



Notat 15: Notat om simulering af implementering af modellen 0-12 og 0-6

Karakterdata bruges i forhold til adgang til ungdomsuddannelser og videregående uddannelser. Der er derfor brug for at tage stilling til, hvordan elever/studerende med eksamensbeviser fra hhv. den gamle karakterskala og den nye karakterskala rangordnes i forhold til hinanden. Dette vil ofte give anledning til etablering af omregningstabeller, der oversætter enkeltkarakterer og gennemsnit fra den gamle karakterskala til den nye karakterskala. En proces, som nødvendiggør en antagelse om, hvordan karaktererne og gennemsnittene i den nye skala forventes at fordele sig – i hvert fald i forhold til perioden frem mod, at faktisk karakterstatistik forefindes.

Endelig er der brug for at tage stilling til, hvordan elever/studerende ”undervejs” i en uddannelse skal behandles ved overgangen fra en skala til en anden – skal de gøre deres eksamen færdig efter den gamle skala, eller skal de allerede givne karakterer oversættes til karakterer efter den nye skala? At gøre eleverne/de studerende færdige efter den gamle skala, mens en ny er introduceret, etablerer den vanskelighed, at den nye og den gamle karakterskala skal sameksistere i en periode hos både elever/studerende og bedømmere.

Dette notat undersøger en alternativ implementering med udgangspunkt i ekspertgruppens videreudviklede 7-trins-skala (hvor -3 afskaffes, og 7 erstattes med et 6 tal og et 8 tal).

Spørgsmålet der stilles er, **om der findes en fordeling af karaktererne efter den videreudviklede 7-trins-skala, der resulterer i en fordeling af gennemsnit, der ligner den fordeling af gennemsnit, vi har efter den aktuelle 7-trins-skala så meget, at etablering af omregningsskalaer mellem gennemsnit efter den gamle og den nye skala kan undgås?** Og videre, om problematikken med igangværende elever/studerende kan elimineres ved, at karakterer efter den aktuelle 7-trins-skala og den nye (den videreudviklede 7-trins-skala) kan sameksistere på samme eksamensbevis, uden at elevens/den studerendes gennemsnit påvirkes i ugunstig retning?

Konklusionen på notatet er, at en sådan fordeling findes. Dog skal det påpeges, at der vil være mindre forskelle mellem gennemsnit efter 7-trins-skalaen og den videreudviklede 7-trins-skala, men disse forskelle vil i langt overvejende grad være mindre end 0,1 karakterpoint (altså forsvinde i afrundingen). Endeligt vises i notatet, at resultaterne er relativt stabile, også hvis det ikke lykkedes at ramme den forventede fordeling perfekt, og at forskellene mellem gennemsnit de to skalaer imellem er betydeligt mindre, end de forskelle i gennemsnit der historisk har været imellem forskellige eksamensår.

Notatet tager sit udgangspunkt i scenariet for implementering af den videreudviklede 7-trins-skala, hvor omregning og løbende indfasning af skalaen kan undgås, men simuleringerne kan også danne baggrund for diskussioner af implementering af 0-6-skalaen. Med udgangspunkt i samme ovennævnte fordeling for 0-6 karaktertrinnene i 0-6-skalaen, kan implementeringen af 0-6-skalaen også foretages med en ”simpel omregning”, hvor karakterer og gennemsnit fra den nuværende 7-trins-

skala i en overgangsperiode divideres med to. Overvejelser omkring implementering af 0-6-skalaen uddybes nærmere som afslutning på nærværende notat.

Simuleringerne nedenfor er gennemført med udgangspunkt i etablering af studentereksamensgennemsnit, men forventes at kunne genfindes i forhold til gennemsnit i øvrige dele af uddannelsessystemet.

Hvis den i notatet anvendte forventede fordeling ikke ønskes tilstræbt betyder det selvsagt ikke, at den videreudviklede 7-trins-skala eller 0-6-skalaen ikke kan implementeres, blot at der i så fald vil skulle etableres omregningstabeller.

En ny 7-trins-skala (den videreudviklede 7-trins-skala) med samme fordelingsmæssige egenskaber som 7-trins-skalaen

For at to skalaer er direkte sammenlignelige, skal deres fordeling af gennemsnit, dvs. sandsynligheden for at få et givet karaktergennemsnit, være den samme. Ved at indføre den videreudviklede 7-trins-skala er det ikke muligt at opnå den nøjagtigt samme fordeling af gennemsnit som ved 7-trins-skalaen i dag. Dog er det muligt at komme tæt på, og på den måde sikre at fluktuationer ved et skalaskift minimeres.

