



Den professionelle  
udøver skøn i  
ubestemtheds-  
situationer

- ud fra et bredt  
vidensbaseret  
kendskab til  
evalueringsfeltet

## Karaktergivning på professionelt grundlag

Jens Dolin

Institut for Naturfagenes Didaktik



## Oversigt

Karakterfeltet - hvad er karakterer, hvilken rolle spiller de, hvad er deres konsekvenser?

Kvalitetsparametre ved evalueringer og karaktergivning. Specielt validitet og reliabilitet – og hvorledes de kan optimeres.

Problemer forbundet med karaktergivning.

Forskellen på formativ og summativ brug af evalueringer  
Kan man kombinere tænkningen om den formative og den summative evaluering?

Hvordan man kan gøre karaktergivningen til et mindre privat anliggende og tænke karaktergivningen ind i professionelle læringsfællesskaber og gøre den til en del af skolens pædagogiske fokus.



## Hvad er evaluering?

Evaluering af læring handler om at indsamle og vurdere data (i bred forstand) relateret til målene for den læring, der evalueres

Evaluering indebærer *måling* af en præstation og *vurdering* i forhold til nogle kriterier

Ved al måling måler man det måleinstrumentet viser - og ikke nødvendigvis det man er interesseret i! (Spørgsmål om *gyldighed/validitet*)

Al vurdering har et element af subjektivitet i sig. (Spørgsmål om *pålidelighed/reliabilitet*)



# Forskellige formål med summative evalueringer/karakterer

## **kontrol** af den studerendes viden

Her er validitet central (måler testen det kursusformålet siger)

## **sorteringsmekanisme**

Her er reliabilitet vigtig (er testen retfærdig), men også forudsigelsesvaliditet

## **pejlemærke for undervisning og læring**

Undervisning og eksamen må udgøre et hele,  
Washback effect, teach to the test

## **motivation**

Instrumentalisme vs. intentionalitet

## **evaluering af undervisningen/kurset/skolen/udd.systemet**

Kvalitetsindikator - på individniveau, institutionsniveau og nationalt niveau

## **styringsinstrument**

Accountability, ressourceallokering (fx via løfteevnemål)

Data samlet på ét niveau til et bestemt formål  
bruges ofte på et andet niveau til andre formål



## Problemer forbundet med karaktergivning

Opfyldelse af kvalitetskrav til evalueringsprocessen (validitet og reliabilitet)

Summative evalueringer har negative tilbagevirkninger på undervisningen (indhold og form) og på eleverne (motivation, selvværd og læringsadfærd) – **over hele uddannelseskæden**

Summative evalueringer overskygger formative evalueringer – det standpunktsvurderende blokerer for det læringsfremmende

Manglende præcision og åbenhed i bedømmelseskriterier og karaktergivningsprocedurer – og høj grad af individualitet

Lærerens dobbeltrolle som underviser og bedømmer

Underviseren indgår i formative processer, hvor eleverne viser svage sider mhp at styrke dem

Bedømmeren trækker elevernes svage sider ned, hvorfor de skjuler dem

Karakterskalaen

Karaktergivning baseres på mangler og fejl i tjekliste over målformuleringen – frem for elevens personlige tilgang til det faglige

Den ikke-ækvivalente skala vanskeliggør nuancering i midterfeltet



## Validitet (gyldighed)

**Validitet** er et udtryk for, om et instrument måler det, som det er tænkt at skulle måle, om resultaterne (tolkningen af testresultaterne) er gyldige for et specifikt formål. Et udtryk for hvilke slutninger man berettiget kan drage på baggrund af testresultater

Validitet kan ikke måles, men skal vurderes.

*Begrebsvaliditet* (construct validity): Begrebsvaliditeten udtrykker hvorvidt evalueringens målemetode er udviklet på baggrund af den rigtige teori/framework for det, der skal testes. Evalueres fx ved at se hvorvidt testens varians i elevscore korrelerer med variable som man ved er relateret til det givne construct.

*Indholdsvaliditet* (content validity): Bliver alle relevante aspekter evalueret af metoden?

