægge opgavernes niveau.



Vi laver en forhåndsvotering ved nogle eksaminer, hvor vi læser de samme opgaver og diskuterer for at sikre os, at vi vurderer ens. Det fungerer rigtig godt. Men der er enighed næsten hver gang. Vi er højst én karakter fra hinanden.

En enkelt bedømmer påpeger dog, at der også kan være en risiko ved forcensur. Eksempelvis kan en dominerende bedømmer være med til at bestemme niveauet for samtlige opgaver, hvilket kan være problematisk.

Andre bedømmere fortæller, at der internt på deres uddannelse er blevet igangsat forskellige initiativer med henblik på at drøfte bedømmelse, fx at man har læst en række opgaver og så drøftet, hvad man lagde vægt på i bedømmelsen:



Vi har sammen bedømt et konkret bachelorprojekt. Det giver rigtig meget. Vi var ret uenige, så det var rigtig brugbart at kunne diskutere.

Censorerfaringer understøtter ensartet bedømmelse

Bedømmerne vurderer, at censorformandskabet og censorinstitutionen generelt spiller en stor rolle ift. at kvalitetssikre bedømmelsen. Dette sker blandt andet, ved at man som underviser møder forskellige bedømmere og dermed bliver bedre til at vurdere de studerendes niveau. Bedømmerne fortæller derudover, at censorformandskabet er med til at sikre kvaliteten af bedømmelsen gennem deres arbejde med at udpege censorer. Et eksempel på dette er, at formandskabet er meget opmærksomme på parringen af censor og bedømmer – eksempelvis ved at koble uerfarne bedømmere til erfarne censorer eller omvendt. En bedømmer siger: "Ekstern censur betyder også meget. At man vælger de rigtige censorer. Der får man ensrettet sin vurdering." På nogle uddannelser er censorbeskikkelsen blevet omorganiseret, hvilket de berørte bedømmere vurderer, har haft negative konsekvenser med hensyn til kvaliteten af de censorer, som udpeges til eksamenen.



Det nye sensorsystem er en katastrofe. Der skal censorerne selv angive, hvad de er gode til, uden nogen checks and balances. Rutinen er væk, fordi det kommer i udbud centralt. Det er så centraliseret, at det er et stående censorkorps. Der sendes så et tilbud ud, og den første, der melder sig, får lov. Den kæmpe overlevering af rutine, der var før, forsvinder.

Bedømmerne oplever i varierende grad, at der er ledelsesmæssigt fokus på karaktergivning

Det varierer meget, om bedømmerne oplever, at ledelsen har fokus på karaktergivning. Nogle bedømmere forklarer, at karaktergivning ikke fylder noget i dialogen med ledelsen, mens andre fortæller, at karaktergivning og bedømmelsespraksis bliver adresseret af ledelsen på forskellig vis. Der er da også forskellige eksempler i fokusgrupperne på, at ledelserne reagerer på karaktergivning, hvilket er illustreret nedenfor:



Hvis der er mere end 30 % dumpede studerende hos os, skal underviseren vise en handlingsplan for, hvordan han/hun vil ændre på resultaterne.



Der er fokus på praktikeksaminer, fordi der gives markant højere karakterer end i de andre fag. Det diskuterer vi. For hvor sandfærdigt er det billede? Der er heller ikke eksterne eksaminatorer, så der sidder jeg alene, og det er meget problematisk. Jeg kan jo kun lave en andenordens-bedømmelse, hvor jeg bedømmer deres refleksioner over deres praktikophold. Og det fortæller jeg også vores studerende.



Jeg fik en besked fra studieadministrationen om, at mit karakterniveau havde været stigende i en periode. Jeg kunne se på statistikkerne, at jeg havde givet mange 10-taller. Der tænkte jeg, at jeg skulle tage et kig på læringsmålene og se, hvad de beskriver.



Jeg tror, at vi alle kender situationer, hvor vi har en fornemmelse af, at karaktersnittet for et fag er for højt år ud og år ind. Nogle gange kan det også være lavt. Fra studienævnets side er vi blevet bedt om at se på kurser, hvor karaktersnittet er forbavsende lavt. Der er vi blevet bedt om at holde øje med, hvad årsagen til det kunne være. Men det er meget ud fra fornemmelsen. Vi har også statistikker, vi kan følge. Det har ikke gjort det nemmere, at det er absolut karaktergivning, men over tid skal være relativt. Det er især svært at forstå for udenlandske bedømmere. Det gør det svært at udforme læringsmålene. Man må acceptere, at der findes kurser, hvor karakterskalaen bruges på mærkværdig vis.

6 Karakterskalaens udformning

6.1 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 – bilag til *Evaluering af 7-trins-skalaen*” for yderligere beskrivelse af sammenhængen). Det betød, at der var større spring mellem de tre midterste beståede karakterer 4, 7 og 10 end mellem de øvrige karakterer. Samtidig var det en prioritet for Karakterkommissionen at understøtte, at der var en klar trindifferentiering og dermed en større sikkerhed 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 demonstrerer 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, hvorfor man endte med karakteren -3.

I dette kapitel sætter vi fokus på evalueringsspørgsmål 3: ”Hvad betyder karakterspringene og karakteren -3 for karaktergivningen?”

Evalueringen har primært fokuseret på den summative brug af karakterskalaen, selvom den formative 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 summative vurdering, idet man her skal hjælpe den studerende til at forbedre sine præstationer.

6.2 Survey blandt bedømmere på grundskoleområdet

6.2.1 Sammenfatning

Surveyen blandt bedømmere på grundskoleområdet viser overordnet set, at:

- 71 % oplever, at der er nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, og at det særligt er mellem 4 og 7 og 7 og 10.
- Størstedelen af bedømmerne oplever, at de brede trin i 7-trins-skalaen dækker over store faglige forskelle og rummer for forskellige niveauer i præstationerne.
- Lidt over halvdelen af bedømmerne oplever, at grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen, og det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 eller 0. Samtidig oplever de også, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende.

6.2.2 Gennemgang af resultater

Tabel 6.1 viser, at 71 % af bedømmerne på grundskolerne svarer ja i forbindelse med udsagnet "Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja i forbindelse med udsagnet "Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?", er statistisk signifikant forskellig mellem bedømmere fra dansk (76 %) og bedømmere fra matematik (62 %).

TABEL 6.1

Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?

	Procent (antal)
Ja	71 % (475)
Nej	29 % (190)
Total	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 6.2 viser, at 69 % af bedømmerne på grundskolerne angiver "4-7" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. 60 % angiver "7-10" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. Ved de øvrige steder på skalaen angiver markant færre, at det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har angivet "00-02" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere fra dansk (12 %) og bedømmere fra matematik (22 %).
- Andelen, der har angivet "7-10" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere fra engelsk (69 %) og bedømmere fra matematik (51 %).

TABEL 6.2

Hvor på skalaen er det særligt svært at vurdere, hvilken karakter eleven/den studerende skal have? (n=475)

	Procent (antal)
-3-00	7 % (33)
00-02	17 % (83)
02-4	19 % (90)
4-7	69 % (326)
7-10	60 % (285)
10-12	20 % (96)
Total	192 % (475)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Note: Spørgsmålet er stillet som et multiple-choice spørgsmål, hvorfor procenterne summer til mere end 100.

Note: Spørgsmålet er stillet til bedømmere, der synes, at der er et bestemt område af skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have.

Tabel 6.3 viser, hvordan bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn. I det følgende oplistes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 84 % af bedømmerne på grundskolerne er enige i udsagnet "De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevens progression mellem den ene standpunktskarakter/årskarakter og den næste".
- 82 % er enige i udsagnet "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne".
- 43 % er enige i udsagnet "Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter".

- 41 % er enige i udsagnet ”De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter”.
- 27 % er enige i udsagnet ”Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen”.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen”, er statistisk signifikant forskellig mellem bedømmere fra dansk (19 %) og bedømmere fra matematik (35 %).

TABEL 6.3

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter	6 % (37)	35 % (230)	22 % (147)	27 % (177)	10 % (66)	1 % (8)	0 % (0)	100 % (665)
Karaktertrinnene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne	37 % (249)	45 % (298)	8 % (53)	8 % (51)	1 % (6)	1 % (8)	0 % (0)	100 % (665)
Springene mellem 4, 7 og 10 fører jævnlige til en højere bedømmelse for ikke at give den studerende en for lav karakter	11 % (73)	32 % (213)	28 % (189)	23 % (156)	4 % (29)	1 % (5)	0 % (0)	100 % (665)
Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen	5 % (35)	22 % (149)	18 % (122)	36 % (239)	16 % (108)	2 % (12)	0 % (0)	100 % (665)
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.	41 % (269)	43 % (284)	8 % (50)	7 % (44)	1 % (9)	1 % (6)	0 % (1)	100 % (663)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

Tabel 6.4 viser, hvad bedømmerne på grundskoleområdet svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der har svaret at de er enige (enten ”Meget enig” eller ”Enig”) i udsagnene:

- 73 % af bedømmerne på grundskolerne er enige i udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen".
- 73 % er enige i udsagnet "Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende".
- 72 % er enige i udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse".
- 68 % er enige i udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen".
- 60 % er enige i udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse", er statistisk signifikant forskellig mellem bedømmere fra dansk (75 %) og bedømmere fra matematik (65 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen", er statistisk signifikant forskellig mellem bedømmere fra engelsk (56 %) og bedømmere fra matematik (70 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0", er statistisk signifikant forskellig mellem bedømmere fra dansk (53 %) og bedømmere fra matematik (65 %).

TABEL 6.4

Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen	53 % (353)	20 % (133)	10 % (65)	9 % (61)	7 % (46)	1 % (7)	100 % (665)
Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse	32 % (215)	40 % (265)	12 % (79)	13 % (87)	2 % (15)	1 % (4)	100 % (665)
Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende	48 % (316)	25 % (166)	9 % (58)	12 % (82)	5 % (36)	1 % (7)	100 % (665)
Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen	14 % (94)	54 % (361)	15 % (99)	12 % (80)	4 % (25)	1 % (6)	100 % (665)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0	21 % (137)	39 % (261)	13 % (88)	17 % (113)	8 % (50)	2 % (16)	100 % (665)

Kilde: EVA, survey maj 2018 blandt bedømmere på grundskoleområdet.

6.3 Survey blandt bedømmere på de gymnasiale uddannelser

6.3.1 Sammenfatning

Surveyen blandt bedømmere på de gymnasiale uddannelser viser overordnet set, at:

- 69 % oplever, at der er nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, og at det særligt er mellem 4 og 7 og 7 og 10.
- Størstedelen af bedømmerne oplever, at de brede trin i 7-trins-skalaen dækker over store faglige forskelle og rummer for forskellige niveauer i præstationerne.
- Lidt over halvdelen af bedømmerne oplever, at grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen, og det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 eller 0. Samtidig oplever de også, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende.

6.3.2 Gennemgang af resultater

Tabel 6.5 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn om springene på karacterskalaen. I det følgende oplistes andelen, der har svaret at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 82 % af bedømmerne på de gymnasiale uddannelser er enige i udsagnet "De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevens progression mellem den ene standpunktskarakter/årskarakter og den næste".
- 81 % er enige i udsagnet "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne".
- 47 % er enige i udsagnet "Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter".
- 34 % er enige i udsagnet "De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter".
- 23 % er enige i udsagnet "Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter", er statistisk

signifikant forskellig mellem bedømmere i dansk (32 %) og bedømmere i matematik (44 %), såvel som mellem bedømmere fra engelsk (26 %) og bedømmere fra matematik (44 %).

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne”, er statistisk signifikant forskellig mellem bedømmere i dansk (87 %) og bedømmere i matematik (72 %), såvel som mellem bedømmere fra engelsk (84 %) og bedømmere fra matematik (72 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter”, er statistisk signifikant forskellig mellem bedømmere i dansk (51 %) og bedømmere i matematik (38 %), såvel som mellem bedømmere fra engelsk (49 %) og bedømmere fra matematik (38 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen”, er statistisk signifikant forskellig mellem bedømmere i dansk (18 %) og bedømmere i matematik (34 %), såvel som mellem bedømmere i engelsk (21 %) og bedømmere fra matematik (34 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevens progression mellem den ene standpunktskarakter/årskarakter og den næste”, er statistisk signifikant forskellig mellem bedømmere i dansk (85 %) og bedømmere i matematik (77 %), såvel som mellem bedømmere i engelsk (85 %) og bedømmere fra matematik (77 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hhx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne” er statistisk signifikant forskellig mellem bedømmere på stx (83 %) og bedømmere på htx (67 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter”, er statistisk signifikant forskellig mellem bedømmere på stx (46 %) og bedømmere på hhx (86 %), såvel som mellem bedømmere på hf (41 %) og bedømmere på hhx (86 %), såvel som mellem bedømmere på hhx (86 %) og bedømmere på htx (39 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen”, er statistisk signifikant forskellig mellem bedømmere på stx (22 %) og bedømmere på hf (32 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”De brede trin i 7-trins-skalaen dækker over store faglige forskelle, og det gør det ofte vanskeligt at synliggøre elevens progression mellem den ene standpunktskarakter/årskarakter og den næste”, er statistisk signifikant forskellig mellem bedømmere på stx (83 %) og bedømmere på hf (61 %), såvel som mellem bedømmere på hf (61 %) og bedømmere på hhx (81 %).

TABEL 6.5

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter	6 % (75)	28 % (333)	24 % (280)	27 % (322)	12 % (147)	3 % (30)	0 % (0)	100 % (1187)
Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne	38 % (451)	43 % (516)	8 % (98)	7 % (87)	2 % (20)	1 % (15)	0 % (0)	100 % (1187)
Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter	14 % (167)	33 % (389)	23 % (269)	23 % (278)	4 % (46)	3 % (38)	0 % (0)	100 % (1187)
Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen	5 % (59)	18 % (211)	18 % (210)	35 % (412)	20 % (241)	5 % (54)	0 % (0)	100 % (1187)
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.	45 % (534)	37 % (438)	8 % (98)	5 % (58)	2 % (18)	2 % (22)	1 % (8)	100 % (1176)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Tabel 6.6 viser, at 69 % af bedømmerne på de gymnasiale uddannelser svarer ja i forbindelse med udsagnet "Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Andelen, der har svaret ja, er ikke statistisk signifikant forskellig på tværs af fag. Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Andelen, der har svaret ja, er ikke statistisk signifikant forskellig forskellige gymnasiale uddannelser.

TABEL 6.6

Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?

	Procent (antal)
Ja	69 % (816)
Nej	31 % (371)
Total	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Tabel 6.7 viser, at 69 % af bedømmerne på de gymnasiale uddannelser angiver "4-7" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. 68 % angiver "7-10" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. Ved de øvrige steder på skalaen angiver markant færre, at det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har angivet "00-02" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere fra dansk (9 %) og bedømmere fra matematik (18 %), såvel som mellem bedømmere fra engelsk (15 %) og bedømmere fra matematik (18 %).
- Andelen, der har angivet "02-4" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere fra dansk (10 %) og bedømmere fra matematik (20 %).
- Andelen, der har angivet "4-7" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere fra dansk (75 %) og bedømmere fra matematik (60 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, htx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har angivet "-3-00" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere på hf (11 %) og bedømmere på htx (30 %), såvel som mellem bedømmere på htx (30 %) og bedømmere på stx (13 %).
- Andelen, der har angivet "10-12" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere på htx (16 %) og bedømmere på htx (48 %), såvel som mellem bedømmere på hf (13 %) og bedømmere på htx (48 %), såvel som mellem bedømmere på stx (16 %) og bedømmere på htx (48 %).

TABEL 6.7

Hvor på skalaen er det særligt svært at vurdere, hvilken karakter eleven/den studerende skal have? (n=816)

	Procent (antal)
-3-00	14 % (111)
00-02	13 % (108)
02-4	14 % (113)
4-7	69 % (566)
7-10	68 % (553)
10-12	17 % (142)
Total	195 % (816)

Note: Spørgsmålet er stillet som et multiple-choice spørgsmål, hvorfor procenterne summer til mere end 100.