Udover at fordelingen af karaktergennemsnit skal være sammenlignelig, så må den enkelte ikke være stillet anderledes i fordelingen af karakterer efter den nuværende 7-trins-skala og karakterer efter den videreudviklede 7-trins-skala. Det betyder, at studenternes indbyrdes rangering af eksamenskvoteinter i gennemsnit gerne skulle være den samme, svarende til at to studenter med den samme kvotient efter hver deres skala ideelt set skal have præcis den samme rangering i fordelingen af karaktergennemsnit ved begge skalaer.

For at sikre sammenlignelighed i karaktergivning over årene og i forhold til den internationale ECTS-skala, etablerede Karakterkommissionen en forventet fordeling for de beståede karakterer i 7-trins-skalaen. Der er tale om en klokkeformet fordeling med 7-tallet som toppunkt. Det er muligt at etablere en forventet fordeling for den videreudviklede 7-trins-skala, hvor 7-tallet er erstattet med karaktererne 6 og 8, som har samme fordelingsmæssige karakteristika for de beståede karakterer (klokkeformet og med samme gennemsnit og spredning) som den nuværende 7-trins-skala¹. For at opnå disse fordelingskarakteristika skal desuden en mindre andel af 4-taller konverteres til karaktererne 2 og 6, og en mindre andel af 10-taller konverteres ligeledes til karaktererne 8 og 12, jf. tabel 1.

¹ Den statistiske spredning måler fordelingen af observationerne omkring gennemsnittet.

Tabel 1

Fordelingskarakteristika for gældende 7-trins-skala og den videreudviklede 7-trins-skala

Forventet karakterfordeling for 7-trins-skalaen								
Karakter	2	4	7	10	12	Gen-nemsnit	Statistisk spredning	
Fre-kvens	10 %	25 %	30 %	25 %	10 %	7	3,08	
Forventet karakterfordeling for den videreudviklede 7-trins-skala								
Karakter	2	4	6	8	10	12	Gen-nemsnit	Statistisk spredning
Fre-kvens	12 %	17 %	21 %	21 %	17 %	12 %	7	3,08

Den forventede karakterfordeling vil skulle realiseres gennem systematisk kvalitetssikring af karaktergivningen (tilpasning af prøver og læringsmål). Samtidig vil det forhold, at studerende i en overgangsperiode har eksamensbeviser, hvor nogle karakterer er givet efter den ene skala, fungere som udglatning af evt. mindre forskelle i fordelingerne efter den gamle og den nye skala.

Den nye, forventede fordeling af beståede karakterer samt omgørelsen af dumpekarakteren -3 til 0 anvendes i statistiske simuleringer til at afprøve, om den videreudviklede 7-trins-skala med nogle af ovennævnte karakteristika matematisk set understøtter, at eksamenskvothenter efter de to skalaer kan sammenlignes direkte (c).

For at undersøge om eventuelle afvigelser fra ovenstående fordeling af karakterer efter den videreudviklede 7-trins-skala har stor betydning for analysens konklusioner, er der udført en følsomhedsanalyse, hvor nogle yderligere scenarier ift. karakterfordelingen tages i betragtning. Scenarierne afspejler alle, at den videreudviklede 7-trins-skala ikke er blevet brugt som ovenstående forventede fordeling tilsiger, fx hvis der har været en større tilbøjelighed end forventet til at give karakteren 8 i stedet for 6.

Simuleringsmetode

For at give et bud på, hvordan indførelsen af den videreudviklede 7-trins-skala vil påvirke eksamenskvothenterne, har vi udført simulationer baseret på den faktiske karaktergivning i hf-, hhx-, htx- og stx-studentereksamenerne bestået i 2010 til 2019.

Karaktererne efter 7-trins-skalaen er ved lodtrækning (tilfældigt) oversat til den videreudviklede 7-trins-skala med udgangspunkt i den nye, forventede karakterfordeling, så halvdelen af 7-tallerne oversættes til karakteren 6 og den andel halvdelen til karakteren 8. For at opnå den samlede, forventede fordeling konverteres også en mindre andel af 4-taller tilfældigt til karaktererne 2 og 6, og en mindre andel af 10-taller konverteres ligeledes tilfældigt til karaktererne 8 og 12. Derudover konverteres -3 til 0 jf. tabel 2.