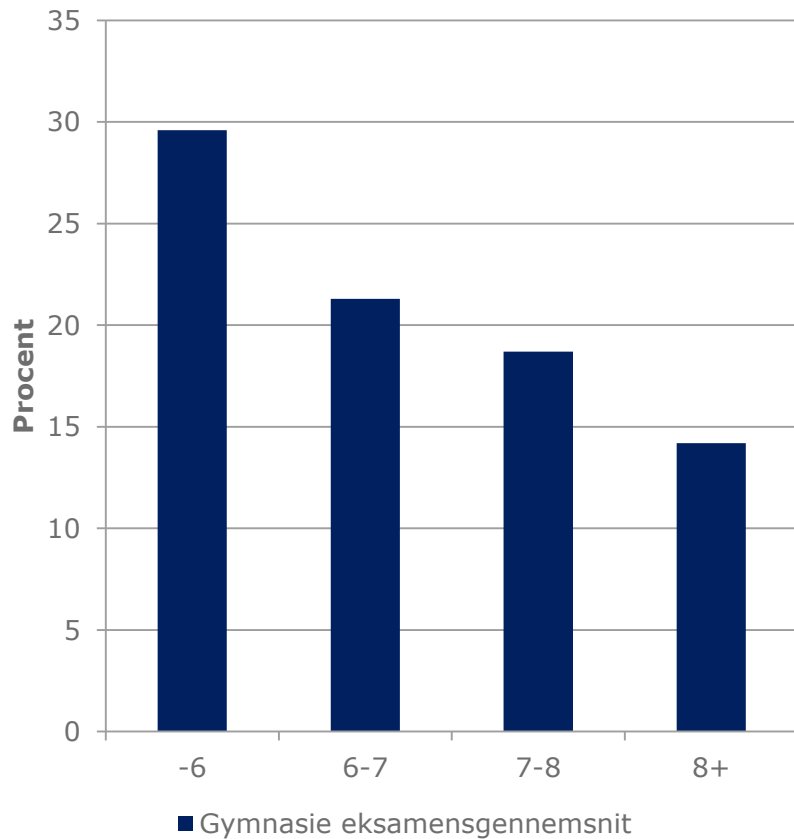
*Kriterie-relateret validitet* (criterion validity): Korrelerer evalueringens resultater med andre anerkendte metoders resultater (det gyldne standardmål)?

*Prædiktiv validitet*: Evalueringens korrelation med andre mål, som den bør kunne forudsige (fx fremtidige præstationer)

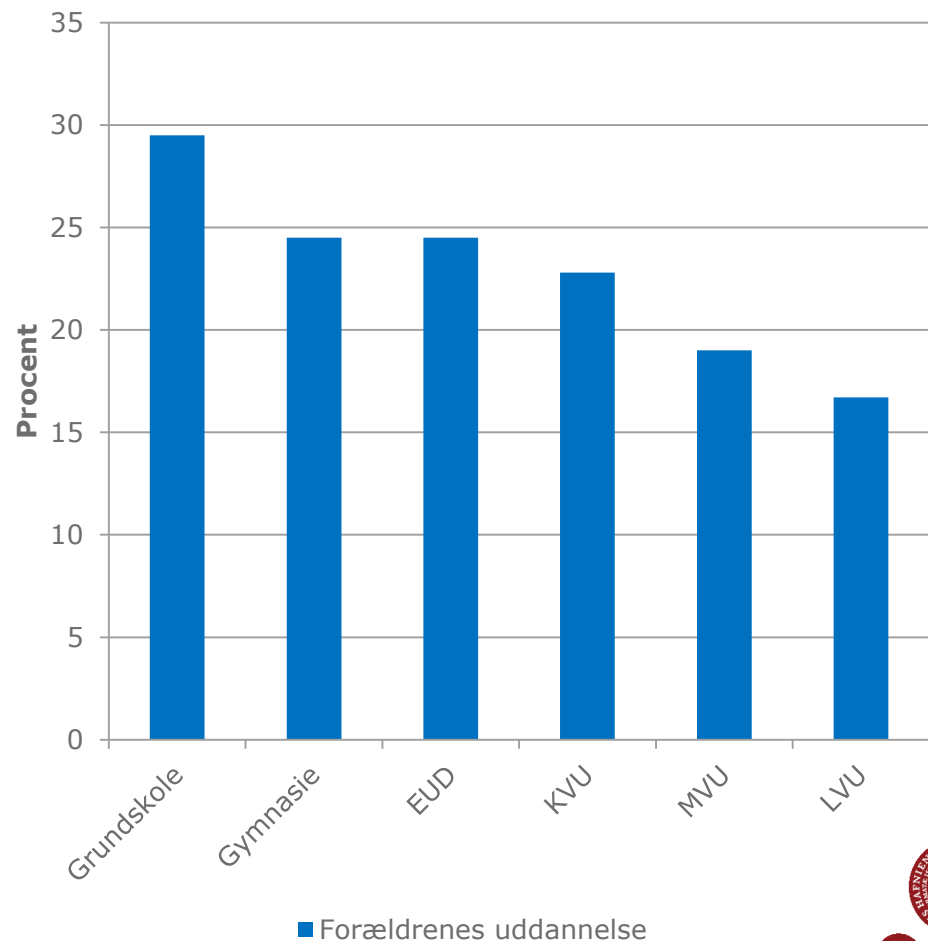


# Procentdel som afbryder inden for de første tre år på en teknat-uddannelse, studenterår 2000-2004

## Gymnasie eksamensgennemsnit



## Forældrenes uddannelse



Kilde: Lars Ulriksen, IND



## Reliabilitet (pålidelighed)

**Reliabilitet** referer til *stabiliteten/reproduktionsegenskaben*, dvs. evalueringsmetodens evne til at nå det samme resultat ved gentagne målinger.