Note: Spørgsmålet er stillet til bedømmere, der synes, at der er et bestemt område af skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

Tabel 6.8 viser, hvad bedømmerne på de gymnasiale uddannelser svarer i forbindelse med en række udsagn. Det følgende oplystes andelen, der har svaret at de er enige (enten "Meget enig" eller "Enig") i udsagnene, listet fra højeste andel til laveste andel:

- 72 % af bedømmerne på de gymnasiale uddannelser er enige i udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen".
- 70 % er enige i udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse".
- 62 % er enige i udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen".
- 59 % er enige i udsagnet "Karakteren -3 er svært at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende".
- 59 % er enige i udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen", er statistisk signifikant forskellig mellem bedømmere fra dansk (70 %) og bedømmere fra matematik (48 %), såvel som mellem bedømmere i engelsk (68 %) og bedømmere i matematik (48 %).

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse", er statistisk signifikant forskellig mellem bedømmere fra dansk (78 %) og bedømmere fra matematik (54 %), såvel som mellem bedømmere i engelsk (76 %) og bedømmere i matematik (54 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende", er statistisk signifikant forskellig mellem bedømmere fra dansk (69 %) og bedømmere fra matematik (40 %), såvel som mellem bedømmere i engelsk (64 %) og bedømmere i matematik (40 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0", er statistisk signifikant forskellig mellem bedømmere fra dansk (53 %) og bedømmere fra matematik (73 %), såvel som mellem bedømmere i engelsk (53 %) og bedømmere i matematik (73 %).

Variationer i svar mellem bedømmerne fra hhv. stx, hf, hxx og htx er ligeledes at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen", er statistisk signifikant forskellig mellem bedømmere på stx (61 %) og bedømmere på hxx (72 %), såvel som mellem bedømmere på hf (56 %) og bedømmere på hxx (72 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0", er statistisk signifikant forskellig mellem bedømmere på hxx (57 %) og bedømmere på htx (38 %), såvel som mellem bedømmere på hf (69 %) og bedømmere på htx (38 %), såvel som mellem bedømmere på stx (60 %) og bedømmere på htx (38 %).

TABEL 6.8

Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen	40 % (479)	22 % (264)	14 % (171)	15 % (176)	7 % (78)	2 % (19)	100 % (1187)
Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse	35 % (414)	35 % (417)	14 % (166)	12 % (148)	3 % (35)	1 % (7)	100 % (1187)
Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende	31 % (365)	28 % (329)	12 % (140)	20 % (235)	9 % (110)	1 % (8)	100 % (1187)
Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen	15 % (175)	57 % (679)	15 % (173)	10 % (124)	2 % (25)	1 % (11)	100 % (1187)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0	18 % (217)	41 % (485)	14 % (168)	18 % (216)	8 % (90)	1 % (11)	100 % (1187)

Kilde: EVA, survey maj 2018 blandt bedømmere på de gymnasiale uddannelser.

6.4 Survey blandt bedømmere på erhvervsuddannelserne

6.4.1 Sammenfatning

Surveyen blandt bedømmere på erhvervsuddannelserne viser overordnet set, at:

- 67 % oplever, at der er nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, og at det særligt er mellem 4 og 7 og 7 og 10
- Størstedelen af bedømmerne oplever, at de brede trin i 7-trins-skalaen dækker over store faglige forskelle og rummer for forskellige niveauer i præstationerne
- Omkring to-tredjedele bedømmerne oplever, at grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen, og det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 eller 0. Samtidig oplever de også, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende.

6.4.2 Gennemgang af resultater

Tabel 6.9 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 74 % er enige i udsagnet "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne"
- 40 % er enige i udsagnet "Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter"
- 39 % er enige i udsagnet "De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter"
- 25 % er enige i udsagnet "Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen", er statistisk signifikant forskellig mellem bedømmere fra engelsk (12 %) og bedømmere fra matematik (33 %).

TABEL 6.9

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter	6 % (19)	33 % (103)	24 % (74)	25 % (80)	11 % (34)	1 % (4)	0 % (0)	100 % (314)
Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne	35 % (109)	39 % (121)	11 % (33)	13 % (41)	3 % (8)	1 % (2)	0 % (0)	100 % (314)
Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter	14 % (44)	26 % (81)	29 % (92)	23 % (72)	5 % (17)	3 % (8)	0 % (0)	100 % (314)
Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen	7 % (21)	18 % (58)	18 % (55)	34 % (108)	19 % (60)	4 % (12)	0 % (0)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Tabel 6.10 viser, at 67 % af bedømmerne på erhvervsuddannelserne svarer ja i forbindelse med udsagnet "Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja i forbindelse med udsagnet "Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have", er statistisk signifikant forskellig mellem bedømmere fra dansk (75 %) og bedømmere fra matematik (52 %), såvel som mellem bedømmere fra engelsk (82 %) og bedømmere fra matematik (52 %).

TABEL 6.10

Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?

	Procent (antal)
Ja	67 % (211)
Nej	33 % (103)
Total	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Tabel 6.11 viser, at 63 % af bedømmerne på erhvervsuddannelserne angiver "4-7" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. 62 % angiver "7-10" som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. Ved de øvrige steder på skalaen angiver markant færre, at det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have.

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. For hvert af områderne er andelen, der har angivet netop dette område som særligt svært, ikke statistisk signifikant forskellig på tværs af fag.

TABEL 6.11

Hvor på skalaen er det særligt svært at vurdere, hvilken karakter eleven/den studerende skal have? (n=211)

	Procent (antal)
-3-00	13 % (28)
00-02	18 % (39)
02-4	13 % (27)
4-7	63 % (132)
7-10	62 % (130)
10-12	19 % (40)
Total	188 % (211)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

Note: Spørgsmålet er stillet som et multiple-choice spørgsmål, hvorfor procenterne summer til mere end 100.

Note: Spørgsmålet er stillet til bedømmere, der synes, at der er et bestemt område af skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have.

Tabel 6.12 viser, hvad bedømmerne på erhvervsuddannelserne svarer i forbindelse med en række udsagn. I det følgende oplistes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 71 % er enige i udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen"
- 66 % af bedømmerne på erhvervsuddannelserne er enige i udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen"
- 65 % er enige i udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse"
- 59 % er enige i udsagnet "Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende"
- 49 % er enige i udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0".

Variationer i svar mellem bedømmerne fra hhv. dansk, matematik og engelsk er at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen", er statistisk signifikant forskellig mellem bedømmere i dansk (64 %) og bedømmere i engelsk (94 %), såvel som mellem bedømmere i engelsk (94 %) og bedømmere i matematik (60 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse", er statistisk signifikant forskellig mellem bedømmere i dansk (65 %) og bedømmere i engelsk (88 %), såvel som mellem bedømmere i engelsk (88 %) og bedømmere i matematik (51 %).
- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen", er statistisk signifikant forskellig mellem bedømmere i dansk (60 %) og bedømmere i matematik (77 %), såvel som mellem bedømmere i engelsk (58 %) og bedømmere i matematik (77 %).

TABEL 6.12

Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen	48 % (151)	23 % (71)	11 % (34)	12 % (37)	6 % (20)	0 % (1)	100 % (314)
Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse	32 % (101)	33 % (103)	15 % (46)	16 % (51)	4 % (11)	1 % (2)	100 % (314)

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende	33 % (103)	26 % (83)	14 % (44)	18 % (56)	7 % (22)	2 % (6)	100 % (314)
Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen	11 % (36)	55 % (174)	17 % (52)	10 % (31)	4 % (14)	2 % (7)	100 % (314)
Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0	10 % (30)	39 % (123)	16 % (51)	24 % (74)	10 % (31)	2 % (5)	100 % (314)

Kilde: EVA, survey maj 2018 blandt bedømmere på erhvervsuddannelserne.

6.5 Survey blandt bedømmere på de videregående uddannelser

6.5.1 Sammenfatning

Surveyen blandt bedømmere på de videregående uddannelser viser overordnet set, at:

- 76 % oplever, at der er nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have, og at det særligt er mellem 4 og 7 og 7 og 10.
- Størstedelen af bedømmerne oplever, at de brede trin i 7-trins-skalaen dækker over store faglige forskelle og rummer for forskellige niveauer i præstationerne.
- Omkring to-tredjedele bedømmerne oplever, at grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen, og omkring halvdelen, at det er tydeligt for dem, hvornår en præstation skal bedømmes med -3 eller 0. Samtidig oplever de også, at karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende.

6.5.2 Gennemgang af resultater

Tabel 6.13 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 80 % er enige i udsagnet "Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne".
- 53 % er enige i udsagnet "Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter".
- 29 % er enige i udsagnet "De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter".
- 20 % er enige i udsagnet "Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter”, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (27 %) og bedømmere på erhvervsakademiuddannelser (33 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne”, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (78 %) og bedømmere på professionsbacheloruddannelser (33 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 6.13

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

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Ikke relevant	Total
De brede karaktertrin (4, 7 og 10) gør det lettere at nå til enighed om, hvad der er den rigtige karakter	5 % (141)	24 % (696)	19 % (562)	31 % (916)	20 % (599)	1 % (32)	0 % (0)	100 % (2946)
Karaktertrinene på midten af skalaen (4, 7 og 10) er så brede, at de skal rumme for forskellige niveauer i præstationerne	40 % (1166)	40 % (1184)	10 % (280)	7 % (215)	2 % (54)	2 % (47)	0 % (0)	100 % (2946)
Springene mellem 4, 7 og 10 fører jævnligt til en højere bedømmelse for ikke at give den studerende en for lav karakter	17 % (505)	36 % (1056)	22 % (661)	19 % (574)	3 % (99)	2 % (51)	0 % (0)	100 % (2946)
Det har ikke nogen betydning for karaktergivningen, at der er forskellige afstande mellem trinnene forskellige steder på skalaen	4 % (122)	16 % (458)	14 % (408)	37 % (1101)	26 % (780)	3 % (78)	0 % (0)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 6.14 viser, at 76 % af bedømmerne på de videregående uddannelser svarer ja i forbindelse med udsagnet ”Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?”.

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret ja til spørgsmålet ”Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have”, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (75 %) og bedømmere på professionsbacheloruddannelser (79 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 6.14

Er der nogen områder på skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have?

	Procent (antal)
Ja	76 % (2243)
Nej	24 % (703)
Total	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Tabel 6.15 viser, at 74 % af bedømmerne på de videregående uddannelser angiver ”7-10” som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. 60 % angiver ”4-7” som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have. Ved de øvrige steder på skalaen angiver markant færre, at det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have.

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplistes alle statistisk signifikante forskelle:

- Andelen, der har angivet ”7-10” som et sted på skalaen hvor det er særligt svært at vurdere hvilken karakter, eleven/den studerende skal have, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (77 %) og bedømmere på professionsbacheloruddannelser (71 %), såvel som mellem bedømmere på universitetsuddannelser (77 %) og bedømmere på erhvervsakademiuddannelser (66 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 6.15

Hvor på skalaen er det særligt svært at vurdere, hvilken karakter eleven/den studerende skal have? (n=2243)

	Procent (antal)
-3-00	8 % (176)
00-02	13 % (297)
02-4	12 % (277)
4-7	60 % (1348)
7-10	74 % (1652)
10-12	23 % (512)
Total	190 % (2243)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

Note: Spørgsmålet er stillet som et multiple-choice spørgsmål, hvorfor procenterne summer til mere end 100.

Note: Spørgsmålet er stillet til bedømmere, der synes, at der er et bestemt område af skalaen, hvor det er særligt svært at vurdere, hvilken karakter eleven/den studerende skal have.

Tabel 6.16 viser, hvad bedømmerne på de videregående uddannelser svarer i forbindelse med en række udsagn. I det følgende oplystes andelen, der har svaret, at de er enige (enten "Meget enig" eller "Enig") i udsagnene:

- 70 % er enige i udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen".
- 70 % er enige i udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse".
- 65 % af bedømmerne på de videregående uddannelser er enige i udsagnet "Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen".
- 54 % er enige i udsagnet "Karakteren -3 er svært at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende".
- 49 % er enige i udsagnet "Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0".

Variationer i svar mellem bedømmerne fra hhv. erhvervsakademiuddannelser, professionsbacheloruddannelser og universitetsuddannelser, samt de kunstneriske – og maritime uddannelser at finde i den særskilte tabelrapport. Nedenfor oplystes alle statistisk signifikante forskelle:

- Andelen, der har svaret "Meget enig" eller "Enig" i forbindelse med udsagnet "Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse", er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (69 %) og bedømmere på professionsbacheloruddannelser (73 %).

- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende”, er statistisk signifikant forskellig mellem bedømmere på universitetsuddannelser (51 %) og bedømmere på professionsbacheloruddannelser (58 %), såvel som mellem bedømmere på universitetsuddannelser (51 %) og bedømmere på erhvervsakademiuddannelser (63 %).
- Andelen, der har svaret ”Meget enig” eller ”Enig” i forbindelse med udsagnet ”Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen”, er statistisk signifikant forskellig mellem bedømmere på erhvervsakademiuddannelser (69 %) og bedømmere på professionsbacheloruddannelser (62 %).

Variationer i svar mellem bedømmerne fra de forskellige hovedområder er ligeledes at finde i den særskilte tabelrapport.

TABEL 6.16

Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen?

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet i 13-skalaen	48 % (1406)	22 % (661)	11 % (334)	12 % (347)	5 % (156)	1 % (41)	100 % (2946)
Der er for få trin i skalaen til, at det er muligt at indfange alle vigtige nuancer i forbindelse med bedømmelse	35 % (1032)	35 % (1020)	12 % (344)	14 % (415)	4 % (124)	0 % (11)	100 % (2946)
Karakteren -3 er svær at bruge, fordi den sender et unødigt negativt signal til eleven/den studerende	29 % (844)	25 % (725)	15 % (445)	17 % (499)	10 % (293)	5 % (141)	100 % (2946)
Grænsen mellem bestået og ikke-bestået er tilstrækkelig tydelig i skalaen	13 % (390)	52 % (1521)	17 % (499)	13 % (389)	4 % (112)	1 % (34)	100 % (2946)
Det er tydeligt for mig, hvornår en præstation skal bedømmes med -3 eller 0	13 % (369)	36 % (1065)	17 % (512)	20 % (603)	10 % (281)	4 % (115)	100 % (2946)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

6.6 Fokusgruppeinterviews med bedømmere

I det følgende beskrives bedømmernes perspektiver på karakterskalaens udformning, som de er kommet til udtryk i fokusgruppeinterviewene.

6.6.1 Bedømmere på grundskoleområdet

Nogle bedømmere fortæller, at de store spring i midten af skalaen letter karaktergivning, andre, at de gør bedømmelsen sværere

Nogle bedømmere beskriver, at de store spring i midten af skalaen mellem 4, 7 og 10 gør det lettere for dem at give karakterer. En bedømmer siger:



Når de er så brede, så dækker de også et større område. Når man ikke skal overveje, om det er 7, 8 eller 9, så kan man bare give 7. Derfor er der nogle ting, der er mere enkle.

Andre bedømmere forklarer omvendt, at spændet i 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. Dette er, fordi eksempelvis 4- og 7-tallet kan rumme så forskellige faglige præstationer, at det bliver svært at afgøre, om der er tale om et stort 4-tal eller et lille 7-tal. En bedømmer siger:



Man kan ud fra kriterierne altid argumentere for både 4 og 7. Så der er noget vurdering og tolkning, og der kan vi være forskellige.

Men når karaktergivningen i midten af skalaen giver anledning til diskussioner, drejer det sig ikke nødvendigvis kun om skalaen. Det kan også dreje sig om, at det generelt kan være sværere at bedømme middelpåstættelser, hvor der er både gode og mindre gode ting og mange nuancer frem for klare top- og bundpræstationer.

Bedømmere savner at kunne nuancere og differentiere i midten af skalaen

Et element, der efterspørges af bedømmerne, er muligheden for at kunne nuancere og differentiere karaktergivningen mere i midten af skalaen. Det drejer sig om at kunne bedømme forskellige elever mere retvisende og nuanceret og om, at de store spring i 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 progression.

De store spring i midten af skalaen kan virke demotiverende for eleverne

Bedømmerne beskriver, at de store spring i midten af skalaen mellem 4, 7 og 10 kan have negativ indvirkning på elevernes motivation. For det første gør en bred karakter som 7-tallet det svært at synliggøre elevernes progression. Nogle bedømmere savner derfor et 8- eller 9-tal til at nuancere middelkarakteren. En bedømmer fortæller:



I de mundtlige eksaminer er der ofte brug for et 9-tal, hvis eleverne ligger lige mellem 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, der ligger på et stort 7-tal, et 10-tal.