Tabel 2

Oversigt over omregning af enkeltkarakterer i simulering af indførsel af den videreudviklede 7-trins-skala

		Den videreudviklede 7-trins-skala						
		0	2	4	6	8	10	12
7-trins-skala	-3	100 %						
	0	100 %						
	2		100 %					
	4		8 %	68 %	24 %			
	7				50 %	50 %		
	10					24 %	68 %	8 %
	12							100 %

For at sikre mod tilfældigt skæve lodtrækninger gentages simulationen 80 gange, og alle resultater medregnes i simuleringen.

Eksamenskvoteinterne er derefter genberegnet med de nye karakterer efter den nye skala. Omregningen tager både højde for enkeltkarakterernes vægte i de faktiske karaktergennemsnit og for den gældende ordning for bonus for fag på A-niveau.

Resultater af simuleringen

Der er gennemført forskellige analyser af simuleringen efter ovennævnte metode. Alle analyserne gennemgås detaljeret i notat 17 i bilag 4 i rapporten ”Modeller for en ændret karakterskala”. Overordnet set viser resultaterne af simuleringen, at:

- fordelingen af eksamenskvoteinter er tæt på uforandret af ændringen af skalaen, hvilket vil sige, at den videreudviklede 7-trins-skala ikke ændrer betydende ved fordelingen af eksamensgennemsnit.
- rækkefølgen af studenternes eksamenskvoteinter er næsten komplet uforandret, og der er en meget stærk korrelation mellem eksamenskvoteinterne efter de to skalaer. Således vil de allerfleste studenter opnå samme eksamenskvoteinter og rangering som ved 7-trins-skalaen i dag.
- 8 procent af eksamensresultaterne må forventes at flytte sig 0,2 karakterpoint, mens de resterende 92 procent må forventes at blive inden for 0,1 karakterpoint fra den eksisterende skala.

Mens forskelle i fordelingen af eksamenskvoteinter og rangering mellem de to skalaer er tilstede, så er de små i størrelsesorden. Udover at påpege de faktiske forskelle, så kan simuleringen også give et fingerpeg af, om forskellene overhovedet er væsentlige. Indtil videre har undersøgelsen taget udgangspunkt i, at fordelingen af eksamenskvoteinter efter 7-trins-skalaen er stabil over årene, men dette er en forsimpning. I virkeligheden varierer fordelingen af eksamenskvoteinter mellem de forskellige studenterårgange.

Hvis der beregnes et samlet karaktergennemsnit for alle percentiler i de forskellige studenterårgange efter både 7-trins-skalaen og den videreudviklede 7-trins-skala, så kan man beregne variationen i karaktergennemsnittene, som kan tilskrives et skalaskift (svarende til variationen mellem skalaerne),

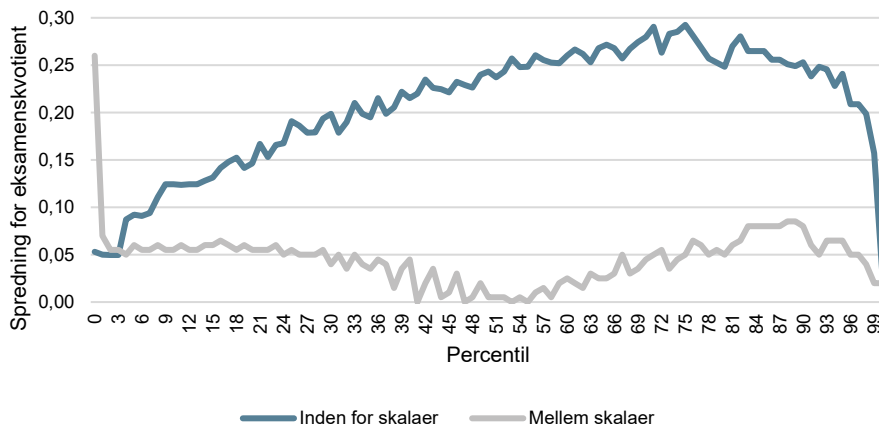
samt variationen som skyldes forskelle i karaktergivning over årene (svarende til variationen inden for de enkelte skalaer). Resultatet af analysen er vist i figur 1.

For alle percentiler gælder det, at eksamenkvotienterne svinger over tid, og at den samlede fordeling af karaktergennemsnit dermed heller ikke er stabil.² For den enkelte betyder dette, at hvilken studenterårgang man kommer fra, har betydning for ens rangering blandt de øvrige studerende.