Reliabiliteten kan udtrykkes som forskellen mellem den 'sande' værdi – som jo er ukendt - og en tilfældig målefejl (systematiske målefejl hører under validitet).

Reliabilitet kan også udtrykke *korrelationen* mellem sæt af observationsværdier – fx to eller flere individers bedømmelse af et antal objekter.

Reliabiliteten mindskes ved at:

- Forskellige bedømmere giver forskellig bedømmelse for samme præstation
- Samme elev præsterer forskelligt til forskellige tider
- En bestemt elevs præstation afhænger af hvilket spørgsmål der trækkes

The luck of the draw





## Pålidelighedsproblemer ved evaluering

”Et århundredes forskning har konsistent vist at lærere ikke er pålidelige bedømmere af elevers læring, hvis ikke de bruger strategier til at reducere målefejl” (McMillan 2013, p. 110)

Jo mere narrativt orienterede opgaver/problemer, jo lavere pålidelighed.

300 essays rettet af 53 forskellige censorer. 94 % af opgaverne fik 7 forskellige karakterer på en 7-trin skala. (ibid)

Strategier til at reducere vurderingsupålidelighed:

1. Reducere eller eliminere behovet for menneskelig bedømmelse
2. Etablere 'best practice' ved udarbejdelse af opgaverne
3. Opstille retningslinjer for bedømmelse eller skemaer (rubrics)
4. Fremme læreres forståelse af elevers læring (mhp bedre tolkning af elevsvar)
5. Danne praksisfællesskaber til at opbygge en fælles forståelse af hvad man kan forvente af elever.



## Årsager til bedømmelsesinkonsistens

Bedømmere har (vidt) forskellige forståelser af fælles kriterier og standarder, p.gr.a.

- Forskellig viden hos lærerne, forskellig erfaring og forskellige værdier
- Forskellige forventninger til elevers muligheder for at opfylde standarden
- Ignorering af eller manglende brug af kriterier
- Forskellig fortolkning af kriterier og standarder
- Forskellig vægtning af hvilke dele af standarden der er vigtigt
- Focus på forskellige dele af elevarbejder

Rapporten konkluderer (bl.a.) at mere detaljerede kriterier og stræben efter 'den rette karakter' er ikke svaret. Men sociale processer blandt bedømmerne er nødvendig. Og deling af bedømmelsesprocesserne med eleverne.

(Bloxham, 2015)



## Censorpålidelighed ved Folkeskolens afgangsprøver

Pålideligheden er beregnet på baggrund af Folkeskolens prøver i dansk og matematik ved sommerprøven 2016. Besvarelsene for 150 elever er rettet på tre forskellige måder, nemlig af en censor under almindelige rettevilkår, af en kontrolretter og af en ekspertretter.

### *Dansk:*

For to karakterer givet af to forskellige censorer for den samme besvarelse vil der være 40% sandsynlighed for at de to karakterer er ens, og 87% sandsynlighed for at de højst afviger med en karakter.

### *Matematik:*

For to karakterer givet af to forskellige censorer vil der være 72% sandsynlighed for at de to karakterer er ens, og 99% sandsynlighed for at de højst afviger med en karakter.

"... tallene (viser) med stor tydelighed hvor forsigtig man skal være med at lægge prøvekarakterer til grund for elevseleksion. Det er simpelthen ikke muligt at give fuldt pålidelige bedømmelser af elevers prøveresultater." (s. 17)

(Dolin et al, 2018)



## Hvor pålideligt skal det blive?

Rapporten om Folkeskolens afgangsprøver anbefaler (bl.a.):

- Censorrettepålideligheden øges ved at censorerne i højere grad end nu opbygger en fælles standard. Dette kan ske ved at etablere og facilitere strukturerede sociale processer blandt censorerne og ved at der etableres mere præcise kriterier.
- Censormøderne i større udstrækning bidrager til en professionalisering af censorerne
- Der strammes op på kriterier for bedømmelser i dansk og sprogfagene (udarbejdet af fagdidaktikere)

((Dolin et al, 2018, s. 7)

Men: Skal det tilstræbes at alle vurderer den samme præstation ens?