En anden bedømmer fortæller, at de brede karakterer ikke blot betyder, at det kan være svært at synliggøre den enkelte elevs progression, men at de brede karakterer også betyder, at elever, der befinder sig på forskellige faglige niveauer, kan få samme karakter. En bedømmer siger:



I fx retskrivning og matematik kan det antal point, der giver en bestemt karakter, variere meget, hvilket virker demotiverende for eleverne, fordi de ikke kan se på karaktererne, at de rykker sig. Det er klart, at når man har haft den gamle skala, så føler man en mangel. I retskrivning, hvor man får point, der går det op for eleverne, at det er så bredt. Eleverne ved jo godt, hvad de kan, og hvad deres klassekammerater kan. Så to elever, der ligger vidt forskelligt, kan få 7.

En anden bedømmer beskriver, at det, at forskellige elever på forskellige faglige niveauer får samme karakterer, kan virke frustrerende for eleverne:



Jeg har også elever, der bliver frustreret over at få det samme som en anden. Men ud fra de kriterier, der er, tænker jeg, at de egentlig har løst det hhv. godt og mindre godt, men begge ligger på 7-taller.

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

Bedømmerne fortæller, at eleverne har en tendens til kun at betragte 10 og 12 som gode karakterer, mens 7-tallet ikke opleves som en god karakter af eleverne. En bedømmer siger: "Eleverne hænger sig i at være 7-tals-elever. De anser ikke 7 for at være en særligt god karakter." Nogle bedømmere mener, at en del af forklaringen kan være, at elevernes forældre selv er bedst bekendt med 13-skalaen og derfor betragter et 7-tal som en under middel-karakter.



Jeg synes ofte, at man møder en fordom fra forældrene. Der er mange af dem, der ikke synes, at 7 er en god karakter, og derfor synes eleverne det heller ikke.

En anden bedømmer fortæller:



Jeg læste lige en artikel om, at forældre poster, når børnene får 10 eller 12, men ingen poster 4-tallet. Det giver jo en skævvridning af, hvad der er den rigtige karakter at få. Når man kun hører om dem, der får 10 eller 12, så føler de andre sig skidt tilpas.

En tredje bedømmer mener, at en del af forklaringen kan være, at der gives mange 12-taller, og at det devaluerer værdien af et 7-tal for eleverne:



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.

-3 er en undtagelseskarakter med stor symbolsk betydning

Bedømmerne fortæller, at de meget sjældent anvender karakteren -3. Nogle beskriver, at de ikke engang har den inden for radaren som en karakter, man kan give. Adspurgt om, hvad forskellen på -3 og 00 er, og hvornår -3 kan komme i spil, lyder det, at ved 00 har man som elev gjort et forsøg, men er ikke lykkedes, mens man ved -3 slet ikke har gjort et forsøg eller decideret har udvist modvilje. En bedømmer 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.

Mens en anden bedømmer fortæller, at -3 også kan gives, hvis eleverne udviser decideret modvilje:



Jeg har en enkelt gang givet -3 i standpunktskarakter til en elev, som virkelig bare var kropumulig, og som jeg skulle ud og hente i cykelskuret for at få til time.

Bedømmerne fortæller, at -3 har stor symbolsk betydning, og bedømmerne er generelt kritiske over for symbolikken i at have en minuskarakter. Af denne grund benytter nogle bedømmere af princip ikke karakteren -3. En bedømmer fortæller:



Altså, jeg synes jo, at bare man møder op til en mundtlig prøve, så kan man ikke få en minuskarakter. Jeg ved godt, at man burde kunne. Men som menneske kan man jo ikke sige ”du er minus værd”.

Derudover begrundes nogle bedømmere også deres kritik af det at have en minuskarakter med den betydning, som en sådan karakter har for elevernes gennemsnit. En bedømmer udtaler:



Jeg tænker, at -3 er en decideret uskik. Det med, at den trækker de andre karakterer negativt ned ift. bare at have et 00. Jeg synes, at det er foruroligende.

Nogle bedømmere savner at kunne belønne den exceptionelle præstation

Nogle af bedømmerne savner at kunne belønne den exceptionelle præstation, hvor eleven gør noget ud over det sædvanlige. Bedømmerne beskriver 12-tallet som en meget bred karakter, ligesom 7-tallet, som kan dække over meget forskellige præstationer. En bedømmer siger: ”12-tallet er ikke exceptionelt, men det er til en, som har fået fat i det hele.” Men det gør det ifølge nogle bedømmere svært at honorere den virkelig gode opgave eller præstation. En bedømmer siger:



Jeg synes også, vi kan snakke om, at der mangler et tal oppe i toppen. Altså et tal, hvor man kan præmiere den helt sublime præstation. Altså ligesom det gamle 13-tal. Det er for nemt at få 12, og hvorfor skal man have en topkarakter for noget, der har mangler.

Men nogle bedømmere understreger, at de oftere savner en ekstra karakter i midten af skalaen end at kunne belønne den exceptionelle præstation. Samtidig er der også bedømmere, der af hensyn til sammenligneligheden med udlandet ikke ønsker en karakter som 13-tallet tilbage. En bedømmer siger:



Jamen, var årsagen til, at man fjernede 13-tallet, ikke også, at folk skulle kunne komme ind på Ivy League-skolerne. Det skal jo være sammenligneligt med andre lande. Jeg synes, at det ville være ærgerligt at have 13-tallet. Og en præstation med få fejl og mangler er da stadig en flot præstation.

Nogle bedømmere oplever et pres ift. at give højere karakterer

Nogle bedømmere beskriver, at de på forskellig vis oplever at være pressede ift. at give høje karakterer. Bedømmerne forklarer, at dette pres har forskellige årsager og kommer fra forskellige sider. Nogle beskriver, at det hænger sammen med, at karakterer er det, skolerne bliver målt på. En bedømmer siger:



Når man ser på kommunens fokus, så er det nationale tests etc. 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?

En anden bedømmer 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.

En tredje bedømmer problematiserer det, at man med en intenderet normalfordeling lægger et niveau for karaktererne:



Jeg synes, at det [intenderet fordeling efter normalfordelingen] er et issue, der virkelig skal tages op og gøres noget ved. Jeg tror, at det handler om at lægge et niveau. Jeg ved ikke, om der ligger noget ubevidst i, at alle føler sig pressede ift. at skulle levere høje karaktergennemsnit og høje resultater. Sidste år var jeg fx 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.

En bedømmer fortæller, at oplevelsen af pres ift. at levere høje karakterer har betydning for undervisningens indhold, som kan få præg af teaching to the test:



Jeg føler mig presset af, at der skal leveres bestemte karakterer – og helst så høje som muligt. Det tror jeg, har haft en indvirkning ift., at det bliver mere teaching to the test. Jeg er ikke sikker på, at det er den rigtige vej frem. Men det betyder da også noget for os. Jeg er meget mere nervøs for afgangsprøver nu, end jeg var tidligere. Der er også kollegaer, der fravælger overbygningen af den her grund.

En bedømmer supplerer ift., hvilken betydning det karakterpres, bedømmeren oplever, kan have for undervisningen:



Det kan også have betydning for selve undervisningens indhold. Tidligere kunne man bruge mere tid på noget alment dannende. Den tid har man ikke, man har nogle krav og vil have gode, høje karakterer. Så det kommer også til at gå lidt ud over indholdet.

6.6.2 Bedømmere på det gymnasiale område

Nogle bedømmere fortæller, at de store spring i midten af skalaen letter karaktergivning, andre, at de gør bedømmelsen sværere

Nogle bedømmere beskriver, at de store spring i midten af skalaen mellem 4, 7 og 10 gør det lettere for dem at give karakterer. En bedømmer siger: ”Det er blevet meget nemmere at give karakterer. Man behøver ikke gøre sig lige så umage, fordi man ikke skal finjustere.” En anden bedømmer op-ridser, hvordan der er både fordele og ulemper ved de store spring:



Jeg synes, at den går begge veje [ift. fordele og ulemper ved de store spring]. Det er dejligt nemt at give 7, fordi det er så bredt. Men samtidig kan jeg ikke lade være med at give 7 med pil op til nogle elever, og så har jeg jo allerede delt min karakterskala op i flere intervaller. Jo flere intervaller, du har at vælge mellem, jo sværere bliver det også at give karakterer. Men omvendt er det meget rart at kunne sige, at det er en god 7-tals-opgave, ikke en dårlig.

Andre bedømmere fortæller omvendt, at spændet i 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. Dette er, fordi eksempelvis 4- og 7-tallet kan rumme så forskellige faglige præstationer, at det bliver svært at afgøre, om der er tale om et stort 4-tal eller et lille 7-tal.

Men når karaktergivningen i midten af skalaen giver anledning til diskussioner, drejer det sig ikke nødvendigvis kun om skalaen. Det kan også dreje sig om, at det generelt kan være sværere at bedømme middelprestationer, hvor der er både gode og mindre gode ting og mange nuancer frem for klare top- og bundpræstationer.

Bedømmere savner at kunne nuancere og differentiere i midten af skalaen

Et element, der efterspørges af bedømmerne, er muligheden at kunne nuancere og differentiere karaktergivningen mere i midten af skalaen. Det drejer sig om at kunne bedømme forskellige elever mere retvisende og nuanceret og om, at de store spring i 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 progression.

Samtidig påpeger nogle bedømmere, at tanken med en karakterskala og brug af tal til at bedømme eleverne må være at give en præcis indikation på elevernes niveau. Men når de samme tal kan rumme store forskelligheder, bliver det en uklar indikator for, hvor eleven ligger, både for eleven selv og for andre, der skal forholde sig til elevens karakterer og gennemsnit. En bedømmer siger: "Karakterskalaen svarer ikke til virkeligheden. Det er som en termatarmåling, hvor nogle af temperaturerne mangler." En anden bedømmer siger:



Karakterskalaen er upræcis og ikke dækkende ift. virkeligheden. Det er upræcist, at den samme karakter rummer så mange forskelligheder. Vi er vant til, at tal er præcise. 7 er upræcist.

De store spring i midten af skalaen har stor betydning for elevernes gennemsnit

En del af bedømmerne fortæller, at de er meget opmærksomme på, at de store spring i midten af skalaen betyder, at det har stor betydning for elevernes gennemsnit, om det eksempelvis bliver et 7-tal eller et 10-tal. En bedømmer siger: "De store spring får store konsekvenser for elevernes snit, hvilket påvirker eleverne, men også de lærere, der skal give karakterer med stor betydning." Nogle bedømmere fortæller, at fordi det har stor betydning for elevernes gennemsnit, om det bliver eksempelvis 7 eller 10, så kan det for dem som bedømmere komme til at have den indvirkning, at de ofte lader karakteren vippe opad, hvis eleverne ligger på vippen.



Der er jo også et elevhensyn, for det kan jo have konsekvenser for dem, og det betyder jo meget, om det er 7 eller 10. Så man kan jo godt, når den ligger på vippen, have tendens til at tage den høje. Jeg kan ikke finde ud af, om det er hensigtsmæssigt, men det er sådan, det er blevet, siger min erfaring.

Andre bedømmere udtrykker, at de er opmærksomme på at undgå at runde op for at differentiere, fx ved at lade et stort 7-tal blive til 10 – for at undgå inflation i karaktergivning.

De store spring i midten af skalaen kan virke demotiverende for eleverne

En del af bedømmerne beskriver, at de store spring i midten af skalaen mellem 4, 7 og 10 kan have negativ indvirkning på elevernes motivation. For det første gør en bred karakter som 7-tallet det svært at synliggøre elevernes progression. Bedømmerne fortæller, at der eksempelvis er "meget langt fra bunden af 7-tallet til 10-tallet", hvilket betyder, at eleven kan "arbejde meget og blive bedre uden at rykke sig karaktermæssigt". Hvor man på den gamle skala, "hvis man fik 8, kunne arbejde sig op på 9. Men nu er det bare et bredt tal, så midtereleven har svært ved at rykke sig". En bedømmer siger:



Det kan være rigtig svært for elever at flytte sig. De starter måske på et lille 7-tal, og så bevæger de sig op, men det er stadig 7. Eleverne er pisseirriterede, for de føler bare, at de har fået et stempel, og nærmest, at man ikke læser deres opgave. På en eller anden måde kan jeg godt forstå det. Det kan være demotiverende for dem.

For det andet betyder de brede karakterer også, at elever, der befinder sig på forskellige faglige niveauer, kan få samme karakter. Dette har ifølge bedømmerne også indvirkning på elevernes motivation, fordi de har tendens til at sammenligne sig meget med hinanden. En bedømmer siger:



Jeg synes, at 7-tallet er det værste. For eleverne, de sammenligner sig med hinanden, og det burde de ikke, men det gør de. Så to, der begge har fået 7, kan sidde bagefter og se på hinanden og vide, at deres faglige niveau er vidt forskelligt, men 7 er bare det, der rummer dem begge. Jeg synes, at den er alt for bred.

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

Bedømmerne fortæller, at eleverne har en tendens til kun at betragte 10 og 12 som gode karakterer, mens 7-tallet ikke opleves som en god karakter af eleverne, hvorfor bedømmerne fortæller, at karakteren 7 fra deres side ofte får følgeskab af bemærkninger som ”og det er jo en god karakter”. Nogle bedømmere, særligt dem, der også har bedømt ud fra 13-skalaen, beskriver, at de også selv ”skal lære at betragte både 4 og 7 som gode karakterer. Jeg synes jo heller ikke, at de er det, når man lige ser dem på papir. Og eleverne synes i allerhøjeste grad det samme”.

-3 er en undtagelseskarakter med stor symbolsk betydning

Bedømmerne fortæller, at de meget sjældent anvender karakteren -3. Nogle beskriver, at de ikke engang har den inden for radaren som en karakter, man kan give. Adspurgt om, hvad forskellen på -3 og 00 er, og hvornår -3 kan komme i spil, lyder det, at ved 00 har man som elev gjort et forsøg, men er ikke lykkedes, mens man ved -3 slet ikke har gjort et forsøg eller decideret har udvist modvilje. En bedømmer fortæller om forskellen på -3 og 00:



Det er nok signalværdien. Hvis der er en elev, der ikke kan noget fagligt og heller ikke prøver og måske ligefrem viser foragt for det. Med 00 har man prøvet.

Bedømmerne fortæller, at det er sværere at give -3 ved en mundtlig bedømmelse end ved en skriftlig. Dels fordi man ved en mundtlig bedømmelse for det meste kan få hevet et forsøg på faglighed ud af eleven, mens der ved skriftlig bedømmelse kun er det at bedømme, man har på skrift. Dels fordi det opleves som sværere at give -3 til en person, man sidder ansigt til ansigt med, frem for til en anonym elev på skrift. Det hænger sammen med, at bedømmerne oplever, at -3 har stor symbolsk betydning. Bedømmerne er enige om, at -3 sender et stærkt signal til eleverne. Nogle bedømmere beskriver, at -3 kan bruges til at signalere, at det ikke er realistisk for eleven at bestå faget. En bedømmer fortæller:



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.

Andre bedømmere påpeger, at sådanne budskaber ikke bør gives til eleverne med en karakter, men kræver, at man tager en snak med eleven om, at der er alvorlige udfordringer.

Bedømmerne er generelt kritiske ift. symbolikken i at have en minuskarakter og spørger, om det skal forstås som, at man kan mindre, end da man kom. Af den grund anvender nogle bedømmere principielt ikke karakteren -3. En bedømmer fortæller:



Jeg tror kun, at jeg har givet -3 én gang. Det virker som at gnide salt i såret. Der er ikke nogen grund til at tvære folk ud.

Dumpegrænsen er tydelig, men vigtig at diskutere ekstra grundigt

Bedømmerne fortæller, at de ser dumpegrænsen som tydelig. Alligevel fortæller bedømmerne, at det er omkring dumpegrænsen, at censor og eksaminator drøfter bedømmelsen mest. Det skyldes dog ikke uklarhed med hensyn til dumpegrænsen, men at det har så stor betydning for eleven, om den pågældende dumper eller ej. Derfor mener bedømmerne, at det er vigtigt at bruge ekstra tid

på at være ekstra grundig i gennemgangen af, hvilke argumenter der taler for hhv. dumpet og bestået, og undersøge, om der er noget, der gør, at eleven ikke skal dumpe. En bedømmer fortæller:



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 00.

Forskellige bud på, hvad stigningen i andelen af 12-taller kan skyldes

Nogle bedømmere fortæller, at de selv har givet flere 12-taller over tid, mens andre bedømmere fortæller, at antallet af 12-taller for deres eget vedkommende er uændret over tid. Bedømmerne har forskellige bud på, hvad stigningen i andelen af 12-taller kan skyldes.

En del af bedømmerne beskriver, at de giver flere 12-taller nu end tidligere, fordi det har taget dem noget tid at omstille sig til 12-tallet og vænne sig til, at 12-tallet ikke svarer til 13-tallet, men rummer mere end det. I starten, da 7-trins-skalaen blev indført, anvendte nogle bedømmere derfor 12-tallet som et 13-tal.

En anden medvirkende forklaring på tendensen til højere karakterer er, ifølge bedømmerne, at eleverne har stort fokus på deres gennemsnit, og dermed på deres enkelte gymnasiekarakterer, med henblik på at kunne komme ind på konkrete videregående uddannelser eller for ikke at lukke nogle døre ift. videregående uddannelse. Det betyder ifølge bedømmerne, at en del elever er villige til at gøre en ekstra og meget målrettet strategisk indsats for at forbedre deres karakterer. Bedømmerne har forskellige perspektiver på, om eleverne reelt er blevet mere bevidste om, hvad der stilles af krav for at komme ind på videregående uddannelse, eller om gymnasieelever altid har været bevidste herom, og det er kravene til dem, der er blevet skærpet. Lærerne fortæller, at de oplever karaktergivning som et kæmpe ansvar, fordi karaktererne betyder så meget for eleverne, og de store spring kan gøre en stor forskel for elevernes snit.

Bedømmerne fortæller også, at de oplever, at der har været en særlig opmærksomhed på 12-tallet – og karakterer og gennemsnit generelt. Fx beskriver nogle bedømmere, at der har været ”en landsudmelding om, at der blev givet for få 12-taller”. Lærerne mener dog ikke, at det er deres opgave at ramme normalfordelingen og derved give relative karakterer, men derimod at give absolutte karakterer ift. skala og mål, og så kan de nogle gange ligge højt, andre gange lavere. Den opmærksomhed over for 12-tallet, som bedømmerne har oplevet fra centralt hold, kan dog også hænge sammen med den tidligere nævnte tendens til, at nogle bedømmere i starten anvendte 12-tallet som et 13-tal.

Endelig beskriver bedømmerne generelt, at hvis der er uenighed eller tvivl om, hvorvidt en elev skal have den ene eller den anden af to karakterer ved siden af hinanden på skalaen, så drøfter bedømmerne, hvad der taler for det ene eller det andet, men lader også tvivlen komme eleven til gode og vælger den højeste karakter ved fortsat tvivl. Nogle bedømmere kæder også dette sammen med oplevelsen af en tiltagende klagekultur blandt eleverne.

Bedømmerne peger dog også på, at prøveformerne kan have noget at sige ift., hvor let eller svært det er at få et 12-tal. Generelt oplever bedømmerne, at det er lettere at få et 12-tal mundtligt end skriftligt. Det kan ifølge bedømmerne hænge sammen med de krav, der bliver stillet mundtligt kontra skriftligt, men også med, at eleverne generelt opleves som stærkere mundtligt end skriftligt.

Nogle bedømmere savner at kunne belønne den exceptionelle præstation

En del af bedømmerne savner at kunne belønne den exceptionelle præstation, hvor eleven gør noget ud over det sædvanlige. Bedømmerne beskriver 12-tallet som en meget bred karakter – ligesom 7-tallet, som kan dække over meget forskellige præstationer. Det, bedømmerne savner, er derfor at kunne differentiere og nuancere bedømmelsen af de elever, der toppræsterer, sådan at man kan skelne mellem de elever, der får 12, fordi de ikke havde nogen væsentlige fejl og mangler, og de elever, der får 12, fordi de leverede en helt særegen præstation. En bedømmer udtaler:



Jeg synes, at der mangler et 13-tal. I matematik er det så rigtigt, fordi man tæller rigtige. Man kan have to elever, hvor den ene virkelig har indsigt og forstår at kommunikere, og han får også 12 – ligesom ham, der bare har lavet alle beregningerne rigtigt.

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



Det er utrolig svært at belønne den virkelig gode elev. Men det er ikke, fordi jeg savner 13-tallet. Der var så få, der fik det. Det er ikke svært at få 12 i dag. Det ved jeg heller ikke, om det skal være. 12-tallet er næsten lige så bredt som 7-tallet. Der kan være milevid forskel på to kursister, der får samme karakter.

En anden bedømmer udtaler:



Jeg savner ikke 13-tallet. Jeg gav så få, så jeg mangler det ikke. Men jeg kunne godt tænke mig, at man kunne give kryds og slange. Så man i hvert fald fik sendt et signal. Det måtte også gerne stå der, men behøvede ikke nødvendigvis at indgå i gennemsnittet. Det skulle kun være til 12-tallet – ikke til de øvrige karakterer.

Der findes dog også bedømmere, der ikke savner 13-tallet eller at kunne markere, når eleven gør noget ekstra, men er tilfredse med 12-tallet som topkarakteren. For disse bedømmere handler det om, at de ikke mener, at man bør bedømme eleverne på andet og mere end det, som de er blevet undervist i, og som det eksplicit er formuleret, at de skal lære. En bedømmer fortæller: ”De skal kunne det, som de har lært, og behøver ikke vise noget udover for at få topkarakter,” mens en anden bedømmer spørger: ”Hvorfor skal vi forvente mere af dem end det, de har lært?”

6.6.3 Bedømmere på de videregående uddannelser

Nogle bedømmere fortæller, at de store spring i midten af skalaen letter karaktergivning, andre, at de gør bedømmelsen sværere

Bedømmerne fortæller, at springene i midten af skalaen mellem 4, 7 og 10 opleves som store i kraft af deres numeriske værdi. Men bedømmerne har forskellige perspektiver på, hvorvidt de brede karakterer i midten af skalaen letter eller vanskeliggør karaktergivningen.

På den ene side forklarer nogle bedømmere, at karaktergivningen bliver lettere og enigheden mellem censor og eksaminator større, når karaktererne er brede, og der skal differentieres mindre. En bedømmer siger: ”Når man ser på eksamenssituationen, har det lettet karaktergivningen i praksis med færre karakterer.” En anden bedømmer siger: ”Jo flere trin, des nemmere er det at differentiere, men det bliver også sværere at bedømme.”

På den anden side fortæller andre bedømmere, at netop det, at trinene i midten af skalaen er så brede, kan gøre det vanskeligt at afgøre, om der er tale om den ene eller den anden karakter, og

om en studerendes præstation eksempelvis er til et stort 7-tal eller et lille 10-tal, fordi karaktererne kan rumme meget forskelligartede præstationer. For nogle bedømmere ligger det svære i at afgøre, om det skal være den ene eller den anden karakter i midten af skalaen ift. trinbeskrivelsens formuleringer. To bedømmere siger:



7 er virkelig en problematisk karakter. Det, der beskriver 7-tallet, er ”den gode præstation med en del mangler”, og det, der beskriver 10-tallet, er ”den fortrinlige præstation med få mangler”. Men man kan godt stå med en fortrinlig præstation med en del mangler og en god præstation med få mangler.



Jeg har et stort problem med, hvornår jeg skal give 4. Hvis man ser på 4 og 7, så står der den gode og den jævne præstation. Det kan gradbøjes, og man siger ofte, at det skal komme den studerende til gode.

For andre bedømmere er det, der vanskeliggør fastlæggelsen af karakteren, den numeriske størrelse på springet mellem eksempelvis 7 og 10, som opleves som betydelig, og valget af karakter bliver derfor et valg med stor betydning for den studerende. En bedømmer siger:



Jeg er enig i, at det oftest er i spændet [mellem 7 og 10], at tvivlen ligger. Vi er bevidste om, at det betyder meget for de studerende, om det bliver det ene eller det andet, måske især når det gælder 7 og 10.

Generelt beskriver bedømmerne, at hvis der er tvivl om, hvilken middelkarakter de studerende skal have, så kommer tvivlen de studerende til gode.

Nogle bedømmere efterlyser at kunne differentiere i midten af skalaen, andre efterlyser ensartede spring

En gruppe af bedømmerne efterlyser i højere grad at kunne nuancere og differentiere karaktergivning i midten af skalaen ved at have flere forskellige karakterer at vælge mellem, mens andre bedømmere ikke oplever et behov for flere karakterer, men i højere grad for, at man ensarter afstanden mellem trinene.

Når nogle bedømmere savner i højere grad at kunne differentiere karaktergivningen i midten af skalaen, så er det dels af hensyn til præcisionen i bedømmelsen af den studerende, og hvor klar en indikator karakteren er på den studerendes niveau, dels af hensyn til den studerendes motivation.

De store spring i midten af skalaen gør bedømmelsen upræcis

En gruppe af bedømmerne oplever, at den nuværende skala gør bedømmelsen af de studerende, der befinder sig i midten af skalaen, upræcis, fordi der ikke er flere karakterer, og karaktererne er så brede, at de dækker over meget forskelligartede faglige præstationer. En bedømmer fortæller, at disse faglige forskelle jo så ikke afspejler sig på eksamensbeviset:



Det er et problem, at det høje 7-tal er langt bedre end det lave 7-tal, men ser man på eksamensbeviser, vil der være mange, der har mange 7-taller, men som er ret ringe, eller omvendt. Men man kan ikke differentiere mellem de to typer.

En gruppe af bedømmere undrer sig over, at man netop i midten af skalaen, hvor hovedparten af de studerende ligger, har valgt så store spring og dermed begrænset mulighed for differentiering, mens springene er mindre andre steder på skalaen, hvor færre studerende befinder sig. På den baggrund efterlyser nogle bedømmere flere karakterer i midten af skalaen. En bedømmer siger:



Der er nogle "10-taller, hvor der er stor forskel på de to, der får 10. Men jeg synes ikke, at nogen skulle have 7. Jeg mangler at kunne signalere, at det er bedre end 7, men ikke godt nok til 10.

En anden bedømmer fortæller:



Det er ved 7-tallet, at jeg har min primære anke. Det dækker simpelthen for bredt. Jeg savner et 6- eller 8-tal. Jeg vil gerne tilkendegive, hvad præstationen afspejler.

De store spring i midten af skalaen kan virke demotiverende for de studerende

Det, at samme karakter kan rumme meget forskellige faglige præstationer og niveauer, beskriver nogle bedømmere også, kan være demotiverende for de studerende. Det drejer sig om, at det kan være svært at afspejle den enkelte studerendes egen progression med hensyn til karaktererne. Men også om, at det kan være svært for de studerende at forstå, at en gruppe af studerende med meget forskellige faglige niveauer vurderes ens. En bedømmer siger: "Det er svært, for hvis en studerende er gået fra et lille 7-tal og op til et stort 7-tal, så kan man ikke vise vedkommende det."

Nogle bedømmere finder springenes numeriske værdi problematisk

Andre bedømmere ønsker ikke flere karakterer i midten af skalaen, men derimod, at springene mellem de forskellige trin mindskes og ensartes, sådan at skalaen fremstår lineær. En bedømmer siger:



Jeg må indrømme, at jeg ikke har brug for flere tal. Men springene er for store. Jeg vil hellere bare have fra 1 til 5 og så uden springene. Når jeg læser karakterbeskrivelsen, giver springene ikke engang mening.

En anden bedømmer beskriver, at den numeriske værdi af trinene har betydning for både studerende og bedømmere, også selvom der kun ligger et karaktertrin imellem:



Jeg er ikke tilhænger af flere karakterer. Men springene er talmæssigt for store. Vi skal passe på med at sende signaler. Der er også en grund til, at de har fået højere karakterer. Det numeriske spring omkring 7 betyder meget – både for de studerende og for bedømmerne. 4-tallet opleves som noget, man næsten ikke kan tillade sig.

Nogle bedømmere påpeger også, at de ikke finder den betydning, som de forskellige spring mellem trinene har for de studerendes gennemsnit, rimelig. En bedømmer siger:



Så 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.

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

Bedømmerne fortæller, at de studerende har en tendens til kun at betragte 10 og 12 som gode karakterer, mens 7-tallet ikke opleves som en god karakter. En bedømmer siger: "Mange studerende bliver kede af at få 7. 10 er de som regel godt tilfredse med." En anden bedømmer udtaler: "Vi har fortsat en pædagogisk opgave i at forklare de studerende, at 7-tallet er defineret ved at være en god karakter." En anden bedømmer fortæller:



Jeg tænker også, at flere tager skolen meget seriøst og knækker uddannelseskoden. Jeg oplever fænomenet 12-talspiger. Der er mange af dem hos os. Jeg kan også sige 12-talsdreng, men dem har vi ikke så mange af. De her piger er meget perfektionistiske og græder, når de får 10. De svageste kommer ikke ind til studievejledningen, men det gør 12-talspigerne, fordi de vil høre, hvordan de kan få 12 næste gang. Der er en psykisk udvikling hos de studerende, som er meget bekymrende.

-3 er en undtagelseskarakter med stor symbolsk betydning

Bedømmerne fortæller generelt, at -3 er en karakter, de sjældent benytter. Nogle bedømmere benytter den en gang imellem, andre benytter den aldrig. Om brugen af -3 i sammenligning med 00 fortæller bedømmerne, at -3 kan komme i spil ved en fuldstændig uacceptabel eller decideret fraværende præstation. En bedømmer fortæller om sin brug af -3:



Der er et par per årgang, som præsterer så uacceptabelt, at vi ikke kan lægge dem på linje med dem, som får 00. Det er signalet i, at den studerende skal arbejde sig hele to karakterer op for at bestå.

Nogle bedømmere fortæller, at -3 særligt sjældent kommer i brug ved mundtlig bedømmelse, da man her kan have en dialog med den studerende, og man kun yderst sjældent oplever studerende, der decideret nægter at sige noget. Mens man ved de skriftlige prøver alene kan bedømme den studerende på det skrevne, hvor der eksempelvis kan være fuldstændigt fravær af fagtermer.

Nogle bedømmere gør brug af -3 til at signalere, at det ikke er inden for rækkevidde at bestå faget, og at den studerende måske endda bør genoverveje sit studievalg. Andre bedømmere mener dog, at det er vigtigt, at sådanne snakke tages på anden vis, i andre sammenhænge, frem for gennem en karakter. To bedømmere fortæller om brugen af -3 som en symbolsk karakter:



Jeg har faktisk givet -3 til en bachelorprøve som et signal om, at der manglede alt for meget, og at vedkommende ikke skulle blive sygeplejerske. Der var det en symbolsk karakter.



Man kan også give -3 for at signalere, at man ikke mener, at den studerende bør prøve at gå til endnu en reeksamen.

Andre bedømmere har svært ved at se meningen med to dumpekarakterer og er særligt kritiske over for den symbolik, der ligger i at operere med en minuskarakter. Nogle af disse bedømmere benytter derfor slet ikke -3. Bedømmerne siger blandt andet om -3: "Der er noget helt upædagogisk ved -3 på alle måder", "det er utilstadeligt – den gør ikke noget godt for de studerendes læring og burde smides væk" og "det er lidt et udtryk for, at man er dummere, end da man kom".

Vigtigt, at det er tydeligt for både bedømmere og studerende, når studerende dumper

Bedømmerne fortæller, at de som regel ikke er i tvivl om, hvorvidt en studerende skal dumpe eller ej – dog oplever nogle, at beskrivelserne af dumpegrænsen ikke er helt tydelige. Men bedømmerne beskriver, at uanset om dumpegrænsen opleves som tydelig eller ej, er dumpekarakteren noget af det, de diskuterer allermest ved karaktergivning, "fordi den er så afgørende for den studerendes fremtid", men også fordi "argumenterne skal være i orden, hvis der opstår en klagesag". Bedømmerne gør således meget ud af at argumentere for, om en studerende skal bestå eller ej, og at det står helt klart for den studerende. Men en del af bedømmerne er også meget opmærksomme på, at det skal stå klart for den studerende, at præstationen var utilfredsstillende, hvis den studerende dumper. Det drejer sig om at have gode argumenter i formidlingen af karakteren til den stude-

rende. Men flere bedømmere fortæller også, at de er opmærksomme på, at dette også skal indikeres undervejs under eksempelvis en mundtlig prøve. Bedømmerne beskriver, at i og med at trinbeskrivelsen har fokus på fejl og mangler, så er man ved en mundtlig eksamen også nødt til at stille spørgsmål, de studerende ikke kan svare på, og dermed få konkrete fejl og mangler frem i lyset, også selvom bedømmerne oplever, at de også på anden vis kan spore sig ind på de studerendes niveau. En bedømmer siger:



Ved mundtlige eksaminer skal der også være ting, de ikke kan svare på. Hvis man skal kunne dokumentere det i forbindelse med en klage, er det vigtigt. Ellers kan de ikke forstå, hvorfor det ikke blev 2, 4 eller 7. Man skal have fejlene frem i lyset, ellers forstår de det ikke, hvis de har svaret på alt.

Perspektiver på anvendelse og stigning i andelen af karakteren 12

Som en del af fokusgrupperne er bedømmerne blevet præsenteret for en række grafer, der viser karakterfordelingen på deres uddannelse over tid og på tværs af prøveformer. Tallene viser, at der generelt har været en stigning i karaktererne, herunder i andelen af 12-taller, over tid, samt at prøvekaraktererne ligger højere ved mundtlige prøver og afsluttende projekter end ved skriftlige prøver.

Nogle bedømmere genkender billedet med en stigning i andelen af 12-taller, mens andre ikke oplever denne stigning i deres egen praksis. Bedømmerne har forskellige perspektiver på anvendelsen af karakteren 12, som beskrives i det følgende.

Forskellige oplevelser af udvikling i de studerendes faglige niveau og kravene til dem

En mulig forklaring på stigningen i andelen af 12-taller kunne være, at de studerende er blevet dygtigere. Men bedømmerne har forskellige perspektiver på, hvorvidt dette er tilfældet. Nogle bedømmere oplever, at de studerende på deres uddannelse generelt er blevet dygtigere. Andre bedømmere oplever, at de studerende på deres uddannelse generelt er blevet fagligt svagere. Mens andre fortæller om en polarisering på deres uddannelser, hvor en gruppe af studerende er enormt dygtige, mens en anden gruppe af studerende er fagligt meget svage. Endelig mener nogle bedømmere, at man ikke kan tale om ensidigt dygtigere eller dårligere studerende, men at de studerende over tid er blevet dygtigere på nogle områder og dårligere på andre områder. Udsagn fra to bedømmere illustrerer, at bedømmerne dog ikke er enige om, hvor de studerende er blevet hhv. fagligt stærkere og fagligt svagere:



De studerende er blevet bedre til metode, og så får de studerende også bedre projekter. Men hvad angår de matematiske fag, så ser vi på vores skole, at vores elever er blevet dårligere til matematik.



De er blevet bedre til nogle ting og dårligere til andet. De er blevet bedre til matematik, men dårligere til retstavning.

På nogle uddannelser mener bedømmerne, at stigningen i andelen af 12-taller også kan hænge sammen med, at de faglige krav til de studerende er mindsket. En del af disse bedømmere kæder dette sammen med de besparelser, som sker på uddannelserne, og som gør, at man ikke kan klæde de studerende på og stille og bedømme samme krav som tidligere. En bedømmer siger:



Hvis jeg kigger i læreplaner, oplever jeg, at der er sket en udtæring af læringsmål. Man har pillet punkter i læreplanerne ud i takt med manglen på ressourcer. Med hensyn til studieordninger sker der revisioner, hvor læringsmål er røget ud, men derudover er de røget ud pga. besparelser.

De studerende har fokus på karakterer, målopfyldelse og at knække eksamenskoden

Bedømmerne fortæller, at karakterer betyder mere for de studerende, og at 12 for en del af de studerende betragtes som den eneste acceptable karakter. En bedømmer fortæller:



Dem, 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.

Dette fokus på karakterer og herunder for nogle studerende specifikt på 12-taller betyder, ifølge bedømmerne, at de studerende er blevet mere fokuseret på og bedre til at forventningsopfylde og knække eksamenskoden. En bedømmer siger: "Det er interessant at se, hvor gode de er blevet til at kunne det, de bliver bedømt på." En anden bedømmer fortæller:



Vi snakker tit om, at vores studerende, som går på linjen for kliniske diætister, har knækket koden 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.

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



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.

Nogle bedømmere peger på, at når en del af de studerende opleves som bedre til at afkode og opfylde målene på uddannelserne, så behøver det ikke alene hænge sammen med de studendes eget fokus herpå, men også med, at man på en del uddannelser i de senere år har haft fokus på i højere grad at tydeliggøre læringsmålene for de studerende.

Generelt peger de interviewede bedømmere på, at det at være bevidst om mål, og hvordan de nås, naturligvis er en vigtig del af det at være studerende, men at der også har brug for, at de studerende kan andet end dette:



Jeg er enig i, at de er blevet gode til at forventningsopfylde. Det er ikke nødvendigvis dårligt, men vi har brug for nogle, der kan andet end at opfylde forventninger.

Omstilling fra 13- til 12-tallet

Endelig peger nogle bedømmere på, at stigningen i andelen af 12-taller også kan hænge sammen med bedømmernes omstilling fra 13-tallet til 12-tallet, og at bedømmerne har skullet vænne sig til, at 12-tallet rummer andre og bredere faglige præstationer end 13-tallet. En bedømmer fortæller:



I starten med den nye karakterskala erindrer jeg, at vi ikke havde nok topkarakterer. Vi kunne ikke vænne os til at give topkarakter til nogle, der havde lavet nogle fejl. Da vi først vænnede os til det, gav vi flere 12-taller. Der har vi måske haft svært ved at kalibrere mellem 10 og 12. Det er bare min fornemmelse. Det er en form for udsving. Først den ene vej og så den anden vej. Måske skal vi bare finde balancen igen.

Prøveformen har betydning for karakterfordelingen

Bedømmerne peger på, at prøveformen har betydning for karakterfordelingen. Det drejer sig om, om det er en mundtlig eller skriftlig prøve, ligesom der også gør sig noget særligt gældende ved afsluttende projekter med vejledning.

Karakterfordelingerne for de videregående uddannelser fordelt på prøveformer viser, at de studerende generelt får højere karakterer ved mundtlige end ved skriftlige prøver. Dette billede er generelt genkendeligt for bedømmerne i fokusgrupperne. De oplever generelt, at de studerende er bedre mundtligt end skriftligt. En bedømmer fortæller:



Det er også meget sværere at gå op til skriftlige eksaminer. Hos os er de studerende dårligt skolede i at formulere sig. De går jo til en masse mundtlige prøver hos os. Så det er en genre, de mestrer. Til gengæld har de kun en skriftlig. Så det afspejler sig jo helt klart i lavere karakterer ved de skriftlige.

Naturligt med høje karakterer ved afsluttende projekter

Særligt ved afsluttende projekter ligger karaktererne højt. Dette overrasker ikke bedømmerne. De finder det derimod af flere årsager naturligt, at karaktererne vil ligge højt ved netop afsluttende projekter. For det første beskriver bedømmerne, 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. En bedømmer siger:



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

For det andet mener bedømmerne, at det har betydning, at de studerende selv vælger, hvad de vil beskæftige sig med i det afsluttende projekt, 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 inden for. Nogle bedømmere oplever også som ovenfor nævnt, at de studerende ofte vælger projekter, de oplever som sikre ift. at få gode karakterer, frem for at vælge projekter, der udfordrer, men hvor de måske ikke får 12.

For det tredje peger nogle bedømmere også på, at det har betydning for niveauet, at de studerende i særlig grad får vejledning undervejs i arbejdet med de afsluttende projekter. En bedømmer siger:



Det er meget tydeligt, at karaktererne for bacheloropgaven ligger højere end for de løbende skriftlige eksaminer. Det har at gøre med formen. En stor skriftlig opgave, hvor de studerende har individuel kontakt med vejleder. De vil gerne have dedikeret vejledning i mindre grupper eller individuelt. Og det kan være med til at løfte mellemgruppen op. Men det har også at gøre med ressourcer. Det er umuligt at give alle individuel vejledning hele vejen gennem uddannelsen.

På den baggrund finder bedømmerne det generelt naturligt, at karaktererne ved de afsluttende projekter ligger højt, og ville finde det mere problematisk, hvis det ikke var tilfældet. En bedømmer siger:



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

6.7 Enkeltinterviews med bedømmere

I det følgende beskrives bedømmernes perspektiver på karakterskalaens udformning, som de er kommet til udtryk i de individuelle interviews.

6.7.1 Bedømmere på grundskoleområdet

Hovedpointerne ift. karakterskalaens udformning er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere på grundskoleområdet. I dette afsnit går en del af overskrifterne fra afsnittet om fokusgrupperne således igen, men pointerne illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere oplever, at de store spring i midten af skalaen letter karaktergivningen, andre, at de gør bedømmelsen sværere

Nogle bedømmere beskriver, at de store spring i midten af skalaen mellem 4, 7 og 10 gør det lettere for dem at give karakterer, fordi der er færre karakterer at vælge mellem, og springene er større. En bedømmer siger:



Den nye karakterskala er nemmere, og der er ikke mange diskussioner. Der er ikke nogen petitesseer, noget med små forskelle, så det er lettere. Der er spring mellem dem, men det gør ikke noget.

Andre bedømmere oplever omvendt, at spændet i 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. Dette er, fordi eksempelvis 4- og 7-tallet kan rumme så forskellige faglige præstationer, at det bliver svært at afgøre, om der er tale om et stort 4-tal eller et lille 7-tal. En bedømmer beskriver, at de brede karakterer i midten af skalaen og beskrivelserne heraf gør, at de kan rumme så meget, at det kan være vanskeligt at afgøre, om der skal være tale om den ene eller den anden karakter:



Mellem 02 og 7 kan man altid argumentere for en af de andre karakterer med et trins mellemrum. Karaktererne er så brede, at de overlapper hinanden. Så man er helt nede og pille ved detaljer i bedømmelsen. Det er uklart, hvornår det er "en del" eller "adskillige væsentlige" mangler. Det vil sige, at brugen af tillægsord gør skalaen vanskelig at bruge.

Men når karaktergivningen i midten af skalaen giver anledning til diskussioner, drejer det sig ikke nødvendigvis kun om skalaen. Det kan også dreje sig om, at det generelt kan være sværere at bedømme middelprestationer, hvor der er både gode og mindre gode ting og mange nuancer frem for klare top- og bundpræstationer.

Bedømmere savner at kunne nuancere og differentiere i midten af skalaen

En gruppe af bedømmerne savner at kunne nuancere og differentiere karaktergivningen mere i midten af skalaen. Det drejer sig om at kunne bedømme forskellige elever mere retvisende og nuanceret og om, at de store spring i midten af skalaen har stor betydning for elevernes karaktergenemsnit, men også om, at de store spring kan have betydning for elevernes motivation og muligheden for at synliggøre progression. En bedømmer siger: "Der er for stort et spring mellem 4 og 7 og 7

og 10. Ikke kun matematisk, men også i forståelsen hos lærere og elever.” En anden bedømmer udtaler:



Der er langt mellem 4 og 7. Hvad med at bruge bogstaver, så man ikke kan se den matematiske forskel? Afstanden betyder meget for gennemsnittet. Forskellen mellem 7 og 8 betød ikke så meget for gennemsnittet før, og selvom vi ikke snakker om det, så er det vigtigt for eleverne. Man kunne have brug for et 8-tal.

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

Bedømmerne fortæller, at eleverne har en tendens til kun at betragte 10 og 12 som en god karakter, mens 7-tallet ikke opleves som en god karakter. Bedømmerne beskriver, at denne forståelse ikke bare kommer fra eleverne selv, men også fra, hvordan 7-tallet betragtes i elevernes omverden, hvilket blandt andet kan hænge sammen med, at nogle forældre og lærere kan bære rundt på den forståelse af 7-tallet, som gjaldt for 13-skalaen. En bedømmer siger: ”Nogle hænger fast i 13-skalaen, og så er alt under 8 middel.” En anden bedømmer udtaler:



Eleverne synes, at 7 er kedeligt eller dårligt. Det handler om italesættelse fra mig og forældre og kollegaer – måske samfund. Man kunne tale om 7 som en god karakter i samfundet.

-3 er en undtagelseskarakter med stor symbolsk betydning

Bedømmerne fortæller, at de meget sjældent anvender karakteren -3 og for nogles vedkommende aldrig. En del af bedømmerne er kritiske over for det at have en minuskarakter. En bedømmer siger: ”-3 burde ikke være der. Det er ikke i orden over for eleverne.” Mens en anden bedømmer siger: ”-3 fungerer ikke. Man kan ikke dumpe mere eller mindre. Man skal ikke kunne få mindre end 00.” En bedømmer beskriver, at i og med at nogle bedømmere slet ikke anvender -3, betyder det i praksis, at der kun er tale om en 6-trinsskala, og at dette forrykker ved forståelsen af, hvad der er en middelkarakter.



Det er egentlig kun en 6-trinsskala. For vi bruger ikke -3, og vi giver også sjældent 00 i folkeskolen. Og så bliver 4 ikke middel. Når der er seks karakterer, er der ikke nogen middelkarakter. 7 bliver tit opfattet som middel.

Perspektiver på brugen af karakteren 12

Bedømmerne har ad forskellige veje hørt det beskrevet, at der gives for mange og over tid flere 12-taller. Det har de forskellige perspektiver på. Et perspektiv, der går igen, er, at det hænger sammen med beskrivelsen af 12-tallet, der gør, at et bredt udsnit af præstationer kan falde inden for kategorien 12. En bedømmer beskriver:



12-tallet er uden fejl og mangler. Det er blevet lidt udvandet. Fx havde vi en pige, som vi var i tvivl om. Hun havde ingen fejl og mangler, men jeg var ikke blæst bagover. Men vi kunne ikke give hende 10. Det er ikke et problem, men der er mange superstuderenter, så vi skal være klar over, at 12 ikke er exceptionelt – det kan almindelige mennesker få.

En anden bedømmer fortæller, at forståelsen blandt bedømmerne af, hvad der skal til for at få 12, er blevet justeret over tid.



Jeg har læst, at der går inflation i det. Jeg tror ikke, at eleverne er blevet klogere, men vi har sat barren lavere. I starten tænkte vi, at 12 var mellem 11 og 13, men nu er det mere bare 11.

En tredje bedømmer udtaler, at hvis man ønsker færre 12-taller, så må man indsnævre, hvad der skal til for at få 12, eksempelvis i trinbeskrivelsen:



Karakterbeskrivelserne sørger for, hvor mange der får 12. Man kan bare gøre 12 sværere. Fx med ord som ””udtømmende”” eller ””fremragende””.

Nogle bedømmere savner at kunne belønne den exceptionelle præstation

Nogle af bedømmerne oplever 12-tallet som for bredt og unuanceret og savner at kunne belønne den exceptionelle præstation, hvor eleven gør noget ud over det sædvanlige. To bedømmere siger:



Der burde være en karakter mere i toppen. Toppen burde gøres mindre bred, og det gælder også de andre karakterer. Den er ikke skarpt nok formuleret. 12 er reelt ikke en topkarakter, men dækker også en del af 10-tallet.



Jeg vil gerne kunne give en karakter for den ekstra gode præstation. Vi giver mange 12-taller, som ligner gamle 10-taller, fordi man har fejl. Topkarakter må være for toppræstation.

Bedømmerne efterlyser dog ikke nødvendigvis yderligere karakterer i toppen af skalaen, men mulighed for at kunne markere, når et 12-tal gives for en helt særlig præstation. En bedømmer siger: ”Jeg kan godt leve med, at der ikke er et 13-tal. Men der bliver hurtigt en top og en bund [i 12-tallet].” En anden bedømmer siger: ”Jeg kan savne den særlige karakter. Fx 12+, fordi jeg synes, de skal have noget for at gøre det ekstra.”

Nogle af bedømmerne har dog en opmærksomhed over for, hvad det vil betyde for fordelingen af karakterer, hvis der indføres en mulighed for at differentiere karakteren 12. De overvejer, om det i praksis vil svare til at indføre en ny karakter og kunne betyde, at der gives endnu flere topkarakterer. En bedømmer siger: ”Hvis vi indfører 12+, så får vi måske endnu flere 12-taller, fordi 12 så ikke er exceptionelt.”

6.7.2 Bedømmere på det gymnasiale område

Hovedpointerne ift. karakterskalaens udformning er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere fra de gymnasiale uddannelser. I dette afsnit går en del af overskrifterne fra afsnittet om fokusgrupperne således igen, men pointerne illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere oplever, at de store spring i midten af skalaen letter karaktergivning, andre, at de gør bedømmelsen sværere

Nogle bedømmere oplever, at de store spring i midten af skalaen letter karaktergivningen, mens andre bedømmere omvendt oplever, at karaktergivningen vanskeliggøres af store spring, og derfor oplever, at karaktererne i midten af skalaen er de sværeste at fastlægge. En bedømmer fortæller:



Hvis man er i tvivl om, om det skal være 4 eller 7, og skal blive enig med censor eller eksaminator, så er man nogle gange ude i noget næsten absurd, når man sidder og skal blive enige, for hvad er det lige, der skiller det ad? Det synes jeg, er det sværeste sted at finde ud af, hvad der er rimeligt, fordi det også gør en stor forskel med hensyn til gennemsnittet for den enkelte elev, om det er et 4- eller 7-tal.

En anden bedømmer forklarer, hvorfor middelkaraktererne kan være særligt svære at fastlægge:



Jeg plejer at kalde det mudder i midten. Det er, fordi de har opfyldt nogle faglige krav rimeligt, og nogle har de måske kun lige berørt. Men det er jo ikke nødvendigvis det samme for forskellige elever. Altså, man kan jo have tre elever, der hver især har været gode med hensyn til nogle forskellige faglige mål og dårlige med hensyn til nogle forskellige faglige mål. Hvad er det så lige, der gør, at man skal vælge et 7-tal?

Endelig beskriver nogle bedømmere, at springenes størrelse og karakterernes talværdier ikke har nogen betydning for deres karaktergivning, men alene for elevernes karaktergennemsnit. En bedømmer siger:



Skalaens trin har ikke nogen indflydelse på min bedømmelse. Jeg kunne lige så godt tænke det som a, b og c. Talværdierne har kun noget at sige, når der skal beregnes gennemsnit. Karaktererne er kun interessante i en gennemsnitssammenhæng.

Bedømmere savner at kunne nuancere og differentiere i midten af skalaen

En gruppe af bedømmerne savner at kunne nuancere og differentiere karaktergivningen mere i midten af skalaen, fx i form af en ekstra karakter, blandt andet af hensyn til, hvor retvisende karaktererne er, og af hensyn til elevernes karaktergennemsnit. En bedømmer siger:



Det er problematisk med de store spring i midten, fordi det har store konsekvenser for det samlede karaktergennemsnit, om man får 4 eller 7. Uanset hvor gerne man vil skille fårene fra bukkene, så ligger de fleste i midten, og det er rigtig svært at få de præstationer skilt ad, for der er nuancer, og man mangler nogle bedre middelkarakterer end 7 og 4. Til eksamen får man givet dem lidt for meget eller lidt for lidt, fordi de tit ligger midt imellem. Så jeg savner nogle middelkarakterer. Primært mellem 7 og 10, men også mellem 4 og 7.

En anden bedømmer påpeger, at de store spring kan være med til at presse karaktererne op. Hun siger:



De store spring presser en til at give lidt højere karakterer. Man vil jo ikke give en lavere karakter end det, der er fortjent – så vil man hellere gå op end ned, og man vil hellere lade det komme eksaminanden til gode.

Der er dog også bedømmere, der ikke savner flere karakterer i midten, men som, som før nævnt, oplever, at få, brede karakterer i midten gør karaktergivningen nemmere og er uproblematisk. En bedømmer siger:



Rigtig mange af mine kollegaer siger, at de mangler 9. For nogle gange er det måske ikke helt til 10, men bedre end den gennemsnitlige 7. Men det synes jeg ikke. For enten ligger det jo tættest på 10, eller også ligger det tættest på 7. Så jeg synes faktisk, det er nemmere.

-3 er en undtagelseskarakter med stor symbolsk betydning

Bedømmerne beskriver generelt -3 som en karakter, der meget sjældent kommer i brug, fordi bedømmerne oplever, at det hører til sjældenhederne, at elevernes præstation er til -3. Ved 00 har der været et forsøg, mens der ved -3 intet forsøg på faglighed har været. Bedømmerne beskriver, at -3 næsten aldrig optræder som standpunkts- og årskarakter, fordi disse karakterer gives for elevens præstationer over en længere periode, og der stort set altid vil være noget fra eleven, der berettiger til en karakter over -3. Derudover beskriver bedømmerne også, at -3 er særligt sjælden ved mundtlig eksamen, fordi eksaminator her har mulighed for at stille spørgsmål til eleven, og man derfor som regel kan finde noget, hvor eleven kan sige noget. Mens man til en skriftlig eksamen ikke på samme måde kan udforske elevens faglighed, men alene har det, eleven har skrevet i opgaven, at bedømme ud fra.

Men når -3 sjældent kommer i brug, handler det for nogle bedømmere også om, at de slet ikke anvender karakteren. En bedømmer siger: "Jeg overvejer den faktisk ikke. Jeg tænker faktisk ikke på den. I mit hoved har jeg faktisk kun seks karakterer." For en del af bedømmerne drejer det sig om, at de ikke bryder sig om det signal, de oplever, det er at give en minuskarakter. En bedømmer siger:



Jeg synes, at -3 er pædagogisk ganske forfærdelig. Jeg synes, at der er noget helt pædagogisk forkert i, at man er dummere, når man går ud, end når man kommer ind. Man må sige, at det maksimale lavpunkt er 00.

Andre bedømmere mener, at det er vigtigt, at man bruger hele den skala, man har til rådighed. En bedømmer siger:



I starten var folk meget forsigtige ift. -3. Der var nogle, der af princip ikke ville give -3. Hvilket også var lidt interessant, for når vi har en skala, så skal vi bruge den. Men det synes jeg henover årene, at der er kommet et lidt afslappet forhold til – at man kan faktisk godt kan se forskel på, om det er 00 eller -3.

Dumpegrænsen opleves generelt som tydelig, men vigtig at diskutere ekstra grundigt

Bedømmerne oplever generelt dumpegrænsen som tydelig. Alligevel fortæller bedømmerne, at det er omkring dumpegrænsen, at censor og eksaminator drøfter bedømmelsen mest. Det skyldes dog ikke uklarhed med hensyn til dumpegrænsen, men at det har så stor betydning for eleven, om den pågældende dumper eller ej, at bedømmerne mener, at det er vigtigt at bruge ekstra tid på at være ekstra grundig i gennemgangen af, hvilke argumenter der taler for hhv. dumpet og bestået, og undersøge, om der er noget, der gør, at eleven ikke skal dumpe.

Ifølge dumpegrænsens tydelighed oplever en bedømmer dog trinbeskrivelsens formuleringer om dumpegrænsen som lidt vage, men er samtidig ikke sikker på, at de kan gøres mere præcise, når trinbeskrivelsen skal være generelt gældende:



Man kan godt være usikker på dumpegrænsen. "Tilstrækkelig" og "utilstrækkelig". Det bliver lidt vagt. Men jeg ved ikke, om det kan være anderledes. I nogle fag kan man i højere grad lave pointgivning, men ikke i mit.

Bedømmerne har forskellige bud på, hvad stigningen i andelen af 12-taller kan skyldes

Nogle bedømmere oplever, at de selv har givet flere 12-taller over tid, mens andre bedømmere oplever, at antallet af 12-taller for deres eget vedkommende er uændret over tid. Bedømmerne vurderer, at stigningen i andelen af 12-taller blandt andet skal ses i sammenhæng med, at bedømmernes forståelse af skalaen har ændret sig, men også at elevernes tilgang har ændret sig. For det første fortæller nogle bedømmere nemlig, at de oplever over tid at have fået en bedre forståelse af karakterskalaen og herunder af, hvad der skal til for at udløse et 12-tal. En bedømmer udtaler:



Jeg giver måske lidt flere 12-taller. Jeg tror, det kan skyldes større erfaring med skalaen. Vi har på skolerne og lærerne imellem udviklet en fælles forståelse af skalaen og af, at der godt må være fejl ved 12-tallet. Det er ikke den helt perfekte præstation.

Samtidig oplever nogle bedømmere også, at fordi gennemsnittet af mange elever i højere grad opleves som vigtigt nu end førhen, så er flere elever også villige til at gøre en ekstra indsats af hensyn til karaktererne. En bedømmer fortæller:



Elevernes indstilling har ændret sig. For mange elever er gennemsnittet pludselig blevet vigtigt, og de vil derfor godt gøre den ekstra indsats, det kræver at stige fra 10 til 12. Ikke nødvendigvis af interesse for faget, men af hensyn til gennemsnittet. Jeg giver flere af den type 12-taller end tidligere. Det handler ikke om, at elevernes bevidsthed har ændret sig, men om, at de krav, der bliver stillet til deres gennemsnit, har ændret sig. Eleverne har hele tiden været opmærksomme på de krav, der bliver stillet, men kravene har ændret sig.

Nogle bedømmere savner at kunne belønne den exceptionelle præstation

En del af bedømmerne savner at kunne belønne den exceptionelle præstation, hvor eleven præsterer ud over det sædvanlige og forventelige. En bedømmer siger:



Som det er nu i matematik, så tæller vi point sammen, og når vi når op over en bestemt tærskel, så siger vi 12. Jeg kunne godt tænke mig, at jeg kunne belønne den elev, som ikke bare regnede korrekt, men som også kom bag om opgaven, hvor det blev mere abstrakt og perspektiverende. Jeg savner at kunne markere den helt exceptionelle elev.

En anden bedømmer udtaler, at selvom hun generelt finder store spring på skalaen problematiske, så er et stort spring netop meningsfuldt, når det handler om at markere den exceptionelle præstation. Hun siger:



Jeg savner at kunne belønne, at der er elever, der er uovertrufne og leverer præstationer på universitetsniveau. Der, hvor man får meget mere, end man egentlig kunne forvente af en eksaminand. Der savner jeg 13-tallet – og der havde det der spring mellem 11 og 13 sin berettigelse, fordi det er over alt det andet.

Nogle bedømmere understreger dog, at godt nok savner de at kunne belønne den ekstraordinære præstation, men i dagligdagen er det ikke et problem, da de ekstraordinære præstationer trods alt er sjældne. Derfor peger nogle bedømmere på, at det er et større problem med manglende nuancer i midten af skalaen, da rigtig mange elever befinder sig i midten af skalaen.

6.7.3 Bedømmere på de videregående uddannelser

Hovedpointerne ift. karakterskalaens udformning er i stor udstrækning overlappende på tværs af fokusgrupper og individuelle interviews med bedømmere fra de videregående uddannelser. I dette afsnit vil der derfor være pointer, der går igen, men de illustreres og nuanceres her med udsagn fra bedømmerne i de individuelle interviews.

Nogle bedømmere oplever, at de store spring i midten af skalaen letter karaktergivningen, andre, at de gør bedømmelsen sværere

Bedømmerne oplever springene i midten af skalaen mellem 4, 7 og 10 som store i kraft af deres numeriske værdi og dermed trinene som brede. Men bedømmerne har forskellige perspektiver på, hvorvidt de brede karakterer i midten af skalaen letter eller vanskeliggør karaktergivningen. På den ene side oplever nogle bedømmere, at karaktergivningen bliver lettere, og enigheden mellem censor og eksaminator større, når karaktererne er brede, og der skal differentieres mindre:



Jeg kan bedre lide den her skala, end jeg kunne lide den gamle. Og det kan jeg, fordi når springene er så store, så er det vigtigt, at den studerende får den rigtige karakter. Der var ikke så stor forskel på, om man fik 8 eller 9, men der er stor forskel på, om man får 7 eller 10. Det vil sige, at det tvinger mig som vejleder til at være meget skarp på, hvad for læringsmål der er, og hvordan den studerende har nået dem. Hvad er argumenterne for, at det her så er et 7-tal eller et 10-tal? Det synes jeg, er en stor kvalitet ved den her skala. Det er en kvalitet, at det ikke bliver ”lala” at give karakter. Sådan ”nåh ja, er det et stort 8-tal eller et lille 9-tal?”.

På den anden side oplever andre bedømmere, at netop det, at trinene i midten af skalaen er så brede, kan gøre det vanskeligt at afgøre, om der er tale om den ene eller den anden karakter, og om en studerendes præstation eksempelvis er til et stort 7-tal eller et lille 10-tal, fordi karaktererne kan rumme meget forskelligartede præstationer. En bedømmer fortæller:



De [springene] gør selvfølgelig, at der er større diskussion og overvejelser. Også fordi 7 ikke nødvendigvis bliver set som en god karakter, selvom det er den gode præstation. Det er primært blandt de studerende, men måske også lidt blandt os selv. Også fordi det er så stort et spænd, og vi måske har siddet med en opgave inden, der også gav et 7-tal. Så giver det anledning til en længere votering, at der er de her store spring.

Bedømmerne i de individuelle interviews oplever generelt ikke udfordringer ved springet mellem 4 og 7:



Der er oftere nogle, der ligger på vippen mellem 7 og 10, end der er mellem 4 og 7. Tit, hvis det er en præstation til 4, så er man faktisk sikker på det. Også fordi man ikke kan lade være med at tænke på den gamle skala. Der var jo kun én karakter der.



4-tallet ligger jo fuldstændig op ad det gamle 7-tal. Jeg synes ikke, at udfordringen ved brugen af 4-tallet er stor, og jeg synes også, at der er en fin forståelse for, hvornår noget er 4, og hvornår noget er 2. Der ser jeg heller ikke nogen stor udfordring.

Nogle bedømmere efterlyser at kunne differentiere i midten af skalaen

Nogle af de interviewede bedømmere efterlyser i højere grad at kunne nuancere og differentiere karaktergivning i midten af skalaen ved at have flere forskellige karakterer at vælge mellem:



Og ligeledes synes jeg også nogle gange – og jeg har også snakket med mine kollegaer om det, som siger det samme – at spændet mellem 7 og 10 er rigtig stort. Den 7”er dækker bare over rigtig meget. Og der kunne vi godt nogle gange savne, at der var noget indimellem der. Fordi der er en stor gruppe, som performer meget forskelligt, som ligger i den gruppe.



Jeg vil sige, at savnet af 8-tallet også er et, jeg har mærket. Men det er også en del af konceptet, at der skal være klart markerede spring. De studerendes præstationer i den virkelige verden placerer sig langs et kontinuum. Det bliver på en måde skærpet af talværdierne. Hvis man havde gjort det, som jeg synes, var det rigtige at gøre – altså det at køre med ECTS-skalaens bogstaver – så havde man fået en jævner fordeling. Så havde man ikke – med de samme karaktertrin – haft den her oplevelse af, at der er nogle studerende, der får en karakter, som er urimeligt hård, og andre, hvor det virker urimeligt rundhåndet.

-3 er en undtagelseskarakter

Bedømmerne og censorerne fortæller generelt, at -3 er en karakter, de sjældent benytter. Nogle benytter den en gang imellem, andre benytter den aldrig. Det går igen i interviewene, at -3 særligt sjældent kommer i brug ved mundtlig bedømmelse, da man her kan have en dialog med den studerende, og at man kun yderst sjældent oplever studerende, der decideret nægter at sige noget.



Jeg bruger kun -3 til folk, der forlader eksamen. Hvis de trækker et spørgsmål, er eksamen påbegyndt. Hvis de kommer fra forberedelseslokalet, og klappen er gået ned, så får de -3. Hvis de gør et forsøg på at bestå eksamen, så mener jeg ikke, at der er grund til at føje spot til skade. Der kan man sige, at dem, der ingenting siger, heller ikke har noget problem med at få -3. Der er det okay for dem. Men hvis man har studerende, der faktisk prøver at gennemføre ... der tror jeg ikke, at man har meget gavn af at give dem -3. Det virker for lektor Blomme-agtigt. Det er okay at have en karakter til dem, der slet ingenting vil.



Jeg synes, at -3 er åndssvagt. Den giver jeg kun, hvis de studerende ingenting har skrevet. Ellers er det en fornærmelse at give dem -3.



Ved mundtlige eksaminer i forlængelse af en hovedopgave har jeg aldrig oplevet det. Ved mundtlige eksaminer, der har jeg oplevet -3 få gange. Det opstår i den situation, hvor jeg har den her studerende, der ganske enkelt ikke formår at trække i den rigtige retning. Så den anvendes helt klart.

Bedømmerne fortæller, at -3 kommer mere i spil ved skriftlige prøver:



Karakteren -3 kommer i spil ved de skriftlige eksaminer. Der skal jo noget til for at få 00. Når vi har de her skriftlige eksaminer, hvor der er sat point på hver opgave, og vi anvender en omsætningstabel, så er det ret enkelt.



-3 bruger jeg, når de intet har skrevet. Ellers bruger jeg som regel ikke -3. Og 00 bruger jeg i forbindelse med skriftlige opgaver, når de har scoret under 50 % af pointene.

En bedømmer fortæller, at han oplever, at -3 anvendes til at signalere, at det ikke er inden for rækkevidde at bestå faget, og at den studerende måske endda bør genoverveje sit studievalg:



Jeg har brugt -3 tre gange i min karriere som underviser [ni år]. Det er blevet brugt til at signalere, at den studerende ikke bare var dumpet, men at det var seriøst dårligt, det, den studerende leverede. Jeg var engang censor sammen 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.

Perspektiver på anvendelse af karakteren 12

Et gennemgående perspektiv i interviewene med bedømmerne og censorerne er, at 12-tallet er en meget bred karakter, der kan dække over meget forskellige præstationer. Der er dog forskellige perspektiver på, hvorvidt der mangler en karakter for det exceptionelle:



Jeg savner denne her karakter for det exceptionelle. Jeg havde eksamen sidste weekend, og det var simpelthen så exceptionelt. Vedkommende får så et 12-tal. Så kommer den næste studerende ind, og det er bare rigtig godt, men lige med nogle få uvæsentlige mangler, og får 12. Der ønskede både censor og eksaminator at kunne differentiere oppe i toppen af skalaen. Det kan man ikke, men det er måske heller ikke det, man har ønsket fra statens side.



Jeg har det lidt blandet med 12. Jeg synes, spændet for dem, der har fået 12 til denne eksamen, er rigtig stort. Jeg ville ikke ønske mig tilbage til den gamle karakterskala, for der var også uforholdsmæssigt store spring efter 9, 10, 11 og så op til 13. Jeg savner ikke genialitetskriteriet i denne sammenhæng, men jeg kunne godt ønske mig, at der var en mellemkarakter. En 11"er.



Det er også udmærket, at 13-tallet er væk, fordi jeg aldrig nogensinde gav 13. Det er mærkeligt at have en karakter, man aldrig giver.

En bedømmer fortæller, at hun ikke ønsker sig 13-tallet tilbage, men at hun godt kunne tænke sig, at der var en måde at belønne de exceptionelt dygtige studerende på:



Der er en enkelt ting, som vi måske også savner nogle gange. Og nu kan jeg kun sammenligne med den gamle skala, men det der med, at der er en højere karakter end 12-tallet, hvor der er plads til mangler. For en gang imellem har vi en studerende, der virkelig leverer noget helt ud over det sædvanlige og det forventelige. Altså ud over de rammer, som vi måler ud fra, og der ville det være rigtig rart at kunne honorere det på en eller anden måde. Enten gennem en karakter, eller også er vi selv nødt til at lave noget, så de kan få en eller anden distinction. Det er vi i hvert fald begyndt at snakke om. For der mangler vi at kunne honorere det.

En anden ting, som kommer frem under interviewene, er, hvordan det med få uvæsentlige mangler skal fortolkes, hvilket illustreres i de to citater nedenfor:



Hvis der er én alvorlig fejl, kan de ikke få 12. Men problemet er: 1) Hvornår er noget en alvorlig fejl, og 2) hvor mange mellemstore fejl må der være, før de summer op til en stor fejl? Man får 12 til de skriftlige, hvis man har over 90 % rigtige. Jeg oplever nogle censorer, der har villet give 10 til dem, der har haft 89,5 % rigtige. Det synes jeg, er hårdt.



Vi havde diskussionen før, fordi vi havnede på et 12-tal. Og så var der også den der med: ”Er det, som mangler i den her præstation, få og ubetydelige mangler, eller er det ikke?”. Og det blev vi så enige om, at det ikke var. Der var nogle mangler, men de var få og ubetydelige.

7 Belønning af særligt dygtige studerende

7.1 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 følgende undersøgelsesspørgsmål:

- Er der tilstrækkelige muligheder for at identificere særligt dygtige studerende og belønne den ekstraordinære præstation?

7.2 Survey blandt bedømmere på de videregående uddannelser

Tabel 7.1 viser, hvad bedømmerne på de videregående uddannelser svarer på spørgsmålet "Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen? Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet på 13-skalaen". I det følgende oplystes andelen, der er enige i udsagnet (enten "Meget enig" eller "Enig"), for hver sektor:

- 84 % på de kunstneriske uddannelser er enige.
- 72 % på professionsbacheloruddannelserne er enige.
- 72 % på erhvervsakademiuddannelserne er enige.
- 69 % på universitetsuddannelserne er enige.
- 49 % på de maritime uddannelser er enige.

Især bedømmere på de kunstneriske uddannelser er enige i udsagnet. Især bedømmere på de maritime uddannelser er ikke enige. Variationer i svar mellem bedømmerne fra de forskellige hovedområder er at finde i den særskilte tabelrapport. Andelen, der har svaret "Enig" i forbindelse med udsagnet "Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet på 13-skalaen", er ikke statistisk signifikant forskellig på tværs af de forskellige hovedområder.

TABEL 7.1

Hvor enig eller uenig er du i følgende udsagn om 7-trins-skalaen? Skalaen burde have en mulighed for at honorere den ekstraordinært gode præstation, som svarer til 13-tallet på 13-skalaen

	Meget enig	Enig	Hverken enig eller uenig	Uenig	Meget uenig	Ved ikke	Total
Universitetsuddannelser	47 % (861)	22 % (405)	12 % (215)	13 % (231)	5 % (97)	2 % (30)	100 % (1.839)
Professionsbacheloruddannelser	47 % (325)	25 % (175)	12 % (83)	10 % (70)	5 % (34)	1 % (7)	100 % (694)
Erhvervsakademiuddannelser	53 % (130)	19 % (48)	9 % (22)	11 % (28)	6 % (15)	2 % (4)	100 % (247)
Kunstneriske uddannelser	47 % (9)	37 % (7)	10 % (2)	7 % (1)	0 % (0)	0 % (0)	100 % (19)
Maritime uddannelser	39 % (7)	10 % (2)	23 % (4)	19 % (4)	10 % (2)	0 % (0)	100 % (19)
Total	47 % (1.333)	23 % (638)	12 % (326)	12 % (333)	5 % (147)	1 % (41)	100 % (2.819)

Kilde: EVA, survey maj 2018 blandt bedømmere på de videregående uddannelser.

7.3 Fokusgruppeinterviews med bedømmere

7.3.1 Nogle bedømmere savner at kunne belønne den exceptionelle præstation

Nogle af bedømmerne savner at kunne belønne den exceptionelle præstation, hvor den studerende gør noget ud over det sædvanlige. Bedømmerne beskriver 12-tallet som en meget bred karakter, der kan dække over meget forskellige præstationer. Det, bedømmerne savner, er derfor 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 nogen væsentlige fejl og mangler og lever op til målbeskrivelsen, og de studerende, der får 12, fordi de leverede en helt særegen præstation. En bedømmer udtaler: "Jeg savner noget i toppen. Der kan komme nogle, der præsterer godt til 12, men så kommer der en, der nærmest sprænger skalaen." En anden bedømmer forklarer:



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.

For andre bedømmere handler det ikke om, at de savner en egentlig karakter for den exceptionelle præstation, såsom 13-tallet, men at de på anden vis gerne vil kunne markere, når den studerende leverer en exceptionel præstation, og at det alene skal være ved topkarakteren, man kan lave en sådan markering og differentiering. En bedømmer siger: "Der er også forskel på dem, som får 12. Nogle brillerer og fortjener et 12+." To bedømmere siger:



Man kunne godt bruge plus. Det skal ikke tælle med. Det skal være sammenligneligt med det internationale karaktersystem.



Hvis jeg sad til påskefrokost, ville jeg nok også være for et 13-tal. Men i dette forum må jeg sige, at jeg hellere ville give de 5 % bedste en distinktion a la den der cum laude. Noget, som hjalp dem, når de skulle søge videre. Men jeg savner virkelig ikke 13-tallet.

Nogle bedømmere peger på, at behovet for at kunne markere den exceptionelle præstation er større ved nogle typer af prøver end andre. En bedømmer siger:



13-tallet var misforstået. Hvor genial kan man være på de kurser, jeg underviser i? Det er rart, at vi ikke har det længere. Men ved afgangsprojektet, der kunne jeg godt ønske mig at kunne give et klap på skulderen.

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 med 13-tallet, men ud 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 dækker både 11 og 13.

Mulighederne i talentbekendtgørelsen anvendes i meget varierende grad

Bekendtgørelse om talentinitiativer på de videregående uddannelser på Uddannelses- og Forskningsministeriets område (talentbekendtgørelsen) giver en række muligheder for at belønne særligt dygtige studerende:

- Udmærkelse på eksamensbeviset
- Anerkendelse af ekstracurriculære aktiviteter på eksamensbeviset
- Ekstra ECTS-point.

Der er få eksempler i fokusgruppeinterviewene med bedømmerne på anvendelse af de muligheder, der ligger i talentbekendtgørelsen. En bedømmer på en universitetsuddannelse fortæller, at han lige er blevet opmærksom på muligheden for udmærkelser på eksamensbeviset:



Jeg har lige fået at vide, at man kan give en stjerne. Det er helt nyt for mig. Men jeg vil også sige, at 13-tallet vil skabe problemer for mange studerende, der vil til udlandet. Jeg vil meget gerne bruge den stjerne.

En anden bedømmer fra en professionsbacheloruddannelse fortæller, at studerende, der kobles til forskningsprojekter, får ekstra ECTS-point på eksamensbeviset.

Der er også et par eksempler på andre muligheder, der tages i anvendelse for at belønne særligt dygtige studerende, og som ikke er en del af talentbekendtgørelsen. Eksempelvis fortæller en underviser, at man uddeler legater i samarbejde med virksomheder. Bedømmeren fremhæver dog, at der er lige så meget fokus på den studerendes personlighed som på det faglige i uddelingen. En anden bedømmer fortæller, at man har workshops målrettet de særligt dygtige, hvor de studerende modtager et bevis på deres deltagelse.

Appendiks A – Metode

Denne bilagsrapport bygger på følgende datakilder:

- Spørgeskemaundersøgelse blandt bedømmere på grundskoleområdet, ungdomsuddannelserne og de videregående uddannelser
- Fokusgruppeinterviews med bedømmere fra grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser
- Enkeltinterviews med bedømmere fra grundskoleområdet, de gymnasiale uddannelser og de videregående uddannelser.

Nedenfor gennemgås de enkelte datakilderne.

Spørgeskemaundersøgelse

Formålet med spørgeskemaundersøgelsen er at få indsigt i omfanget/udbredelsen af forskellige elementer i karaktergivning. Det gøres ved at indhente besvarelser på standardiserede spørgeskemaer, hvormed besvarelserne kan sammenlignes mellem forskellige grupper af respondenter.

Kontaktgrundlag

Undersøgelsen er gennemført blandt bedømmere (undervisere og censorer) fra:

- Grundskoler
- Institutioner med almen gymnasiale uddannelser (stx og hf)
- Institutioner med erhvervsrettede uddannelser (EUD, hhx og htx)
- Erhvervsakademier
- Professionshøjskoler
- Universiteter
- Kunstneriske uddannelsesinstitutioner
- Maritime uddannelsesinstitutioner

Kontaktoplysningerne for undervisere blev fremsendt af uddannelsesinstitutionerne, mens kontaktoplysningerne for censorer på de videregående uddannelser blev fremsendt af censorkorpserne, og kontaktoplysningerne for de mundtlige censorer på grundskolerne, de skriftlige censorer på gymnasierne og censorerne på EUD blev fremsendt af Undervisningsministeriet.

Forud for dataindsamlingen fik grundskolerne, gymnasierne og institutionerne med erhvervsrettede uddannelser tilsendt et varslingsbrev fra Undervisningsministeriet med en introduktion til undersøgelsen, samt information om hvilke kontaktoplysninger der skulle fremsendes til Epinion. Ligeledes fremsendte Uddannelses- og Forskningsministeriet et varslingsbrev til de videregående uddannelsesinstitutioner og censorkorpse.

Tabellen nedenfor viser antallet af institutioner, der indberettede kontaktoplysninger for deres undervisere. Af tabellen ses det, at 55 % af de udvalgte institutioner indberettede kontaktoplysninger.

TABEL A.1

Fremsendte kontaktoplysninger for undervisere

Sektor	Antal institutioner	Antal indberetninger	Andel
Grundskoler	400	194	49 %
Gymnasier	157	99	63 %
Erhvervsrettede uddannelser (EUD, hhx og htx)	80	47	59 %
Erhvervsakademier	9	9	100 %
Professionshøjskoler	8	7	88 %
Universiteter	8	8	100 %
Kunstneriske uddannelsesinstitutioner	4	3	75 %
Maritime uddannelsesinstitutioner	10	5	50 %
Total	676	372	55 %

Følgende 7 videregående uddannelsesinstitutioner deltog ikke i undersøgelsen:

- Professionshøjskoler: Den Frie Lærerskole
- Kunstneriske uddannelsesinstitutioner: Kaospiloterne
- Maritime uddannelsesinstitutioner: Marstal Navigationsskole, Nyborg Søfartsskole, Skagen Skiperskole, Skoleskibet "Danmark" og Svendborg Søfartsskole

Tabellen nedenfor viser antallet af censorkorps, der indberettede kontaktoplysninger for deres censorer. Af tabellen ses det, at 75 % af censorkorpse indberettede kontaktoplysninger.

TABEL A.2

Fremsendte kontaktoplysninger for censorer på de videregående uddannelsesinstitutioner

Censorkorps	Antal korps	Antal indberetninger	Andel
Erhvervsakademier	9	7	78 %
Professionshøjskoler	43	32	74 %
Universiteter	101	76	75 %
Total	153	115	75 %

For bedømmere på de videregående uddannelsesinstitutioner modtog Epinion i alt kontaktoplysninger på 31.553 personer. Heraf blev der udtrukket en simpelt tilfældig stikprøve på 5.000 undervisere og 2.500 censorer, som blev inviteret til at deltage i undersøgelsen.

Nedenstående tabel viser undersøgelsens samlede udsendelsesgrundlag. I alt blev 14.177 bedømmere inviteret til at deltage i undersøgelsen.

TABEL A.3

Udsendelsesgrundlag for bedømmere

Sektor	Antal personer
Grundskoler	2.266
Gymnasier	3.228
Erhvervsrettede uddannelser (EUD, hhx og htx)	1.183
Erhvervsakademier	732
Professionshøjskoler	1.744
Universiteter	4.751
Kunstneriske uddannelsesinstitutioner	58
Maritime uddannelsesinstitutioner	45
Tværgående, videregående uddannelsesinstitutioner	170
Total	14.177

Pilottest

D. 14.-18. maj 2018 gennemførte Epinion kvalitative pilottests blandt 13 bedømmere fordelt i tabellen nedenfor. Pilotdeltagerne fik alle tilsendt et unikt link til det webbaserede spørgeskema, hvor de løbende kunne kommentere på spørgsmålene undervejs i deres besvarelse. Umiddelbart efter, at de havde besvaret spørgeskemaet, er de blevet ringet op af den ansvarlige konsulent, som herefter spurgte ind til de mest relevante problemstillinger. Resultaterne fra pilottesten blev samlet i et notat, og på baggrund heraf blev der foretaget en række ændringer i spørgeskemaet forud for dataindsamlingen.

TABEL A.4

Antal pilotdeltagere opdelt på sektor, institutioner og fag

Sektor	Institution	Fag	Antal
Grundskoler - Folkeskoler	Kirkebakkeskolen	Dansk, engelsk	1
Grundskoler - Folkeskoler	Hornslet Skole	Dansk, engelsk	1
Grundskoler - Folkeskoler	Højvangskolen	Engelsk, matematik	1
Gymnasier - stx	Svendborg Gymnasium	Dansk, engelsk	1
Gymnasier - hf	Nordvestsjælland HF & VUC	Engelsk	1
Gymnasier - htx	EUC Nord	Matematik	1

Sektor	Institution	Fag	Antal
Gymnasier – hhx	IBC International Business College	Dansk, engelsk	1
EUD – SOSU-skole	Randers Social- og Sundhedsskole	Dansk, engelsk	1
EUD – Teknisk skole	AMU Nordjylland	Matematik	1
EUD – Handelsskole	IBC International Business College	Engelsk	1
Erhvervsakademiuddannelser	Cphbusiness	Økonomisk-Merkantil	1
Professionsbacheloruddannelser	Københavns Erhvervsakademi	Optometri	1
Universitetsuddannelser	Aarhus Universitet	Informationsvidenskab	1
Total			13

Dataindsamling

D. 29. maj blev der sendt invitationsmails til de 14177 bedømmere. D. 4. juni blev der udsendt e-mailpåmindelser til 11281 bedømmere, som ikke havde besvaret spørgeskemaet. Igen d. 7. juni blev der sendt yderligere en e-mailpåmindelse til 9754 bedømmere.

Svarprocenter

I alt deltog 5.112 bedømmere i undersøgelsen, og dermed blev der opnået en samlet svarprocent på 36 %.

TABEL A.5

Samlede svarprocenter for bedømmere

Sektor	Kontaktgrundlag	Svar	Svarprocent
Grundskoler	2.266	665	29 %
Gymnasier	3.228	1.029	32 %
Erhvervsrettede uddannelser (EUD, hhx og htx)	1.183	472	40 %
Erhvervsakademiuddannelser	732	291	40 %
Professionsbacheloruddannelser	1.744	866	50 %
Universitetsuddannelser	4.751	1.662	35 %
Kunstneriske uddannelser	58	30	52 %
Maritime uddannelser	45	31	69 %
Tværgående videregående uddannelser	170	66	39 %
Total	14.177	5.112	36 %

TABEL A.6

Svarprocenter for bedømmere på erhvervsrettede uddannelser

Skoletype	Kontaktgrundlag	Svar	Svarprocent
AMU-center	6	3	50 %
Handelsskole	319	122	38 %
Kombinationsskole	546	225	41 %
Landbrugsskole	27	5	19 %
SOSU-skole	122	62	51 %
Teknisk skole	163	55	34 %
Total	1.183	472	40 %

TABEL A.7

Svarprocenter for bedømmere på Erhvervsakademier

Erhvervsakademier	Kontaktgrundlag	Svar	Svarprocent
Cphbusiness	72	34	47 %
Erhvervsakademi Dania	24	10	42 %
Erhvervsakademi Kolding	19	9	47 %
Erhvervsakademi MidtVest	8	5	63 %
Erhvervsakademi Sjælland	46	22	48 %
Erhvervsakademi Sydvest	26	12	46 %
Erhvervsakademi Aarhus	87	45	52 %
Erhvervsakademiet Lillebælt	71	34	48 %
Københavns Erhvervsakademi	70	41	59 %
Total	423	212	50 %

TABEL A.8

Svarprocenter for bedømmere på Professionshøjskoler

Professionshøjskoler	Kontaktgrundlag	Svar	Svarprocent
Danmarks Medie- og Journalisthøjskole	12	8	67 %
Københavns Professionshøjskole	273	136	50 %
Professionshøjskolen Absalon	122	64	52 %
Professionshøjskolen Lillebælt University College	78	48	62 %
Professionshøjskolen UC Syddanmark	82	51	62 %
Professionshøjskolen University College Nordjylland	150	83	55 %
VIA University College	273	153	56 %
Total	990	543	55 %

TABEL A.9

Svarprocenter for bedømmere på UNIVERSITETER

Universiteter	Kontaktgrundlag	Svar	Svarprocent
Copenhagen Business School	295	91	31 %
Danmarks Tekniske Universitet	211	89	42 %
IT-Universitetet	48	13	27 %
Københavns Universitet	955	346	36 %
Roskilde Universitet	190	67	35 %
Syddansk Universitet	387	149	39 %
Aalborg Universitet	563	217	39 %
Aarhus Universitet	1302	329	25 %
Total	3951	1301	33 %

TABEL A.10

Svarprocenter for bedømmere på kunstneriske uddannelsesinstitutioner

Kunstneriske uddannelsesinstitutioner	Kontaktgrundlag	Svar	Svarprocent
Arkitektskolen Aarhus	22	13	59 %
Designskolen Kolding	5	4	80 %
KADK	31	13	42 %
Total	58	30	52 %

TABEL A.11

Svarprocenter for bedømmere på maritime uddannelsesinstitutioner

Maritime uddannelsesinstitutioner	Kontaktgrundlag	Svar	Svarprocent
Fredericia Maskinmesterskole	11	6	55 %
MARTEC	9	7	78 %
Maskinmesterskolen København	7	5	71 %
SIMAC	5	3	60 %
Aarhus Maskinmesterskole	13	10	77 %
Total	45	31	69 %

Vægtning

Data for bedømmere på de videregående uddannelsesinstitutioner er efterfølgende vægtet efter hovedområde.

Nedenstående tabel viser fordelingerne inden for hovedområder samt vægtene.

TABEL A.12

Vægtning

Hovedområde	Populationsfordeling	Svarfordeling	Vægt
Erhvervsakademiuddannelser	8,40 %	9,30 %	
Bio- og laboratorietekniske område	0,34 %	0,58 %	0,593
Designfaglige område	0,61 %	0,64 %	0,939
It-faglige område	1,14 %	2,04 %	0,560
Samfunds-faglige område	0,96 %	0,64 %	1,489
Tekniske område	2,92 %	2,14 %	1,365
Økonomisk-merkantile område	2,43 %	3,26 %	0,746
Professionsbacheloruddannelser	23,57 %	29,40 %	
Design	0,41 %	0,31 %	1,349
Medie, kommunikation, it mv.	0,78 %	0,64 %	1,214

Hovedområde	Populationsfordeling	Svarfordeling	Vægt
Pædagogik	8,41 %	10,59 %	0,794
Samfund	1,86 %	2,31 %	0,805
Sundhed	5,97 %	7,98 %	0,749
Teknik	5,49 %	6,42 %	0,856
Økonomisk/merkantil	0,65 %	1,15 %	0,560
Universitetsuddannelser	62,42 %	56,14 %	
Humaniora	12,92 %	13,95 %	0,926
Naturvidenskab	11,31 %	7,06 %	1,602
Samfundsvidenskab	18,59 %	18,23 %	1,020
Sundhedsvidenskab	12,98 %	10,25 %	1,267
Teknisk videnskab	6,61 %	6,65 %	0,994
Kunstneriske	0,65 %	1,02 %	0,638
Maritime	0,63 %	1,05 %	0,602
Tværgående	4,33 %	3,09 %	1,401
Total	100 %	100 %	

Nedenstående tabel viser populationsfordelingerne samt de uvægtede- og vægtede fordelinger på institutionsniveau.

TABEL A.13

Fordelinger på institutionsniveau

Institution	Populationsfordeling	Svarfordeling	Vægtet svarfordeling
Erhvervsakademier	4,04 %	7,20 %	6,02 %
Cphbusiness	0,69 %	1,15 %	0,79 %
Erhvervsakademi Dania	0,28 %	0,34 %	0,26 %
Erhvervsakademi Kolding	0,14 %	0,31 %	0,25 %
Erhvervsakademi MidtVest	0,13 %	0,17 %	0,18 %
Erhvervsakademi Sjælland	0,49 %	0,75 %	0,59 %
Erhvervsakademi Sydvest	0,27 %	0,41 %	0,39 %
Erhvervsakademi Aarhus	0,74 %	1,53 %	1,18 %
Erhvervsakademiet Lillebælt	0,58 %	1,15 %	1,17 %
Københavns Erhvervsakademi	0,72 %	1,39 %	1,21 %
Professionshøjskoler	10,68 %	18,43 %	14,67 %
Danmarks Medie- og Journalisthøjskole	0,14 %	0,27 %	0,33 %

Institution	Populations- fordeling	Svarfordeling	Vægtet svarfordeling
Københavns Professionshøjskole	2,77 %	4,62 %	3,63 %
Professionshøjskolen Absalon	1,35 %	2,17 %	1,70 %
Professionshøjskolen Lillebælt University College	1,14 %	1,63 %	1,27 %
Professionshøjskolen UC Syddanmark	0,90 %	1,73%	1,35 %
Professionshøjskolen University College Nordjylland	1,37 %	2,82 %	2,19 %
VIA University College	3,00 %	5,19 %	4,20 %
Universiteter	47,81 %	44,16 %	48,66 %
Copenhagen Business School	4,79 %	3,09 %	3,12 %
Danmarks Tekniske Universitet	2,28 %	3,02 %	3,00 %
IT-Universitetet	0,54 %	0,44 %	0,71 %
Københavns Universitet	10,38 %	11,74 %	13,37 %
Roskilde Universitet	2,14 %	2,27 %	2,53 %
Syddansk Universitet	4,60 %	5,06 %	5,62 %
Aalborg Universitet	6,74 %	7,37 %	7,99 %
Aarhus Universitet	16,34 %	11,17 %	12,33 %
Kunstneriske uddannelsesinstitutioner	0,65 %	1,02 %	0,65 %
Arkitektskolen Aarhus	0,17 %	0,44 %	0,28 %
Designskolen Kolding	0,04 %	0,14 %	0,09 %
KADK	0,44 %	0,44 %	0,28 %
Maritime uddannelsesinstitutioner	0,63 %	1,05 %	0,63 %
Fredericia Maskinmesterskole	0,17 %	0,20 %	0,12 %
MARTEC	0,08 %	0,24 %	0,14 %
Maskinmesterskolen København	0,12 %	0,17 %	0,10 %
SIMAC	0,07 %	0,10 %	0,06 %
Aarhus Maskinmesterskole	0,19 %	0,34 %	0,20 %
Tværgående uddannelsesinstitutioner	36,19 %	28,14 %	29,36 %
Total	100,00 %	100,00 %	100,00 %

Fokusgruppeinterviews med bedømmere

Fokusgrupper egner sig til at opfange forskellige perspektiver, modsætninger, uenigheder og enighed blandt deltagere omkring et bestemt tema. Formålet med fokusgrupperne er dermed at indhente information om mønstre og variationer i bedømmernes karaktergivningspraksis, og fokus var på følgende temaer:

1. Bedømmernes forståelse og anvendelse af skalaen
2. Bedømmernes anvendelse af bedømmelsesværktøjer til prøver
3. Hvordan bedømmerne klædes på til karaktergivning
4. Kvalitetssikring af karaktergivningen.

Der er gennemført i alt:

- Tre fokusgruppeinterviews med samlet set 18 bedømmere fra grundskoleområdet
- Fire fokusgrupper med samlet set 24 bedømmere på de gymnasiale uddannelser
- Ti fokusgrupper med samlet set 66 bedømmere på de videregående uddannelser.

Rekruttering

På grundskoleområdet blev udvalgt tre skoler ud fra følgende kriterier:

- Geografisk spredning.
- At der på skolen var lærere i fagene dansk, engelsk eller matematik, der var beskikkede censorer.

På hver skole afholdtes en fokusgruppe med 6 bedømmere, som blev udvalgt ud fra følgende kriterier:

- At de var bedømmere i dansk, engelsk eller matematik.
- At så mange af de deltagende lærere som muligt var beskikkede censorer.
- At lærere med forskellige anciennitet og erfaring var repræsenteret i fokusgruppen.

På det gymnasiale området blev udvalgt fire skoler ud fra følgende kriterier:

- Repræsentation af stx, htx, hhx og hf.
- Geografisk spredning.
- At der på skolen var lærere i fagene dansk, engelsk eller matematik, der var skriftlige censorer.

På hver skole afholdtes en fokusgruppe med 6 bedømmere, som blev udvalgt ud fra følgende kriterier:

- At de var bedømmere i dansk, engelsk eller matematik.
- At så mange af de deltagende lærere som muligt var skriftlige censorer.

- At lærere med forskellige anciennitet og erfaring var repræsenteret i fokusgruppen.

På de videregående uddannelser er rekrutteringen varetaget af Uddannelses- og Forskningsministeriet, og fokusgrupperne var sammensat af bedømmere fra forskellige institutioner på udvalgte uddannelser. I tabellen neden for ses en oversigt over fokusgrupperne.

TABEL A.14

Oversigt over fokusgrupper, de videregående uddannelser

Uddannelse	Institutioner
Arkitekt og designer	Designskolen Kolding Arkitektskolen Århus DSKD KADK
Medicin	AAU KU SDU AU
Biologi & Kemi	KU SDU AU
Maskinmester & diplomingeniør	MSK SIMAC AAMS VIA DTU SDU AU
Lærer & Pædagog	KPH VIA Absalon UCL
Sygeplejeske & Ernæring og sundhed	KPH UC Syd UCL VIA Absalon
Finansøkonom & Multimediedesigner	EA Midtvest EA Lillebælt EA Aarhus KEA CBA
Socialrådgiver & Offentlig administration	KPH UC SYD UCL VIA
Dansk & Jura	KU SDU AU
Erhvervsøkonomi & Civilingeniør	AAU DTU AU CBS

Enkelt- og dobbeltinterview med bedømmere

For at opnå mere dybdegående viden om karaktergivningspraksissen, og ud fra antagelsen om at der i karaktergivningen træffes en række hurtige beslutninger, gennemføres en række interviews med bedømmerne i umiddelbar forlængelse af eksamensafholdelse. Interviewene er gennemført som telefoninterviews eller tilstedeværelsesinterview.

Der er gennemført i alt:

- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på grundskoleområdet.
- Fem enkeltinterviews med bedømmere inden for dansk, matematik og engelsk på stx.
- Fire dobbeltinterviews med bedømmere på de videregående uddannelser.
- Fem enkeltinterviews med bedømmere på de videregående uddannelser.

På grundskoleområdet og de gymnasiale uddannelser foregik rekruttering via spørgeskemaundersøgelsen, hvor deltagerne blev spurgt, om de ønskede at deltage i et opfølgende interview. Blandt de bedømmere, der tilkendegav, at de ville deltage i et opfølgende interview, blev de interviewede lærere udvalgt ud fra følgende kriterier:

- At både lærere, der var beskikkede/skriftlige censorer og lærere der ikke var, deltog i interviews.
- At lærere med forskellige anciennitet og erfaring var repræsenteret i interviewene.

På de videregående uddannelser er rekrutteringen varetaget af Uddannelses- og Forskningsministeriet. I tabellen neden for ses en oversigt over enkelt- og dobbeltinterview på de videregående uddannelser.

TABEL A.15

Oversigt over enkelt- og dobbeltinterview på det videregående område

Område	Institutioner
Videregående uddannelser	DTU, bedømmelse af speciale efter mundtligt forsvar, dobbelt - tilstedevær
	KU, mundtlig prøve på humanistisk uddannelse, dobbelt - tilstedevær
	CBA, mundtlig prøve, dobbelt - tilstedevær KPH, prøve med mundtligt element på sygeplejerske, dobbelt - tilstedevær
	AAU, større skriftlig prøve på HA almen, telefon UCN, skriftlig prøve på pædagog, telefon EA Sydvest, afsluttende opgave på designteknolog, telefon
	RUC, ikke-afsluttende prøve med aflevering af projekt-opgave, telefon
	EA Midtvest, interview med bedømmer på multimedie-designer, telefon

Bedømmernes perspektiver på 7-trins-skalaen

© 2018 Danmarks Evalueringsinstitut

Citat med kildeangivelse er tilladt

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

ISBN (www) 978-87-7182-233-5

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