Derudover er det værd at bemærke, at variationen for næsten alle percentiler (bortset fra de allerlaveste gennemsnit) er betydeligt mindre mellem skalaerne end mellem årene, og under afrundingsgrænsen (0,1). Det har således mindre betydning, hvilken skala man har fået karakter efter, end hvilken studenterårgang man kommer fra. På baggrund af ovenstående kan man argumentere for, at ændringerne ved ovenfor beskrevne skalaskift er uvæsentlige ved siden af de årlige udsving, som vi allerede lever med inden for skalaerne.

Figur 1

Spredningen i de årlige eksamenkvotienter for hver percentil i kvotientfordelingen fordelt på årsagen



Kilde: Uddannelses- og Forskningsministeriet på grundlag af data fra Danmarks Statistik.

1.1.1.1 Resultaternes følsomhed

For at undersøge robustheden over for en anden brug af den videreudviklede 7-trins-skala end den forventede fordeling i tabel 1 tilsiger, så er der lavet to alternativscenarier. Alternativscenarierne er symmetriske i deres skævvridning. I det ene af scenarierne er der givet højere karakterer end forventet, svarende til at flere af 4-tallerne er blevet til 6 og flere af 7-tallerne er blevet til 8 i forhold til basisscenariet. I det modsatte scenarie er flere af 10-tallerne blevet til 8 og flere af 7-tallerne er blevet til 6. De nye fordelinger er vist i tabel 3.

² Den 0. percentil er usædvanlig, og der kan ses bort fra den.

Tabel 3

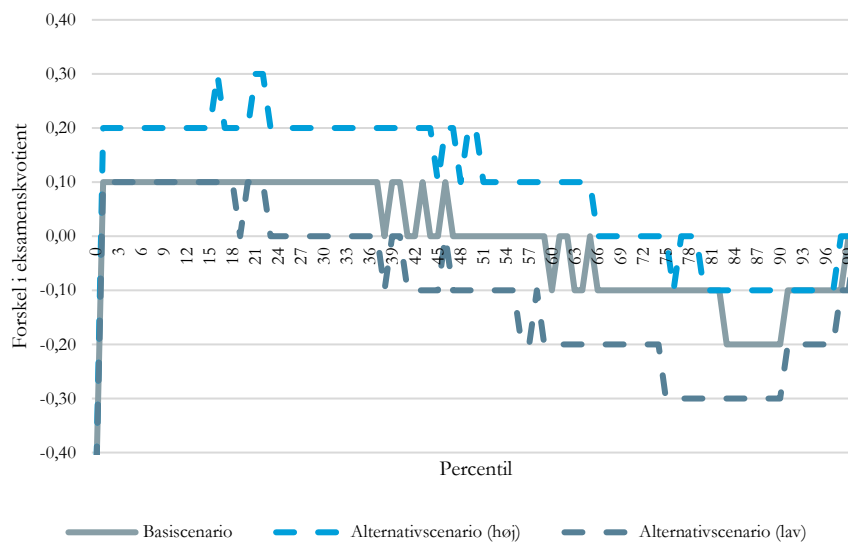
Fordelingskarakteristika for alternativscenarier

Alternativscenario, højere karaktergivning end ventet (høj)								
Karakter	2	4	6	8	10	12	Gen-nemsnit	Statistisk spredning
Fre-kvens	12 %	15 %	20 %	24 %	17 %	12 %	7,1	3,05
Alternativscenario, lavere karaktergivning end ventet (lav)								
Karakter	2	4	6	8	10	12	Gen-nemsnit	Statistisk spredning
Fre-kvens	12 %	17 %	24 %	20 %	15 %	12 %	6,9	3,05

Det er klart, at en brug af den videreudviklede 7-trins-skala som fx ikke formår at fordele 6- og 8-tallerne ligeligt, vil betyde enten højere eksamenskvote eller lavere eksamenskvote end basisscenariet. Dog gør de anderledes fordelinger en overraskende lille forskel. Dette er illustreret i figur 2.