Skal det tillades at lærere (som bedømmere) har forskellige holdninger til hvad der er vigtigt og nødvendigt i deres fag?

Skal man acceptere – og deklarerer – en vis usikkerhed?



## Hvordan virker testning på undervisningen og eleverne?

### Undervisningen

- indsnævret eller fordrejet curriculum, idet faglige tankegange forsimples, faktaviden og mekaniske færdigheder betones på bekostning af kreative og æstetiske perspektiver
- undervisningstid allokeres til det/de fag og fagområder, der skal testes i, på bekostning af de fag og fagområder, der ikke testes i, og
- undervisningen kan forfalde til træning til testen og udenadslæren

### Eleverne

- elevernes testresultater stiger ved indførelse af test, men først efter nogle år
- når en test annonceres, kan det udløse følelsesreaktioner som nervøsitet og angst
- eleven forbereder sig ved at lære udenad og memorere sætninger
- for bedre præsterende elever stiger motivation, mens svagere præsterende taber modet
- det testresultat, som eleven får ved testen, kan virke ind på fremtidig motivation og selvværd.

(Nordenbo et al, 2009)



“Skolens fokus på test og karakterer har nogle omkostninger for børn og unge, som bliver sværere og sværere at ignorere, skriver Eva Secher Mathiasen, formand for Dansk Psykolog Forening”  
(Altinget Debat 23. april 2018)

## Hovedresultater

- Hver 5. studerende ser ikke nogen grund til at lære noget ud over det, de skal kunne til eksamen, og har altså en målorienteret tilgang til deres eksamen.
- Universitetsstuderende har en mere målorienteret tilgang til eksamen end studerende på professionshøjskoler og erhvervsakademier.
- Unge studerende har en mere målorienteret tilgang til eksamen end ældre studerende.
- Studerende med en målorienteret tilgang får oftere en dårlig oplevelse til eksamen.



DANMARKS  
EVALUERINGSINSTITUT

N=7513. Tallene gælder studerende, som startede på en videregående uddannelse i år 2016

26. januar 2018



## Universitetsstuderende bliver "kalkulerende"

"Skal vi bruge det til eksamen" er et almindeligt spørgsmål på landet videregående uddannelser.

Hvis svaret er nej, vælger de studerende i stigende grad læsningen fra.

"De studerende bliver kalkulerende på grund af den incitamentsstruktur, der hersker på universiteterne. Fremdriftsreformen samt de meget skrappe læringsmål sikrer en utrolig stram styring af, hvad de studerende skal lære, og hvordan og hvornår de skal lære det."

Resultatet bliver en stor gruppe dimittender med ens profiler, der ikke har lært at arbejde selvstændigt, og som ikke har haft tid og mulighed for at fordybe sig i emner, der kunne have gjort dem mere attraktive på arbejdsmarkedet"

Ida Trøjborg, Morgenavisen Jyllands-Posten, 26. september 2018, s. 4  
(baseret på iw med ph.d.-studerende i 'studerende og studieliv', Miriam Madsen)



## Præstationsorientering vs mestringsorientering

”En **mestringsorienteret** person fokuserer på den aktuelle opgave og relaterer sig især til at udvikle kompetence og opnå forståelse og indsigt.

En **præstationsorienteret** person fokuserer på selvet og relaterer sig især til hvorledes evner bliver bedømt og hvorledes man selv præsterer, især i forhold til andre.”

( Midgley et al 2001, p.77 – egen oversættelse).

Der er stærk evidens for at præstationsorientering underminerer den *indre motivation*.

Der er indikationer på at *ydre motivation* leder til ‘overflade’-snarere end ‘dyb’ læring

(Harlen 2012, p.174, egen oversættelse)

”Karaktergivning og eksamen er omstændigheder, som tilskynder til en præstationskultur hos mange elever – simpelthen fordi de lægger op til sammenligning og konkurrence”

(Krogh&Andersen, p.174)