Figur 2

Forskelle i de gennemsnitlige eksamenskvote for forskellige scenarier af den videreudviklede 7-trins-skala og 7-trins-skalaen for hver percentil i kvotientfordelingen



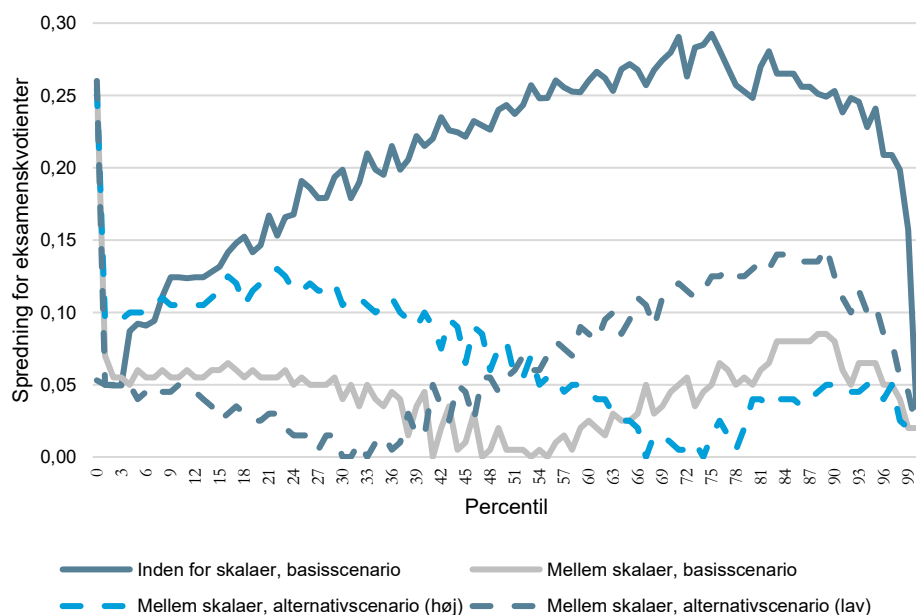
Kilde: Uddannelses- og Forskningsministeriet på grundlag af data fra Danmarks Statistik

Hvis man betragter spredningen af de årlige eksamenskvote for hver percentil, så stiger variationen, som kan tilskrives skalaskift, for alternativscenarierne relativt til basisscenariet. Dog er usikkerheden ved et skalaskift, selvom den nye skala ikke er brugt efter intentionen, for mere end 90 procent af alle studenter mindre end usikkerheden i karaktergivningen over årene.

Hvis 7-trins-skalaen skal være sammenlignelig med den videreudviklede 7-trins-skala, så er det essentielt, at den videreudviklede 7-trins-skala bruges efter intentionen. Dette resulterer i en fordeling af eksamenskvoteinter, som er tættest på den for 7-trins-skalaen. Samtidig vil studenternes rangering af eksamenskvoteinter i allerhøjest grad være bevaret ved et skalaskift. Dog er det samtidig vist, at den videreudviklede 7-trins-skala fortsat er sammenlignelig med 7-trins-skalaen, selv hvis der er en systematisk skævvridning i, hvordan karaktererne gives.

Figur 3

Spredningen i de årlige eksamenskvoteinter for hver percentil i kvotientfordelingen fordelt på årsagen inkl. alternativscenarier



Kilde: Uddannelses- og Forskningsministeriet på grundlag af data fra Danmarks Statistik

Ovenstående forholder sig alene til, hvorvidt fordelingen af gennemsnit kan ensrettes mellem 7-trins-skalaen og den videreudviklede 7-trins-skala. Der er også en problematik i forhold til anvendelsen af enkeltkarakterer de steder, hvor krav om en bestemt enkeltkarakter anvendes i uddannelsessystemet, fx i forhold til specifikke adgangskrav på de videregående uddannelser og ungdomsuddannelser – her skal etableres en omstillingstabel mellem den eksisterende 7-trins-skala og den videreudviklede 7-trins-skala.

Overvejelser om implementering af 0-6-skalaen

Ovenstående overvejelser behandler scenarier, hvor omregning og løbende indfasning kan undgås i forbindelse med implementeringen af den videreudviklede 7-trins-skala. Dette er som beskrevet ovenfor muligt ved den videreudviklede 7-trins-skala med den forventede fordeling, jf. tabel 1. Det helt samme scenarie er ikke muligt med 0-6-skalaen. Dette skyldes først og fremmest, at skalaen fra 0-6 er ”kortere” end den videreudviklede 7-trins-skala skala fra 0-12, og karakterer og gennemsnit kan derfor ikke sammenlignes direkte, hvorfor omregningen i en eller anden form er uundgåelig.

Sammenligneligheden mellem den nuværende 7-trins-skala og 0-6-skalaen kan dog styrkes ved at tage udgangspunkt i samme forventning til fordeling af karaktererne efter 0-6-skalaen som fordelingen for karaktererne i den videreudviklede 7-trins-skala. Den forventede fordeling til karaktererne efter 0-6-skalaen ses af tabel 4.

Tabel 4

Forventning til fordeling af karakterer efter 0-6-skalaen

Forventet karakterfordeling for 0-6-skalaen

Karakter	1	2	3	4	5	6
Frekvens	12 %	17 %	21 %	21 %	17 %	12 %

Ud fra samme simuleringsmetode som for den videreudviklede 7-trins-skala er der foretaget simuleringer af 0-6-skalaen. Således er der på samme vis tilfældigt konverteret eksisterende karakterer fra den nuværende 7-trins-skala til karakterer efter 0-6-skalaen, jf. ovenstående fordeling. Eksamenskvoteinterne er derefter genberegnet med de nye karakterer efter den nye skala. Omregningen tager både højde for enkeltkarakterernes vægte i de faktiske karaktergennemsnit og for den gældende ordning for bonus for fag på A-niveau.

Resultaterne af simuleringerne viser, at 0-6-skalaen resulterer i en lidt svagere korrelation mellem eksamenskvoteinterne efter de to skalaer sammenlignet med den videreudviklede 7-trins-skala³. Således vil nær halvdelen af studenter blive stillet dårligere ved rangering efter 0-6-skalaen sammenlignet med rangeringen som ved 7-trins-skalaen i dag. Dette skyldes, at 0-6-skalaen er ”kortere” end 0-12-skalaen, hvormed den har færre målepunkter i forbindelse med afrunding til eksamenskvoteinter. Det resulterer i, at 0-6 skalaen slår kvotienter målt på 7-trins-skalaen sammen. For eksempel bliver både 3,9 og 4,0 konverteret til 2,0. I dette tilfælde stilles 3,9 bedre, mens 4,0 stilles tilsvarende ringere. 0-6-skalaen differentierer i denne forstand dårligere mellem de studerendes eksamenskvoteinter end 7-trins-skalaen (og den videreudviklede 7-trins-skala) gør. Dette er dog kun ud fra udgangspunktet, at eksamenskvoteinter afrundes til nærmeste første decimal som i dag. Ved at indføre eksamenskvoteinter med to decimaler i stedet for en, vil rangeringen efter 0-6-skalaen blive mere præcis.

Som nævnt kan en grad af omregning ikke undgås ved 0-6-skalaen, da karakterer og gennemsnit ikke kan sammenlignes direkte mellem denne skala og 7-trins-skalaen. Hvis ovennævnte karakterfordeling tilstræbes (og deraf sammenlignelighed med 7-trins-skalaen) kan der imidlertid nøjes med en simpel omregningsmetode. Der er to metoder for denne simple omregning.

Rene eksamensgennemsnit efter 7-trins-skalaen konverteres til gennemsnit efter 0-6-skalaen ved at dividere gennemsnittet med to. Enkeltkarakterer fra 7-trins-skalaen konverteres direkte til karakterer efter 0-6-skalaen – også ved at dividere med to. 7-tallet ved 7-trins-skalaen ville formentlig skulle oversættes til karakteren 3,5, da det ikke på rimelig vis kan afgøres, hvorvidt det er et 3-tal eller 4-tal efter 0-6-skalaen.

³ Pearson r: 0,99549, Spearman: 0,99532.

Valg af omregningsmetode afhænger af, hvorvidt det vurderes hensigtsmæssigt at lade karakterer efter 7-trins-skalaen og 0-6-skalaen sameksistere på samme eksamensbevis. Her ville omregningsmetode 2 blive relevant.

Sameksistens af karakterer efter de to skalaer på samme eksamensbevis kan være at foretrække, da både 0-6-skalaen og 7-trins-skalaen ellers ville skulle anvendes blandt bedømmerne i en overgangsperiode. Omvendt bør det overvejes, om skiftet fra 7-trins-skalaen til en 0-6-skala er så stort, at det, uanset sammenlignelige fordelingskarakteristika ved de to skalaer, er nødvendigt at fortsætte med begge skalaer i en overgangsperiode. Disse overvejelser relaterer sig til de øvrige overvejelser om implementering, herunder informationsindsats over for bedømmere, elever, studerende og omverden, så der skabes konsensus om forståelsen og brugen af den nye skala. Eksempelvis er det vigtigt, at der ikke opstår misforståelser om, hvorvidt et 4-tal på et eksamensbevis er et 4-tal efter 7-trins-skalaen eller 0-6-skalaen. Disse øvrige implementeringsovervejelser behandles ikke i nærværende notat.