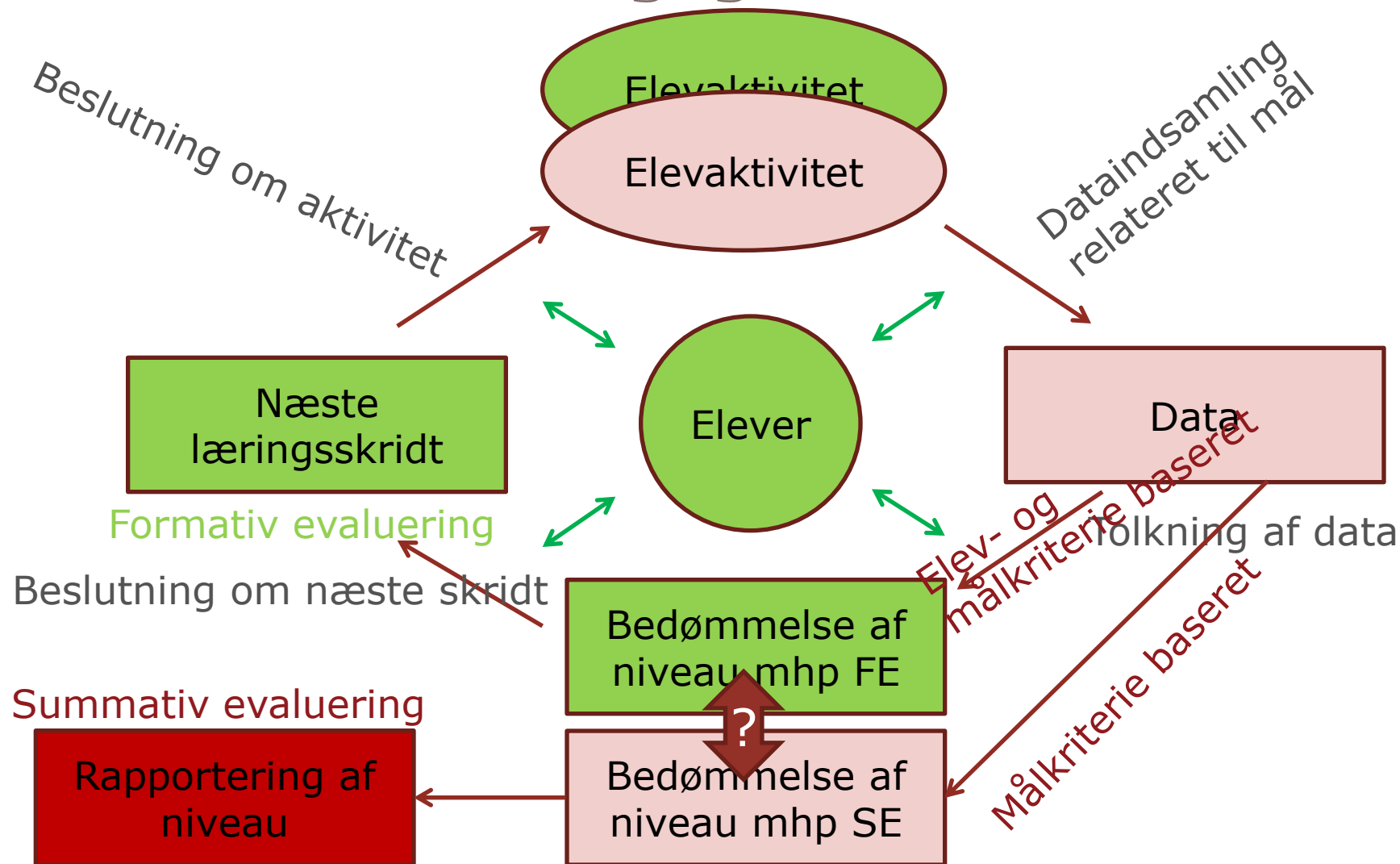


De to orienteringer kan godt være til stede samtidigt, og nogle af præstationsorienteringens positive effekter, såsom bedre eksamensresultater, kan måske kombineres med en mestringsorientering (Skaalvik & Skaalvik, 2013). Men Midgley et al (2001) understreger, at selv om 'præstationsmål kan være nyttige for nogle elever i særlige situationer så længe mestringsorienteringen også er høj' (p. 83, egen oversættelse), så er det problematisk når præstation bliver så vigtigt for elever at det uddriver mestringsmålene.

Men vi er grundlæggende ude i at finde en balance mellem formativ og summativ brug af evalueringer



## Formativ og summativ evaluering af elevlæring

**Faglige mål**

## At kombinere formativ og summativ brug af evalueringer

EVA anbefaler en klar adskillelse mellem øve- og prøverum

Det vigtige er at den didaktiske kontrakt er klar

Hvis formative evalueringer skal bruges summativt, så skal de foretages stringent og pålideligt, hvilket kræver en stor arbejdsindsats

At anvende formative evalueringer summativt vil kræve en stærk målorienteret strukturering af undervisningen (dvs. formulering af præcise krav og kriterier) – som måske vil ændre undervisningsrummet – på godt og ondt

Brug af summative evalueringer til formative formål forudsætter at

- formål og kriterier for den summative evaluering stemmer overens med formål og kriterier i den givne undervisning
- at den summative evaluering fokuserer på at teste hvorledes elever lærer
- at den summative evaluering kan virke motiverende
- at den summative evaluering tester alle dimensioner i det, der skal læres



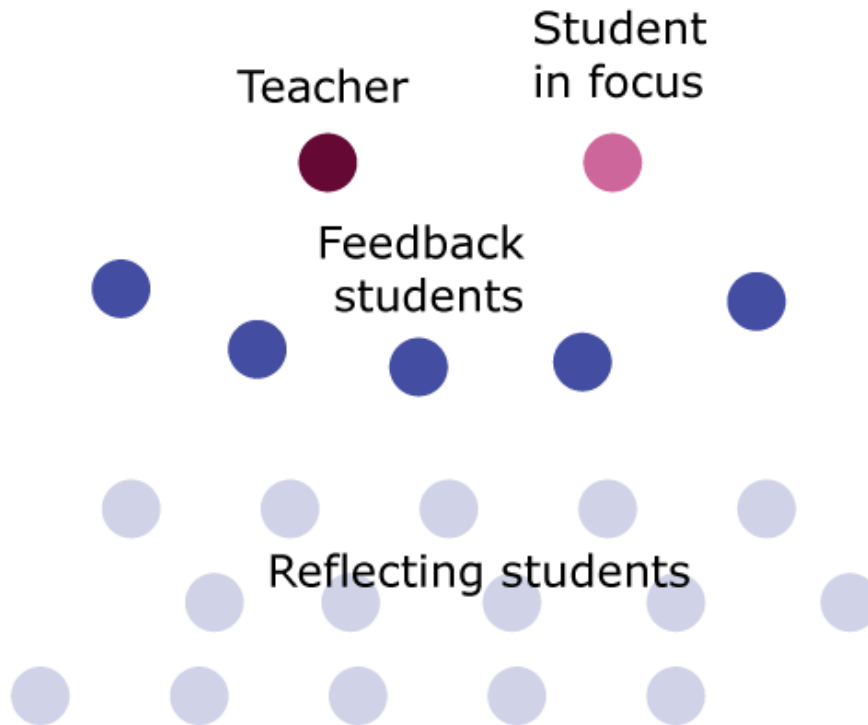
## Formativ og summativ brug af evalueringer kan ses som et kontinuum

	Formativ		<->	Summativ	
	Informal form.	Formal form.		Informal sum	Formal sum
<b>Primært fokus</b>	Hvad er næste læringstrin			Hvad er der opnået til dato	
<b>Eksempel</b>	Klassedialog	Skriftlig fb uden karakterer		Mindre tests Opgaver m karakterer	Eksamen og prøver
<b>Formål</b>	Informere næste læringstrin	Informere næste læringstrin + uvplan		Monitorere elevniveauer ift uvplan	At dokumentere individuelt elevniveau
<b>Bedømmes af</b>	Elever og lærer	Lærer og elever		Lærer	Lærer og censor

(Dolin, Black et al, 2017)



## Struktureret evalueringsdialog



1. Fem min ritualiseret samtale mellem fokuselev og lærer (skal afspejle de faglige krav)

2. Fem min feedback session mellem fokuselev og 5-7 feedback-elever (opøver feedback processer).

3. Tre min refleksion for alle elever og læreren (fremmer selvevaluering)

*Et skematisk overblik over evalueringsdialogen. En elev er i **fokus**, 5-7 andre fungerer som **feedback**-elever og resten af klassen observerer og **reflekterer** over dialogen.*

(Christensen, 2017)



## Konklusion på karakternedtoning på 14 gymnasier i 2016/17

- Arbejdet med at nedtone karakterfokus i et ellers meget karakterdrevet system anerkendes generelt som positivt af lærere, og det nedtonede karakterfokus kan være med til at understøtte en undervisning med en bedre integreret formativ evalueringspraksis.
- De elever der udsættes for færre karakterer ser ifølge lærerne ud til at blive bedre til at forholde sig til deres egen læring.
- Det er sandsynligt, at elever der udsættes for en nedtoning af karakterfokus med tiden udvikler en mere mestringsbetonet motivation for deres skolearbejde og et lavere stress niveau.
- Arbejdet med at nedtone karakterfokus i et ellers meget karakterdrevet system er forbundet med en række udfordringer, herunder at det som lærer kan være tidskrævende at fokusere på formativ evaluering, hvor man normalt blot ville give en karakter.  
(Dolin, Nielsen et al, 2017)



## Mange forsøg med karakterfrihed/nedtoning

Mange privat- og friskoler er karakterfri

### **Karakterfri skole ga samme eksamensresultat**

Etter å ha sammenlignet de tre siste årenes karaktersnitt, er imidlertid den foreløpige konklusjonen at karakterfri ungdomsskole ikke ser ut til å ha påvirket avgangskarakterene i det hele tatt. Snittet er så å si helt likt som det har vært de siste årene.

(Hentet 27. juni 2017 <https://www.nrk.no/vestfold/karakterfri-skole-ga-samme-eksamensresultat-1.13567937>)

Undervisningsministeriet har fra 1. august 2017 igangsatt et forsøg med karakterfrie klasser på 15 af landets gymnasier.

Enkelte steder har man på videregående uddannelser innsatser med karakternedtoning.



Læreres konsekvente og konsistente professionelle arbejde med evaluering og kompetenceorientering kræver

“...at man på nationalt og skoleplan prioriterer at lærere begynder at samarbejde i tæt knyttede teams om kompetencebetoning og -evaluering,”

men

“... dette samarbejde er udfordrende. For det første er det tidskrævende og svært for selv erfarne lærere at udføre dette arbejde konsistent rundt om deres undervisning.

For det andet viser erfaringerne fra ASSIST-ME at der er en hårfin balance mellem at operationalisere komplekse kompetencer og at opstille en tjekliste af målpinde som skal krydses af. Med andre ord er det vigtigt at undgå at evaluering bliver instrumentalistisk. Den primære grund hertil er at en sådan evalueringspraksis direkte understøtter en præstationsorientering “

(Nielsen&Dolin, 2016)





## Mangel på stærke professionelle fællesskaber

Forskning viser at karakterer ikke gives stabilt på et sikkert grundlag og lærerne oplever at stå alene med karaktergivningen.

Behov for at skolen etablerer stærke fælles normer om evaluering og karaktergivning.

I et samarbejde mellem ledelse og lærere skal der opbygges et læringsfællesskab som sikrer transparens, vidensopbygning, vidensdeling, eksterne input, ...

Det forudsætter tillid mellem alle ansatte og det kræver mere ledelse end styring.



## Målprioritering kan have effekt

Baseret på spørgeskema til 73 stx-rektorer og 1353 gymnasielærere fra de samme skoler, omhandlende sammenhænge mellem rektorers prioritering af et mål og gymnasielærernes engagement i dette mål, konkluderer Staniok (2016):

"... rektorer på de gymnasiale uddannelser, som i høj grad prioriterer målsætningen om en høj gennemførelsesprocent, har en positiv effekt på gymnasielærernes engagement i målet, men kun når gymnasielærerne ikke oplever, at der er en høj grad af konflikt imellem det at forfølge målsætningerne om en høj gennemførelsesprocent og et højt fagligt miljø."

"... medarbejderindflydelse ofte spiller en afgørende rolle ... medindflydelse og ikke medinddragelse ... (det betyder) at medarbejdere ikke blot høres, men at det hørte faktisk også inddrages i beslutningsprocessen"

To konklusioner:

"... ledelse i form af målprioritering nytter på de gymnasiale uddannelser ... der er en potentiel risiko for målforskydning ..."



## Opbygning af en evalueringskultur

Eksplicite  
begrundelser og  
procedurer

Er karakteriseret  
ved det implicitte

“En evaluering udføres med sigte på et formål (en eller anden form for anvendelse), hvorimod kulturer udmærker sig ved at holde noget helligt, som ikke lader sig reducere til noget instrumentelt.”

(Dahler-Larsen, 2006, s. 36)

En evalueringskultur ser evaluering som et didaktisk anliggende for hele skolen – med sigte på at øge elevernes læring og vurdere deres læring på et professionelt grundlag



# GYMNASIE PÆDAGOGIK

En grundbog

REDIGERET AF

Jens Dolin, Dion Rüsselbæk Hansen, Gitte Holten Ingerslev og  
Hanne Sparholt Jørgensen



## Teamarbejdet som drivkraft

### **Kap 7.5 Team som element i skoleudvikling** Af Anne Birgitte Klange

”Hvis visionen for skolen imidlertid er at blive en *lærende organisation*, hvor lærere og ledelse hele tiden bliver bedre til at løse opgaverne, kræver det, at team udvikler sig til at blive egentlige professionelle læringsfællesskaber.” (s. 740)

### **Kap 7.3 Profession og pædagogik** Af Jakob Ditlev Bøje

- Et professionelt læringsfællesskab er karakteriseret ved at
- medlemmerne har en fælles forståelse af værdier og visioner
  - de tager kollektivt ansvar for elevernes læring
  - de arbejder med reflekterende og undersøgende metoder
  - de samarbejder på et praktisk plan, fx om undervisning og evaluering
  - de fremmer læring og kompetenceudvikling i fællesskab



## Afsluttende

At opbygge en evalueringskultur er ikke kun et spørgsmål om at bruge de 'rigtige' evalueringsværktøjer.  
Det handler også om at forholde sig til de mange virkninger, værktøjerne vil have.

Karakterer er skolens mest indflydelsesrige værktøj, som påvirker elever, lærere og ledelse.  
De tilskynder eleverne til en præstationsorientering, lærerne til en forsimplet undervisning og ledelsen til at agere på et (måske for) enkelt grundlag.

En professionalisering af karaktergivningen vil involvere hele skolen gennem et samarbejde mellem lærerteams og ledelse – ikke kun om en diskussion af givne karakterer, men om en udvikling af læringsfremmende og konsistente måder at give karakterer på, om opbygningen af en evalueringskultur.



# Referencer

- Black and Wiliam (1998). Inside the Black Box: raising standards through classroom assessment. *Phi Delta Kappa*, 80(2), 139-144, 146-148.
- Black, P., Harrison, C. et al (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1) 8-21.
- Bloxham, S. (2015). The multiple limitations of assessment criteria. Transforming Assessment Webinar.
- Christensen, T. S. (2017). Formativ evaluering. I Dolin, J., Ingerslev, G. H. og Jørgensen, H. S. (red.). *Gymnasiepædagogik*, 3. rev. udg. København: Hans Reitzels Forlag.
- Dahler-Larsen, P. (2006). *Evalueringens kultur. Et begreb bliver til*. Odense: Syddansk Universitetsforlag.
- Dolin, J., Nielsen, K. & Rangvid, B. S. (2018). Rapport fra Følgegruppen for én bedømmer ved folkeskolens prøver. <https://uvm.dk/aktuelt/nyheder/uvm/2018/apr/180425-rapport-censorer-vurderer-folkeskolens-bedoemmerordning-i-ny-rapport>
- Dolin, J., Nielsen, J. A. & Tidemand, S. (2017). Evaluering af naturfaglige kompetencer. I: *Acta Didactica Norge - nasjonalt tidsskrift for fagdidaktisk forsknings- og utviklingsarbeid*. 11, 3, s. 1-28.
- Dolin, J., Black, P., Harlen, W. & Tiberghien, A. (2017). Exploring relations between formative and summative assessment. In Dolin, J. & Evans, R. (eds.). *Transforming Assessment. Through an Interplay Between Practice, Research and Policy*. Springer International Publishing. (s. 53-80).
- Dolin, J. (2016). Idealer og realiteter i målorienteret undervisning. I E. Krogh, & S-E. Holgersen (red.), *CURSIV: Sammenlignende fagdidaktik 4*. (Vol. 19, s. 67-87). København.
- Harlen, W. (2012). The Role of Assessment in Developing Motivation for Learning. In: Gardner, J. (ed.). *Assessment and Learning*. London: SAGE.
- Klange, A. B. (2017). Team som element i skoleudvikling. I Dolin, J., Ingerslev, G. H. og Jørgensen, H. S. (red.). *Gymnasiepædagogik*, 3. rev. udg. København: Hans Reitzels Forlag.
- Marinos, N. (2017). Summativ evaluering. I Dolin, J., Ingerslev, G. H. og Jørgensen, H. S. (red.). *Gymnasiepædagogik*, 3. rev. udg. København: Hans Reitzels Forlag.
- McMillan (2013). I: McMillan (ed.). *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE
- Midgley, C., Kaplan, A. & Middleton, M. (2001). Performance-Approach Goals: Good For What, For Whom, Under What Circumstances, and At What Cost? *Journal of Educational Psychology*, 93(1), s. 77-86.
- Nielsen, J.A. og Dolin, J. (2016). Evaluering mellem mestring og præstation. *MONA 1*, 51-62.
- Nordenbo, S. E., Allerup, P., Andersen, H.L., Dolin, J., Korp, H., Larsen, M.S., Olsen, R.V., Svendsen, M.M., Tiftikçi, N., Wendt, R.E., Østergaard, S. (2009). *Pædagogisk brug af test – Et systematisk review*. København: Danmarks Pædagogiske Universitetsforlag og Dansk Clearinghouse for Uddannelsesforskning
- Parkes, J. (2013). Reliability in Classroom Assessment. In McMillan (ed.). *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE.
- Skaalvik, E.M. and Skaalvik, S. (2013). School Goal Structure: Associations with Students' Perceptions of Their Teachers as Emotionally Supportive, Academic Self-Concept, Intrinsic Motivation, Effort, and Help Seeking Behavior. *International Journal of Educational Research*, (61), s. 5-14.
- Staniok, C. D. (2016). Forsker: Nyttet målprioritering på de gymnasiale uddannelser?. [www.altinget.dk](http://www.altinget.dk) 3.juni 2016.
- Topping, K.J. (2013). Peers as Source of Formative and Summative Assessment. In McMillan (ed.). *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE

