

**DUU TT**

Dansk Universitetspædagogisk Tidsskrift

**TEMA:**

**Fra data til beslutninger**

Årgang 14 nr. 26/2019

# Indhold

Leder <i>Tina Bering Keiding</i>	1-4
Fremmedsprogsstuderendes overbevisninger om sproglæring: Før og efter profilsemestret på kandidatstudiet <i>Susana Silvia Fernández</i>	5-22
A flowchart as a tool to support student learning in a laboratory exercise <i>Dominik K. Grosskinsky, Katrin Hammer úr Skúoy, Kirsten Jørgensen</i>	23-35
Matematiktest som prædiktor af ingeniørstuderendes studiesucces <i>Henrik Skov Midtiby, Nadia Dyrberg Egemose</i>	36-50
Screening for technical flaws in multiple-choice items. A generalizability study. <i>Lotte Dyhrberg O'Neill, Sara Mathilde Radl Mortensen, Cita Nørgård, Anne Lindebo Holm Øvrehus, Ulla Glenert Friis</i>	51-65
Intended and unintended test constructs in a Multiple-Mini admission Interview. A validity study. <i>Lotte Dyhrberg O'Neill, Eva Lykkegaard, Kulamakan Kulasageram</i>	66-81
Kan videregående uddannelser mindske frafald ved hjælp af læringsteknologi? Et systematisk review <i>Maria Hvid Stenalt, Berit Lassesen, Dorte Sidelmann Rossen, Anna Bager-Elsborg</i>	82-102
Takeaway Teaching – A design for redesign <i>Hanne Balsby Thingholm, Anders Hjortskov Larsen</i>	103-114
DUT Guide om akademisk skrivning <i>Tine Wirenfeldt Jensen</i>	115-124
DUT Guide on teaching and learning in the international classroom <i>Karen M. Lauridsen</i>	125-132
Videnskabelse på universitetet. Veje til integration af forskning og undervisning. <i>Jakob Egholm Feldt</i>	133-135
Kig op! Undervisning uden et blåt skær <i>Kim J. Herrmann</i>	136-138

1968: Studenteroprør og undervisningsrevolution. En fortælling om opgøret med traditionel universitetsundervisning  
*Birger Steen Nielsen*

139-142

# Fra data til beslutninger – hvordan får vi en bredere forståelse af universitetsuddannelsernes værdi og betydning?

Tina Bering Keiding<sup>a,1</sup>

<sup>a</sup>Center for Undervisningsudvikling og Digitale Medier, Aarhus Universitet

*Dansk Universitetspædagogisk Tidsskrift har inviteret Tina Bering Keiding til at indlede temanummeret på baggrund af hendes åbningstale ved konferencen med temaet Fra data til beslutninger, som blev afholdt af Dansk Universitetspædagogisk Netværk den 17. til 18. maj 2018.*

I foråret 2018 udkom rapporten "Universitetsuddannelser til fremtiden" (UFM & UUU, 2018). Et digert værk, som forsøger at beskrive værdien af universitetsuddannelserne samt at identificere aktuelle og fremtidige udfordringer og udviklingsmuligheder for disse. I dette indlæg vil jeg diskutere, dels hvorledes universitetsuddannelsernes værdi kommer til udtryk i rapporten, dels hvilke uddannelsespolitiske konsekvenser rapportens iagttagelsesmåde kan have. Afslutningsvist vil jeg rejse spørgsmålet, om de universitetsdidaktiske forskningsmiljøer kan – og skal – gå sammen om at forsøge at udvide den måde, hvorpå vi empirisk beskriver universitetsuddannelsernes værdi og betydning for samfundet og for den enkelte.

## Værdien, der blev væk

Som læser af rapporten og som deltager i Udvalget om bedre universitetsuddannelser (UUU) er jeg ikke i tvivl om, at universitetsuddannelser generelt ses som værdifulde for både samfundet og de studerende, som tager en uddannelse. Et par eksempler fra rapporten:

*Gode uddannelser udvider vores horisont, giver valgmuligheder i livet og engagement til at leve som aktive borgere på arbejdsmarkedet og i vores demokrati. [...] Det er en investering, der kan betale sig, for universitetsuddannelserne giver stor værdi både for den enkelte og for samfundet. Kandidater fra universiteterne løser opgaver på højt niveau, de bringer den nyeste forskningsbaserede viden ud i vores samfund, og de bidrager med kritisk stillingtagen og engagement. (UFM & UUU, 2018, s. 8).*

*Et højt uddannelsesniveau har således betydning for den enkeltes livskvalitet og velfærd. Eksempelvis er der en veldokumenteret sammenhæng mellem uddannelseslængde og sundhed og deltagelse i samfundet, som begge øges i takt med et stigende uddannelsesniveau. [...] Ud fra et demokratisk ideal er det værdifuldt at opnå forudsætningerne for at deltage i samfundet gennem uddannelse, da evnerne til at deltage giver mulighed for at omsætte sine præferencer til indflydelse. Empiriske studier peger på, at deltagelse i såvel politiske, lokale og samfundsmæssige foretagender generelt er højere blandt universitetsuddannede (UFM & UUU, 2018, s. 83-84).*

Efter disse indledende rammesætninger præget af et holistisk blik på værdien af universitetsuddannelser reduceres beskrivelserne af universitetsuddannelsers værdi dog i rapporten

---

<sup>1</sup> Kontakt: keiding@tdm.au.dk

til samfundsøkonomisk værdi (UFM & UUU, 2018, s. 78ff). Der tales om universitetsuddannelsernes bidrag til samfundets velstand, om samfundsøkonomisk værdi af et højere uddannelsesniveau, om samfundsøkonomisk afkast af uddannelse og uddannelsernes økonomiske værdiskabelse, herunder beskæftigelse og livsindkomst for den enkelte. Den økonomiske værdiskabelse opgøres konsekvent i forhold til dimittendernes deltagelse på arbejdsmarkedet, selv om det et enkelt sted anføres, at sundhed og deltagelse i samfundet øges i takt med uddannelsesniveaet (s. 84).

Dermed iagttages værdien af universitetsuddannelser gennem det man, med henvisning til den tyske sociolog Niklas Luhmann, kan betegne som det økonomiske systems kode: overskud eller ikke-overskud (f.eks. Luhmann, 2000, s. 526 for en kort introduktion). Eller sagt på en anden måde: Rapporten afgør værdien af universitetsuddannelser ud fra spørgsmålet om, til hvad og hvor meget det rent økonomisk kan betale sig for samfundet at uddanne dets medborgere og for den enkelte at tage en uddannelse. Og selv et meget hurtigt blik gennem rapporten efterlader ingen tvivl om, at der er umådelige mængder af data til rådighed for sådanne beskrivelser.

I modsætning hertil forbliver beskrivelsen af andre former for værdi og betydning af universitetsuddannelserne på et meget abstrakt niveau. Denne ensidighed i blikket forklares således:

*Mange af de værdifulde effekter, herunder ikke mindst de ikke-økonomiske gevinster, er imidlertid meget svære at måle. Fokus er derfor rettet mod universitetsuddannelsernes samfundsøkonomiske værdi. (UFM & UUU, 2018, s. 20).*

### **Kan det ikke tælles, tæller det ikke**

Det er ikke mit ærinde at anfægte, at spørgsmålet, om hvorledes universitetsuddannelser bidrager positivt til samfundets velstand, er en relevant optik i en uddannelsespolitisk diskussion. Men jeg vil gerne pege på nogle mulige konsekvenser, hvis – eller når – den økonomiske optik og snæver nyttetænkning i forhold til arbejdsmarkedet bliver det eneste empiriske grundlag for de iagttagelser af uddannelsessystemet, der informerer politiske beslutninger.

For det første, og grænsende til det banale, betyder det, at alt, hvad der ikke kan beskrives/opgøres økonomisk i forhold til arbejdsmarkedet, falder uden for værdibegrebet og dermed bliver ikke-eksisterende i forhold til at informere det politiske system. Selvom det nævnes, at uddannelse bidrager med andet og mere end samfundsøkonomisk værdi, forbliver disse beskrivelser på et så abstrakt niveau, at de som empirisk realitet og afsæt for politiske beslutninger mest af alt får karakter af en forlegenhedskategori: Vi ved, at der er noget væsentligt derude, men vi kan ikke beskrive det, og dermed heller ikke inddrage det som beslutningsgrundlag.

For det andet er jeg skeptisk over for det ensidige fokus på værdiskabelse, der lægges hen over universitetsuddannelserne. Selv når begrebet udvides til "bløde former" for værdiskabelse, der ikke opgøres økonomisk men i mere kvalitative kategorier, sådan som vi ser i dele af den nyere entreprenørielle didaktik, er der tale om en optik, der lægger sig tæt op af den økonomiske kode. Vi taler stadig om +/- værdiskabelse, blot med et bredere værdibegreb, som supplerer det økonomiske systems positiv-værdi (overskud/gevinst), såsom: at agere i samfundet, at tage vare på eget liv og karriere og at kunne iværksætte initiativer, der skaber værdi for andre, kulturelt, økonomisk og socialt (fx Rasmussen, Revsbeck, Moberg 2015)

At uddannelse bidrager til værdiskabelse betyder ikke, at *nytte* er den eneste relevante optik til at beskrive, hvorfor samfundet skal sikre mulighed for, at dets borgere kan uddanne sig. I den dannelseseoretiske didaktik formuleres almen dannelse som den dimension af undervisningen, der ikke er rettet mod bestemte erhverv (real dannelse), men mod at forme eller (ud)danne den lærende, fordi denne dermed får mulighed for at repositionere sig i verden, både erkendelsesmæssigt og handlingsmæssigt, uden at uddannelsessystemet på forhånd søger at specificere disse muligheder (Korsgaard, Kristensen, & Siggaard Jensen, 2017). Sagt med andre ord skal uddannelsessystemet i denne optik tilbyde et indhold (matter), men selve betydningen af indholdet, og hvordan indholdet helt konkret åbner verden for den enkelte (meaning), anses som et anliggende mellem den lærende og indholdet (Hopmann, 2007).

For at skabe en både mere nuanceret og mere konkret forståelse af, hvorfor universitetsuddannelser er væsentlige for både samfundet og den enkelte, vil det derfor være relevant *også* at kunne tilbyde robuste beskrivelser af, hvorledes universitetsuddannelser former de studerende som personer og som medborgere, og hvad dette *betyder* på både kort og lang sigt for de uddannedes livsløb, både på arbejdsmarkedet, i civilsamfundet men ikke mindst for deres oplevelse af deres eget liv.

Jeg tror bestemt, at UUU har ret, når de i rapporten anfører, at det er svært at måle værdier og betydninger, som ligger ud over hvad der kan trækkes ud af diverse registre. Men at noget er svært at måle betyder jo ikke nødvendigvis, at man skal undlade at gøre forsøget.

Jeg vil ikke her forsøge at komme med konkrete bud på, hvilke teorier og metoder, der vil kunne løfte et sådant forehavende. Men jeg tror for det første, at vi skal orientere os bredt mod nogle af den pædagogiske filosofis og dannelseseoretiske didaktiks svar på det grundlæggende spørgsmål: Hvorfor uddanne og til hvad? For det andet tror jeg, at analysestrategien (også) skal give mulighed for generaliseringer og sammenligninger. Det vil sige, at den skal vove det ene øje og forsøge at operationalisere de kvalitativt orienterede beskrivelser af uddannelsens betydning for samfundet og den enkelte på en måde, så de kan danne grundlag for kvantitativt orienteret forskning, statistiske analyser og generaliserende udsagn. Ellers forbliver denne nye stemme fra den universitetspædagogiske forskning – alle øvrige kvaliteter ufortalte – for svag og kontekstbundet til for alvor at kunne gå i dialog med og med vægt udfordre og supplere de nuværende analyser.

Jeg vil derfor slutte af med at invitere interesserede til et møde på Aarhus Universitet i forsommeren 2019 med henblik på at drøfte, dels hvordan vi kan *udvide værdibegrebet* i de nuværende beskrivelser af værdien af universitetsuddannelserne, så de inkluderer mere end blot økonomisk gevinst gennem deltagelse på arbejdsmarkedet, dels hvordan vi empirisk kan *beskrive betydningen* af universitets(ud)dannelse for den enkelte i forhold til hans/hendes livsløb. Og gerne med afsæt i sådanne drøftelser udvikle et fælles forskningsprojekt, som forsøger at måle det som nok er svært at måle, men som på den anden side er for betydningsfuldt til ikke at indgå i det uddannelsespolitiske beslutningsgrundlag.

## Referencer

- Hopmann, S. (2007). Restrained Teaching: the common core of Didaktik. *European Educational Research Journal*, 6(2), 109-124.
- Korsgaard, O. F., Kristensen, J. E. F., & Siggaard Jensen, H. (2017). *Pædagogikkens idehistorie*. Aarhus: Aarhus Universitetsforlag.
- Luhmann, N. (2000). *Sociale systemer*. København: Hans Reitzels.
- Rasmussen, Revsbeck, Moberg (2015) Taksonomi for entreprenørskabsuddannelse, Fonden for Entreprenørskab. Tilgængelig på <http://www.ffe-ye.dk/media/538364/taksonomi-i-entreprenoerskabsuddannelse-til-net.pdf>
- UFM & UUU. (2018). *Universitetsuddannelser til fremtiden*. Hentet fra København: <https://ufm.dk/publikationer/2018/filer/rapport-universitetsuddannelser-til-fremtiden.pdf>

# Fremmedsprogsstuderendes overbevisninger om sprog-læring: Før og efter profilsemestret på kandidatstudiet

Susana Silvia Fernández<sup>a,1</sup>

<sup>a</sup>Institut for Kommunikation og Kultur, Aarhus Universitet

Faglig artikel, fagfællebedømt

Formålet med artiklen er at analysere, hvordan sprogstuderendes overbevisninger om sprogsyn og sproglæring bliver påvirket, når de vælger at følge en "underviserprofil" på deres kandidatstudie. Ved at vælge underviserprofilen på det næstsidste semester i kandidatuddannelsen kan den universitetsstuderende tone sit studium mod underviser-/lærerprofessionen med særligt fokus på undervisning i det danske gymnasium. Via faget "Undervisning i teori og praksis", som inkluderer gymnasiepraktik, får fremmedsprogsstuderende mulighed for at beskæftige sig med fremmedsprogsundervisning og didaktik fra et teoretisk og praktisk perspektiv. Gennem et casestudie af otte studerende fra et dansk universitet og på basis af metoder udviklet i forskningsfeltet om lærerkognition viser jeg, hvordan de studerendes viden og tanker, om hvordan man bedst underviser i og lærer fremmedsprog, udvikler sig i løbet af profilsemestret.

## Introduktion

Forskningsfeltet om lærerkognition har eksisteret i over 30 år med fokus på både nye og erfarne lærere samt også til en vis grad på lærerstuderende. I de seneste ca. 15 år har der været en voksende interesse for fremmedsproglærernes kognition, især takket være pionerarbejde af Simon Borg (fx 1998a,b; 2001; 2003), og det er lige præcis sproglærernes (rettere sagt sproglærerstuderendes) kognition, som er i fokus i nærværende artikel.

Målet med undersøgelserne inden for sproglærernes kognition er at skabe en bedre forståelse for, hvordan fremmedsproglærernes indre liv, dvs. deres tanker, overbevisninger, viden og følelser, der relaterer sig til det de underviser i, påvirker deres praksis i klasseværelset og er med til at definere dem professionelt. Samtidig sigter man mod en forståelse af, hvor lærernes tanker og overbevisninger stammer fra, og hvordan de evt. påvirkes af uddannelse, efteruddannelse og andre kontekstuelle faktorer (Borg, 2003). I denne artikel vil jeg se nærmere på en bestemt målgruppe: danske kandidatstuderende inden for fremmedsprog, som har valgt at tone deres studium mod en karrierevej inden for de gymnasiale uddannelser (dvs. kommende gymnasielærere inden for fremmedsprog). Jeg vil undersøge deres kognition før, under og efter deres "profilsemester", dvs. det semester på kandidatstudiet, som er målrettet deres fremtidige undervisningserhverv. Ved at indsamle og analysere forskellige typer data (besvarelser fra refleksionsøvelser, gruppediskussioner, produktionsøvelser, refleksionssays og interviews) vil jeg undersøge, hvordan profilsemestret påvirker disse studerendes kognition, og hvordan de selv oplever og reflekterer over semestret.

---

<sup>1</sup> Kontakt: romssf@cc.au.dk



Artiklen er struktureret således, at der efter denne introduktion følger en kort præsentation af forskningsfeltet lærerkognition med særligt fokus på fremmedsproglærere. Jeg vil beskrive de områder, der er blevet undersøgt, og de metoder, som typisk bliver anvendt. Derefter præsenteres kontekstuel information om, hvordan gymnasielærere bliver uddannet i Danmark efterfulgt af en detaljeret beskrivelse af det uddannelsesmodul, som bliver diskuteret i denne undersøgelse: profilmestret for studerende fra humanistiske kandidatuddannelser på Aarhus Universitet, som ønsker at tone deres studium i en undervisningsorienteret retning (herunder undervisning i danske gymnasier). Når både den teoretiske og den kontekstuelle ramme er på plads, bliver den empiriske analyse præsenteret efterfulgt af en diskussion og konklusion.

### Lærerkognition som forskningsfelt

Begrebet lærerkognition (*teacher cognition*) er blevet defineret af Borg (2003, s. 81) som “the unobservable cognitive dimension of teaching – what teachers know, believe and think”. I forskningsfeltet om lærerkognition prøver man både at udforske og forstå lærernes viden og overbevisninger samt at finde ud af, hvordan disse påvirker deres praksis i klasseværelset. Interessen for lærerens tanker hænger sammen med en voksende forståelse for lærerens centrale rolle i undervisningen og for elevernes læringsproces (Hattie, 2003). Dette indebærer en overbevisning om, at læreren ikke blot mekanisk følger en opskrift til undervisningen, men er en beslutningstager, som er med til at designe, udføre og evaluere sin praksis i et cyklisk forløb – med Schöns (2001) ord: en “reflekterende praktiker”.

I forskningen om lærerkognition har man fundet, at lærernes overbevisninger er hierarkisk organiseret med nogle tanker, som er stærkt rodfæstede, og andre, som er mere perifere og dermed nemmere at påvirke, fx ved uddannelse og efteruddannelse. Ofte er de mere centrale og rodfæstede tanker gamle, og de relaterer sig til almenpædagogiske forhold, såsom arbejdsformer eller elevernes motivation (Phipps & Borg, 2009; Daryai-Hansen & Henriksen, 2017). Når forskellige overbevisninger er i modsætning til hinanden, vil de centrale overbevisninger typisk tilsidesætte de mere perifere i lærerens praksis. Fx har Vijayarathan (2017) fundet, at engelsklærere på Færøerne mente, at det er vigtigt at værne om elevernes motivation og selvværd, og at det også er vigtigt for læring at rette mundtlige fejl, men at de i praksis vælger ikke at rette mundtlige fejl af hensyn til elevernes følelser.

Forskningen har også vist, at de mest hårdføre tanker og meninger stammer fra lærernes egen erfaring som elever, hvad Lortie (1975, s. 67) har kaldt *apprenticeship of observation*. Disse tidlige kontakter med klasseværelset og sprogundervisning danner en stærk basis for lærernes egen konceptualisering af, hvad det betyder at lære og undervise i fremmedsprog. Disse tidlige modeller er stærke i de begyndende stadier af lærerprofessionen, og de kan sagtens fortsætte med at påvirke lærernes praksis gennem hele deres professionelle liv (Borg, 2006, s. 88).

Lærernes kognition er kompleks, ikke alene fordi tankerne kan være i konkurrence med hinanden, som vi lige har set, men også fordi mentale processer ikke umiddelbart er observerbare, og de kan være bevidste og ubevidste, samt affektive og personlige. Derfor er det nødvendigt at udforske kognitionen via forskellige metoder. Man kan både anvende kvantitative metoder til at undersøge teoretiske retninger, tanker og viden af et stort antal lærere via spørgeskemaer og tests, og man kan søge dybdegående indsigt i lærernes tanker via kvalitative instrumenter som interviews, fokusgruppeinterviews, dagbogsskrivning, biografiske beretninger, refleksionsessays, mindmaps osv. Desuden kan disse metoder kombineres, og de

kan komplementeres med observationer af undervisningspraksis, således at de verbaliserede overbevisninger kan sammenlignes med klasseværelsesaktivitet.

Forskellige aspekter af sproglærerkognitionen har været i fokus i forskningsfeltet. Et veludforsket område har været forskellen mellem erfarne og uerfarne lærere, men også forskellige komponenter af sprogundervisning er blevet undersøgt. Ikke overraskende har lærernes overbevisninger vedr. grammatikundervisning været bedst repræsenteret (fx Borg 1998a, 1998b, 2001; Borg & Burns 2008; Phipps & Borg, 2009; Nishimuro & Borg, 2013; Graus & Coppen, 2016), men også overbevisninger om literacy (læsning og skrivning) (fx Graden, 1996; Meijer, Verloop & Beijaard, 1999; Kuzborska, 2011) og det kulturelle fokus i sprogundervisning (fx Young & Sachdev, 2001; Xiaohui & Li, 2011; Fernández, 2015) har vakt opmærksomhed. Andre aspekter, såsom mundtlighed, lytteforståelse, anvendelse af teknologi, motivation for at lære fremmedsprog osv., er i mindre grad repræsenteret, men forskningsfeltet er produktivt, og flere og flere studier er med til at belyse tidligere uudforskede områder (i Danmark se fx Vijayarathan (2017) om mundtlighed, Kjærgaard (2018) om feedback og teknologi, samt Børsting (2018) om sproglæringsstrategier).

Som tidligere nævnt stammer lærerkognitionen i høj grad fra lærernes tidlige erfaringer med skolesystemet, men den bliver i større eller mindre grad også påvirket af andre faktorer: universitets- eller læreruddannelsen, efteruddannelsen, forskningen i sprogdidaktik, den sociale kontekst i form af kolleger, forældre, læreplaner osv., samt ikke mindst deres egen praktiske erfaring i klasseværelset (Borg, 2006). Samtidig er der ikke et én-til-én forhold mellem kognition (viden, tanker, overbevisninger) og praksis, da flere forskellige faktorer ofte kan gøre, at lærerne ikke agerer efter bedste overbevisning i deres pædagogiske beslutninger. Ministerielle retningslinjer, adgang til materialer, praktiske forhold i skolen, krav fra elever og forældre, eksamenskrav osv. kan være med til at påvirke lærernes pædagogiske beslutninger i en retning, som de ikke selv kan stå inde for. Denne uoverensstemmelse mellem kognition og praksis skal dog ikke nødvendigvis ses som udelukkende negativ, fx i tilfældet hvor lærerens holdninger er uegnede eller dysfunktionelle (Vijayarathan, 2017). Uanset hvad kan forskningsfeltet om lærerkognition og dets inddragelse i læreruddannelsen være med til at hjælpe lærerne til at ransage deres overbevisninger og deres praksis, og disses indbyrdes forhold, for at få en bedre forståelse af egen professionelle ageren (Basturkmen, Loewen & Ellis, 2004).

### **Gymnasielæreruddannelsen i Danmark**

I Danmark findes der ikke deciderede gymnasielæreruddannelser, som i så mange andre lande, men der uddannes kandidater i gymnasiefag (fx kandidater i historie, engelsk sprog, litteratur og kultur eller matematik), som så kan påtage sig funktionen som gymnasielærere, oftest med kompetence i to og evt. flere fag. Efter endt kandidatstudium på universitetet og ved fast ansættelse på et gymnasium bliver de nye lærere tilmeldt det ét-årige "Pædagogikum ved de Gymnasiale Uddannelser", som består af en praktisk og en teoretisk pædagogikumuddannelse. Teoretisk pædagogikum, som styres af et konsortium af danske universiteter, omfatter 20 ECTS og inkluderer forskellige kurser og workshops i almen- og fagdidaktik<sup>2</sup>. Praktisk pædagogikum, som hører under Undervisningsministeriet, vægter 40 ECTS og foregår på ansættelsesstedet, hvor en vejleder er med til at kvalificere kandidatens undervisning

---

<sup>2</sup> For information om det Teoretiske Pædagogikum, inkl. studieordningen, se Syddansk Universitet (u.d.)

(Undervisningsministeriet, u.d.). Selvom det er meningen, at nyansatte gymnasielærere skal igennem Pædagogikumuddannelsen i deres første arbejdsår, er virkeligheden en anden for et stort antal gymnasielærere, da mange skal igennem en række vikarstillinger (muligvis i flere år), før de opnår en fast stilling og dermed adgang til Pædagogikumuddannelsen.

Derfor er det kærkomment, at der i de senere år har været en stigende tendens til, at universiteterne i løbet af kandidatstudierne tilbyder nogle uddannelseselementer (med vægtning på op til et halvt årsværk), hvor den studerende kan tone sit studium med henblik på en konkret karrierevej, herunder undervisning. Disse studieelementer, der målretter sig undervisning, inkluderer typisk kurser i almen- og/eller fagdidaktik samt en gymnasiepraktik. På denne måde tager man hul på de kommende gymnasielæreres pædagogiske uddannelse, før det egentlige pædagogikum, og i mens de stadig studerer på universitetet. Disse pædagogiske moduler kan være med til at udruste kandidaterne bedre til et lærerjob, da de supplerer kandidatens faglige viden med principper om almen- og fagdidaktik. Men de kan også bidrage til at kvalificere de studerendes karrierevalg, især under den praksisorienterede gymnasiepraktik, hvor den studerende kan prøve kræfter med en rigtig arbejdsplads.

På Aarhus Universitet udbydes "Profilmag", som beskrives som *tværfaglige, arbejdsmarkedsorienterede fag til kandidatstuderende på de humanistiske uddannelser* (Aarhus Universitet, u.d. a). Udbuddet af profiler kan variere lidt fra år til år. I efteråret 2017 blev der fx udbudt fire profiler: Kommunikationsprofil: web- og viral kommunikation; Kommunikationsprofil: Tillægsuddannelse i journalistisk formidling; Kulturformidlingsprofil; Special profile in international collaboration and intercultural competence; Underviserprofil.

Målet for nærværende analyse er – som tidligere nævnt – profilmaget "Underviserprofilen", og næste sektion handler derfor om denne.

#### *Underviserprofilen: Faget Undervisning i teori og praksis*

De studerende, der vælger denne profil, skal følge faget "Undervisning i teori og praksis", som vægter 20 ECTS og består af tre hovedelementer: Et kursus i almindelig didaktik (ca. 16 timers undervisning), hvor den studerende får information om gymnasiet og om elevernes forudsætninger samt bliver undervist i bl.a. pædagogiske grundbegreber og elevinddragende undervisningsformer; en gymnasiepraktik (20-30 klokketimer), hvor den studerende følger én eller flere lærere i ét eller flere gymnasiefag samt får mulighed for selv at undervise; et kursus i fagdidaktik (ca. 26-39 timers undervisning) som introducerer fagdidaktiske principper og teorier. Der er fem forskellige fagdidaktiske kurser at vælge imellem, afhængigt af hvilket undervisningsfag den studerende har: Danskfagets didaktik, Design- og kunstfagenes didaktik, Historiedidaktik, Religionsvidenskabelig fagdidaktik samt Second Language Acquisition and Pedagogy.

Faget "Undervisning i teori og praksis" med ovennævnte tre elementer består gennem en portfolioeksamen, som inkluderer en praktikrapport, et refleksionsessay, et undervisningsmodul, udviklet og revideret af den studerende, samt en synopsis til et mundtligt oplæg om et fagdidaktisk emne.

Det fagdidaktiske element "Second Language Acquisition and Pedagogy", som er i fokus i denne undersøgelse, kan vælges af alle studerende på Aarhus Universitet fra et fremmedsprogligt studium (fx engelsk, spansk, fransk, tysk, arabisk, kinesisk).

Det bliver beskrevet således:

*This course will introduce students to theories of second language acquisition (SLA) and to the practical application of findings from the study of SLA to actual classroom practice. The course is intended to help students gain a clearer understanding of what is most effective in second language learning and teaching, and what the effects are of different individual factors and contexts of learning. Students will understand their own past and future SLA progress and that of their own students from the perspective of theoretical and applied linguistics. (Aarhus Universitet, u.d. b)*

Arbejdsformen i dette kursus er baseret på holdundervisning og gruppeopgaver, som løses hver uge, og som dels handler om bearbejdning af teoretisk fagdidaktisk litteratur og dels består af at udvikle, designe, præsentere og diskutere *kommunikative tasks* (dvs. problemorienterede kommunikative læringsaktiviteter) med fokus på forskellige aspekter af sprogundervisning (grammatik, ordforråd, udtale, interkulturel kompetence osv.).

Det er værd at nævne i denne kontekst, at de studerende ikke er forpligtede til at vælge underviserprofilen, selvom de har ambitioner om at blive undervisere. Udover at kunne vælge en af de andre profiler har de studerende også mulighed for i dette semester at følge undervisning på et udenlandsk universitet eller tage en erhvervspraktik enten i Danmark eller i udlandet. Dette betyder, at der er færdiguddannede kandidater, som ikke har haft dette pædagogiske element i deres studium, men som stadig har mulighed for at blive ansat på et dansk gymnasium.

### **Empirisk undersøgelse**

I denne sektion vil jeg præsentere resultaterne fra en empirisk undersøgelse foretaget med et hold studerende fra Aarhus Universitet, som valgte Underviserprofilen og derunder Second Language Acquisition and Pedagogy som deres profilssemester. Holdet består af otte studerende, hvoraf syv læser engelsk og én læser tysk. Holdet er lille, da det drejer sig om et "forårshold" for studerende, som har haft et semesters propædeutik i deres studium og derfor er forsinkede/skæve i forhold til andre studerende (om efteråret er der normalt mellem 25 og 40 studerende). Det har været en bevidst beslutning at vælge et lille hold for at kunne foretage en kvalitativ analyse (se næste afsnit). Alle otte studerende på holdet har samtykket i at deltage i undersøgelsen ved at skrive under på en samtykkeerklæring. De er blevet anonymiseret således, at de hver har fået et identifikationsbogstav (fra A til H). Forskeren er også underviseren i det pågældende kursus.

#### *Data og metode*

Jeg har valgt at undersøge emnet kvalitativt (Merriam & Tisdell, 2016), da jeg ønsker at opnå en dybdegående og kontekstualiseret forståelse af de deltagende lærerstuderendes tanker og viden om sproglæring og -undervisning. Med dette formål har jeg gennemført datatrigulering (Merriam, 2009) ved at indsamle og analysere forskellige typer data, som tilsammen er med til at belyse emnet så udtømmende som muligt.

For at belyse udviklingen i de deltagende studerendes overbevisninger omkring sproglæring og -undervisning har jeg derfor indsamlet alt det materiale, der blev produceret som en integreret del af undervisningen og eksamenen, samt andre data, som blev produceret specifikt med henblik på undersøgelsen. Data er således indsamlet ved hjælp af fem instrumenter: 1) *Spørgeskema om sprogsyn*: På første undervisningsdag udfyldte de studerende et spørgeske-

ma med både lukkede spørgsmål (fempunkts Likert-skala) og tre åbne spørgsmål<sup>3</sup> vedr. deres syn på sprog og sproglæring. Den lukkede del af spørgeskemaet kan ses i Tabel 1 og bidrager med en bred oversigt over de studerendes tanker. 2) *Transskriptioner fra gruppediskussioner*: På første undervisningsdag diskuterede de studerende deres besvarelser fra spørgeskemaet. På femte og tolvte undervisningsgang diskuterede de studerende deres besvarelser fra undervisningsgang nr. 1 på basis af deres nye viden. Disse tre sæt diskussioner blev optaget og transskriberet. 3) *Gruppebesvarelser* af i alt 10 ugeopgaver (delt af de studerende i Google docs), som de studerende arbejdede med hver uge som forberedelse til undervisningen i faget. 4) *Individuelle besvarelser af eksamensportfolien*. 5) *Transskriptioner af forsinkede individuelle, semistrukturerede interviews* (Kvale & Brinkmann, 2009) foretaget ét år efter endt profilssemester.

Your ideas about language learning and language teachers	1 (strongly agree)	2	3	4	5 (strongly disagree)
1. Languages (L1 and L2) are learned mainly through imitation		ABDFGH	E	C	
2. People with high IQs are good language learners		B	CDEFG	AH	
3. The most important factor for success in learning an L2 is motivation	AG	BCDEH	F		
4. Some people have a special gift for language learning	ACD	BEFGH			
5. It is best to learn an L2 in the same way as you learn your L1		CD	AHF	BEG	
6. The earlier an L2 is introduced in school, the greater the likelihood of success in learning	CDF	ABEG		H	
7. Most of the mistakes made by L2 learners are caused by differences with the L1		ACD	BEFGH		

<sup>3</sup> Disse tre spørgsmål var: 1) Think of successful and unsuccessful experiences you have had of learning new languages. 2) How would you learn a new language assuming you had all the time and resources? 3) How would you teach someone your first language?

8. Learners' errors should be corrected as soon as they are made in order to prevent the formation of bad habits	A	D	BCFGH	E	
9. When L2 learners interact freely (in groups or pairs), they pick up each other's mistakes		BEH	ACDFG		
10. Students learn what they are taught		CH	ADG	BEF	
11. Teachers should only use the L2 in the L2 classroom		AD	BEFGH	C	
12. Teachers should avoid explicit discussion of grammar		C	E	ABDFGH	
13. Language learners should aim to sound like native speakers		CD	FH	ABEG	
14. Fluency is more important than accuracy in L2 speech		AFH	BCDEG		
15. L2 learners should take responsibility for their own learning	EH	ABCDFG			
16. It is necessary to know the L2 culture(s) in order to speak an L2	BH	CF	DEG	A	
17. Learning an L2 is mostly about learning grammar rules			DH	ABE	CFG
18. Sufficient L2 proficiency is the most necessary to become a qualified teacher		DH	ABEFG		C

Tabel 1. Spørgeskema med de studerendes besvarelser (de studerende er identificeret fra A til H) (Tilpasset fra Lightbown & Spada 2011, s. xvii)

Proceduren for dataanalysen bestod i en tematisk- og indholdsanalyse af materialet (Dörnyei, 2007). Materialet blev læst gentagne gange for at finde de emner og kategorier, som fremstod som relevante for besvarelsen af undersøgelsens mål, med fokus på tre nøglelementer: før, under og efter semestret.

## Analyse og diskussion

### *Før profilsemestret*

Ved besvarelsen af den første refleksionsøvelse på den første undervisningsgang viste de otte studerende ret stor enighed i deres tanker omkring sproglæring. Svarene til de lukkede spørgsmål i spørgeskemaet viser meget lidt spredning (se Tabel 1 ovenfor): I de fleste spørgsmål er der mindst fem studerende, som har valgt den samme valgmulighed, og generelt er maks. tre ud af fem valgmuligheder benyttet (med undtagelse af spørgsmålene 8 og 16) med konsekvent fravælgelse af de ekstreme valg: *meget enig* og især *meget uenig* (sidstnævnte kun valgt i to spørgsmål af i alt tre studerende). Denne tendens til at svare med den neutrale mulighed (3) eller de nærliggende 2 og 4 kan have flere fortolkninger. Det kan forstås som tegn på usikkerhed (Cota Grijalva & Ruiz-Esparza Barajas, 2013) – og én studerende undskylder sig således med en lille note på spørgeskemaet: *I am sorry for all the 3's* – men man kan også se det som et tegn på, at det er svært at svare uden kontekst, da svaret kan afhænge af mange forskellige faktorer. I transskriptionerne fra den første gruppe-refleksion, hvor de studerende reflekterer over deres svar, kan man finde evidens for begge fortolkninger: *I wasn't sure about it, so I put it in the middle one again<sup>4</sup>*, men også *Midt i – altså jeg synes lidt, at det kommer an på så meget*.

I denne første gruppediskussion er der et emne, som meget hurtigt kommer til at optage en del af samtalen, og det er grammatikundervisning. Denne interesse for grammatikkens funktion i sprogundervisning og -læring afspejler den centrale rolle, som grammatikundervisning har haft i den fagdidaktiske litteratur (Fernández, 2014), og her er de fleste studerende enige om, at der skal undervises eksplicit i grammatik (spørgsmål 12), men at grammatikken ikke er alt (spørgsmål 17). Der viser sig nogle modsigelser i de studerendes svar, som vidner om en vis spænding mellem kommunikativ undervisning og mere traditionel grammatikundervisning: På den ene side fremhæver de studerende den kommunikative tilgang, med anvendelse af sproget ved at tale og lytte, som svar på det åbne spørgsmål om deres mest succesfulde oplevelser med sproglæring. Ligeledes som svar på det åbne spørgsmål, om hvordan de gerne vil lære et nyt sprog, foretrækker de fleste at lære gennem at rejse og bruge sproget (kombineret med undervisning); samtidig nævner de kedelig og irrelevant grammatik som svar på det åbne spørgsmål om deres negative oplevelser med sproglæring. På den anden side, når de svarer på, hvordan de ville undervise andre i deres modersmål, vil de fleste dog undervise i grammatik (nok starte med basisordforråd og -sætninger). Lignende lidt modstridende tanker om grammatik og kommunikativ undervisning er fundet i andre undersøgelser om lærerkognition blandt lærerstuderende (fx Graus & Coppen, 2016).

Et andet emne, der optager gruppediskussionen på semestrets første time, er kulturkomponenten i sprogundervisning (spørgsmål 16). Dette er det spørgsmål, som viser mest uenighed blandt de studerende, og en del af forklaringen kan være, at de studerende er usikre på, hvad man mener med kultur:

---

<sup>4</sup> Nogle grupper diskuterede på engelsk, og nogle på dansk. Undervisningen foregik på engelsk.

*G: But I'm like, isn't language such a big part of culture? Like by learning the language you also necessarily learn about the culture.*

*D: But before I started here I didn't know much about the civil war or anything in America, and I kind of could speak American as well...*

*G: Maybe it depends on the language then.*

*D: And it depends on what kind of culture. Because if you need to talk about the civil war of course you need to know about it. There are times where you need to know the vocabulary.*

*G: But also like what part is culture and what part is history. Or is it more like there are some full pass on how you do stuff. Like when you bump into someone you say sorry. Or excuse me. That's more necessary.*

*H: That's more language than culture then. Also saying "how are you doing?" every time you meet someone. That's another thing when you go to an English-speaking country.*

*D: There are all those idioms and stuff like that. That you wouldn't know.*

*G: Yeah and if there are specific words of racism that are kind of bad...*

*H: Yeah well if you look at it from that point, then I think it is important. But in general you don't need to know everything about a country's culture.*

*D: I think I thought of it more as history than as culture.*

*H: Yeah me too. (Optagelse 2)*

Usikkerheden omkring den kulturelle/interkulturelle komponent i sprogundervisning er vel-dokumenteret i litteraturen om lærerkognition, og det gælder ikke alene lærerstuderende men også erfarne lærere (Fernández, 2015), og kulturkomponenten anses som den del, der meget nemt kan blive tilsidesat til fordel for det rent sproglige (Young & Sachdev, 2011).

Andre emner, der bliver berørt i diskussionen, er anvendelsen af apps, som noget flere studerende fremhæver som værende nyttigt for at lære sprog, og vigtigheden af en god lærer, som bliver nævnt af flere som grunden til deres positive erfaringer med sproglæring. I diskussionen om spørgsmål 18 er det tydeligt, at de studerende er enige om, at der er flere faktorer, der udgør en god lærer, og gode sproglige kompetencer på fremmedsproget er kun én af dem. At være en god formidler og at skabe et godt socialt miljø i klasseværelset er også påskønnet. Flere studerende fortæller anekdoter om gode og dårlige lærere, de har oplevet, og emnet virker til at fylde en del i deres tanker om sproglæring. Dette er i overensstemmelse med andre undersøgelser om fx elevmotivation, som viser, at læreren kan være en afgørende faktor (Knudsen, 2017; Pedersen, 2017).

Alt i alt kan man sige, at disse studerende kom til profilmestret med enslydende tanker omkring sproglæring og med visse usikkerheder. Enigheden kan forklares ved, at de deler lignende livshistorier om sproglæring, da de alle har lært sprog i det danske skolesystem, og vi ved fra forskningen i lærerkognition, at det er præcis disse erfaringer med egen læring, som er med til at forme vores tanker om, hvad der er den rigtige og forkerte måde at undervise i og lære sprog (Lortie, 1975; Borg, 2003; 2015). Det er interessant, at de studerende ikke snakker om det, de har lært på universitetet (med undtagelsen af én studerende, som næv-



ner, at han har haft noget om alder og sproglæring (jf. spørgsmål 6) i et fag på universitetet: *Jeg kan bare huske, at jeg blev præsenteret med en masse undersøgelser*). Manglende referencer til teorier og akademisk litteratur stemmer overens med litteraturen om lærerkognition, som viser, at lærernes tanker udformes tidligt (Pajares, 1992, s. 328), og at lærerne sjældent refererer til didaktisk litteratur (Ebsworth & Schweers, 1997 i Borg, 2015).

### *I løbet af profilssemestret*

I løbet af semestret blev de studerende udsat for tre forskellige typer af gruppeopgaver, hvor de skulle: 1) svare på teoretiske spørgsmål, hvor de skulle forholde sig til den akademiske litteratur (fx "What contribution has *behaviorism* made to our understanding of how languages are learned?"); 2) reflektere over forskellige aspekter af sproglæring og dermed eksplicite egne tanker (fx "How important is the teaching of culture in the context of L2 teaching as compared to other components of the language in your opinion and in real practice - you can include insights from your observations at the practicum" eller "Complete the learning style test that I enclose and based on your result, how do you think your particular learning style can affect your teaching style?"); 3) designe kommunikative tasks med forskellige fokusområder (fx "design a writing task that may also incorporate elements of reading, is collaborative, is multimodal and includes some element of feedback").

Disse opgaver havde til formål netop at give de studerende mulighed for at italesætte deres tanker vedr. sproglæring og diskutere og sammenligne dem med deres klassekammerater. Dermed ville de forhåbentligt nå frem til en ny forståelse - i overensstemmelse med den fagdidaktiske litteratur om lærerkognition, som peger på vigtigheden af at diskutere eksisterende tanker i løbet af læreruddannelsen (Pajares, 1992; Fives & Buehl, 2012). De mere praktiske og kreative opgavers sigte var at give de studerende mulighed for at omsætte tankerne til praksis, dog stadig kun "hypotetisk" praksis, da de designede tasks ikke nødvendigvis skulle afprøves i virkeligheden (nogle tasks blev dog faktisk afprøvet i gymnasiepraktikken).

En gennemgang af besvarelserne fra de forskellige typer øvelser viser en vis fremgang i de studerendes sikkerhed samt voksende forståelse. Fx kan man i de kreative opgaver om task-design se en forskel mellem de første tasks og de efterfølgende, og en del af forbedringen kan måske relateres til de klasses Diskussioner, hvor hver designet task blev kommenteret og undersøgt på basis af kravene til en kommunikativ task (= at der er et informationshul, at eleverne kan anvende egne sproglige ressourcer, at der er fokus på indhold, og at der er et kommunikativt output (Svendson Petersen, 2001)).

I undervisningsgang nr. 5 blev de studerende bedt om at diskutere deres besvarelser fra spørgeskemaet fra undervisningsgang nr. 1 i grupper for at se, hvorvidt de stadig ville svare på samme måde. I optagelserne fra denne gruppediskussion bliver det tydeligt, at de studerende mener, at det er for tidligt til at kunne revurdere deres tanker, især fordi de ikke selv har kastet sig ud i praksis:

*- I just feel like I still need a little more information to really feel sure about this*

*- But we've already got a lot of information, it's just that we don't know how to feel about that information yet. (Optagelse 4)*

*- Det er jo bare tekster i sidste ende, og jeg tror, at noget af det skal du selv ud at prøve, før at det... giver mening. (Optagelse 3)*

Ikke desto mindre er der én kommentar, som gentager sig i optagelserne: At det hele er blevet mere "nuanceret". Denne nuancering udmunder sig i, at nogle studerende føler, at de nu har forstået, at der ikke kun er én rigtig måde at gøre tingene på:

*- Jeg synes også, som du siger, at det er bare blevet mere nuanceret. Altså jeg ikke ved rigtig sådan helt sådan... Der er i hvert fald et par stykker, hvor jeg tænker sådan: "Nå, men det kunne godt være, at det ikke var helt rigtigt, det jeg egentlig havde tænkt". (Optagelse 3)*

*- I think I am more lenient now than I was before because I've seen all the methods and also just looking at the tasks today, the way that you can do tasks in the classroom that was really interesting. I think that might have shifted it a bit for me. (Optagelse 4)*

I sidste undervisningsgang blev der plads til en lignende gruppediskussion, denne gang med fokus på tankerne omkring kulturkomponenten, som var timens emne. Igen er "nuancering" et nøgleord i forbindelse med de studerendes udvikling:

*- Jeg har lært, at man skal være lidt mere påpasselig på, hvad det er for nogle stereotyper, man giver videre til eleverne.*

*- Det handler om at være nuanceret, tror jeg. Så meget man nu kan, og så ikke have berøringsangst. (Optagelse 5)*

Men også en ny bevidsthed om kulturkomponentens rolle kommer til syne:

*- Det var det her med, at i læreplanerne at det kommer til at have så stor betydning, at sådan et kulturelt element skal med. Det synes jeg også var meget overraskende det her. Altså. Det er næsten sådan opdragende facetter (...), at de bliver bedre mennesker, og at man skal bruge kulturen måske, og ja, så undervisningen i det til at få dem derhen.*

*- For at blive mere nuancerede og blive ordentlige mennesker, der kan rumme flere... altså det giver god mening, men man... hold da op. Det er godt nok en stor mundfuld i en læreplan, man lige har ansvaret for. (Optagelse 5)*

Ved slutningen af semestret kan man også mærke i de studerendes kommentarer, at praksis har sneget sig ind og påvirket deres kognition. Fx bruger gruppen en del diskussionstid på at snakke om, hvordan man skal forholde sig til kulturkomponenten, når man har en meget multikulturel elevgruppe – og dette på basis af én af de studerendes erfaring fra sin praktik.

#### *Efter profilsemestret*

Ca. to uger efter endt undervisning i profilsemestret skulle de studerende aflevere deres eksamensportfolio, og for denne analyse er der interessant materiale at hente i både refleksionsessayet, hvor de studerende skriver deres vurdering af semestrets samlede udbytte, og praktikrapporten, hvor de er bedt om at forholde sig til den observerede og selvudførte undervisning på basis af relevant teoretisk litteratur fra semestret. Fælles for de otte studerende er, at de vurderer semestrets udbytte meget positivt. Det er uden tvivl gymnasiapraktikken, som især vækker begejstring, men de påskønner også de teoretisk didaktiske forløb som nødvendig basis:

*G: What I learned most from was being out in my practicum, but I would not have felt terribly comfortable out there, if I did not feel I had a good theoretical background to back up what I was doing.*

Semestret virker til at have haft en indflydelse på disse studerendes viden og overbevisninger vedr. sproglæring på forskellige måder. For det første introducerede det fagdidaktiske kursus aspekter om deres fag, som de ikke var blevet præsenteret for tidligere i deres studium:

*A: The fact that there are so many different aspects to second language teaching, and teaching in general, has been a real eye-opener. As I had not had similar subjects before, it was very relevant to get some theory on how to deal with teaching in practice.*

*H: Before this course, my own abilities in regards to my studies in English and classical studies had been very much in focus. When I thought about teaching, I mostly thought about all the interesting components of my subject. The aspect of student learning and all the different ways of teaching had not really crossed my mind.*

Ligesom var tilfældet i løbet af semestret i de studerendes diskussioner, fokuserer disse slutrefleksioner om især tre aspekter: kommunikativ undervisning, tilgangen til grammatik, samt interkulturel kompetence:

*F: Da jeg selv er opdraget i grammatik og oversættelses-metoden og den generative-teori så var det befriende i SLA-forløbet at få præsenteret et mere nuanceret syn på sprogtilgængelsen (...), hvor de grammatiske fænomener præsenteres i en relevant kommunikativ kontekst.*

*E: I noticed an ongoing theme through all the English classes I observed, namely a strong presence of the communicative approach. (...) When I think back on my English classes in gymnasium, I remember fluency and communication skills as important, but it is not until after I have tried being on the teacher's side of the classroom that I see just how prominent the oral part of English language teaching and learning is.*

*F: Denne opdagelsesproces blev på glimrende vis stilladseret af vores lektion om interkulturel kompetence. Efter forløbet er jeg således også langt bedre klædt på i forhold til at drage nytte af de forskelle, som ifølge Claire Kramsch opstår og vedligeholdes gennem sproglige/symbolske interaktioner mellem mennesker. I min praktik hostede jeg også brugbare erfaringer i forhold til at kunne begå mig i det kulturelt heterogene klasseværelse.*

En gennemgang af besvarelserne viser tydeligt, at det fagdidaktiske kursus ikke alene har bidraget med en introduktion til sproglærings- og didaktiske teorier, men gennem de stillede opgaver, klasses Diskussionerne og familiariseringen også med begrebet "kommunikativ task":

*D: It opened up some great discussions and we started sharing our opinions of how we would do things when we get a job sometime in the near future. This notion of sharing our experiences and ideas is a very valuable lesson to have learned, because we will need to do this in our future jobs.*

*F: Dernæst fandt jeg også brugbart at arbejde i praksis med Ellis' begreb om Task-baseret læring. Det gav mig en konkret model for hvordan den kommunikative undervisning kan udformes i praksis. Sammen med Klafkis alment didaktiske model, forkortet FIMME på dansk, har jeg nu langt bedre mulighed for at tilrettelægge kvalitets undervisning.*

De studerendes svar vidner også om tilegnelsen af et fælles fagligt sprog. De har nu begreber, de kan bruge for at snakke om de forskellige aspekter af sprogundervisning og for at analysere deres observationer i gymnasiepraktikken samt deres egen undervisning:

*D: The teaching style that I observed with XX was influenced by the communicative approach, and made use of both Focus on Form and Focus on Forms centered teaching.*

Dette faglige sprog, som også afspejler en tilegnet *pedagogical content knowledge* (Shulman, 1996), kan ifølge lærerkognitions litteratur være med til at udforme de kommende læreres professionelle identitet. Desuden er der i nogle studerende tegn på en spontan erkendelse af deres "apprenticeship of observation", måske mest elokvent i følgende citat:

*F: Helt overordnet har jeg under forløbet på undervisningsprofilen fået transformeret mine forestillinger om det at være underviser, sådan at jeg nu er i stand til teoretisk at kvalificere de sekvenser som kommer til at være grundstenene i min undervisning med udgangspunkt i elevernes behov og motivation, lærerplanernes krav og mine kompetencer og passioner. Hvor jeg før forløbet ubevidst havde opereret med fagformidler-modellen og den dertil hørende forestilling om at viden uproblematisk flyder fra underviseren til eleven, har jeg gennem forløbet og i min praktik erfaret at det at være underviser fordrer en reflektiv tilgang til egen praksis (...) Især introduktionen til det sociokulturelle paradigme med dets insisteren på at læringen primært foregår i et socialt samspil mellem mennesker, der handler i kraft af vores fælles sprog. Dette fokus fik mig til for alvor at få øjnene op for elevens situation og fordringer. Hvor jeg før uden at tænke over det var tilbøjelig til at hævde, at en stærk videns base ville være tilstrækkelig for en underviser til at sikre god undervisning. Jeg kan nok ikke i den sammenhæng udelukke, at jeg ubevidst har overtaget nogle af mine tidligere og nuværende underviseres tilgange.*

Alt i alt har profilssemestret ikke alene kvalificeret de studerende teoretisk og praktisk, men har også bidraget til at bekræfte/bestemme deres fremtidige karrierevalg, som følgende udsagn viser:

*D: The internship has confirmed my choice of career.*

*E: It gave me a much clearer picture of what it is like to be a teacher than I had beforehand. (...) It is lovely reassuring to know that I have chosen the right education and the right career path.*

#### Ét år efter

Fem ud af de otte studerende mødte op til et interview ét år efter, kurset var afsluttet (de resterende tre var enten syge, bortrejste eller ikke til rådighed). Alle fem fremmødte havde brugt året på at skrive deres speciale (dvs. de var alle blevet kandidater), og tre af dem havde fået vikarjobs som gymnasielærere, mens to var jobsøgende og havde ikke fået mere undervisningserfaring (eller meget lidt) siden praktikken året før. Set i bakspejlet var det fælles for alles vurdering af profilssemestret, at det havde været yderst nyttigt, og at de ikke ville have undværet det. En af de studerende gav udtryk for, at profilssemestret skulle været kommet tidligere i studiet:

*E: Jeg synes jo, at hele min universitetsuddannelse egentlig ikke ledte hen imod, at jeg skulle undervise, som sådan. Det er først kommet på profilssemestret. Og selvom jeg har vidst fra starten af, at det var det, jeg ville, så har jeg jo haft en masse fag, som bare var dem, jeg skulle igennem. (...) Fordi jeg kun har haft den der profil et semester, inden jeg skrev speciale... jeg har ikke været inde i alle de der begreber og ting og sager, som kunne have været rare at have haft fra starten, så man var mere klædt på til at skulle ud og undervise.*

Indholdsmæssigt virker det som om, at det mest er de mere generelle didaktiske komponenter af profilstemestret, som er præsente i deres tanker. Når de tidligere studerende bliver spurgt om de nøglebegreber, som de har med fra semestrets undervisning, er *motivation, variation, medbestemmelse* og *at skabe trykke rammer* gengangere i deres svar. Det er disse aspekter, hvor de tre deltagere, som er i job, har lagt deres kræfter i de meget hektiske første måneder som lærere, og det er også det, de jobsøgende tænker, at de vil fokusere på, når de får lov at undervise.

Til gengæld kniber det meget mere med at huske specifikke didaktiske teorier om fremmedsprog, som har gjort indtryk på dem, og nogle af de centrale komponenter, som havde optaget de studerende gennem semestret (grammatik, ordforråd, interkulturel kompetence, kommunikativ undervisning, herunder kommunikative tasks), var tydeligvis blevet forsømt. I forhold til grammatik vidner deltagerens svar om, at de underviser efter meget mere traditionelle modeller, end dem der blev foreslået, diskuteret og øvet i profilstemestret. Studerende C nævner, at hendes undervisning er "meget klassisk" og følger rækkefølgen i grammatikbogen, "hvor jeg vurderer, at det er nogenlunde sådan, at det stiger lidt i sværhedsgrad". Studerende E nævner, at hun vil arbejde "kronologisk" med grammatik og vil gerne gå igennem alle ordklasser og kommatering; kontekstualiseret grammatik er ikke noget, hun har prøvet.

Ordforrådstilegnelse er ikke noget, som disse nye lærere har fokuseret specifikt på endnu (selvom de gerne vil i fremtiden), og ingen af dem har følt sig sikre nok til at arbejde med kommunikative tasks:

*C: Som ny underviser tvivler jeg rigtig meget på, er det nu rigtigt, det jeg gør? Er det godt nok? Der synes jeg, det kan være enormt rart at benytte sig af nogle bøger, eller nogle forløb, der er lavet. Jeg har fået nogle af min kollega. Der tænker jeg, at det der med at skulle lave en task, det er også lidt... får de nu noget ud af det? Gør jeg det rigtigt?*

Kulturkomponenten har alle tre nye lærere inkluderet i deres undervisning i form af forskellige emner, nogle gange med fokus på information om engelsktalende lande eller en periode (fx 1920'erne), men de forskellige facetter af modeller som Byrams (1997), som blev ekstensivt diskuteret i profilstemestret, har ikke været i fokus.

Disse resultater, som viser at de nye sproglærere pga. tidspres og usikkerhed fokuserer mere på klasseværelsesledelse end konkrete sprogdidaktiske aspekter, stemmer overens med internationale resultater (Numrich, 1996 i Borg, 2015). De sprogdidaktiske aspekter, som var centrale i profilstemestret, er derfor ikke nødvendigvis glemte eller tabte, og de vil forhåbentligt komme i spil hen ad vejen, når der er mere overskud:

*E: [Om profilstemestrets indhold] Det er meget at få på én gang. Når man så kommer ud, og skal prøve det i praksis, så kan man bare ikke få det hele med fra starten af, selvom man gerne ville. (...) [Om at arbejde med sproglæringsstrategier] Jeg har, men ikke så meget som jeg gerne ville. Men igen, det har været det der med, at det har jeg egentlig ikke haft overskud til. At skulle tænke det ind i alt. Men det kan komme, på et tidspunkt, og det håber jeg da, at det gør.*

Deltagerne er derfor glade for, at noget af det samme indhold, som blev gennemgået i profilstemestret, vil blive behandlet igen, når de kommer til Teoretisk Pædagogikum. De ser en mulig gentagelse som en fordel snarere end en ulempe.

De forsinkede interviews viser, at disse kandidaters – og splinternye sproglæreres – kognition er blevet påvirket af deres profilsemester, men ikke alle disse tanker og den nyligt tilegnede viden afspejles – i følge egne udsagn – i deres praksis, da eksterne faktorer som tidspres og interne faktorer som usikkerhed kan være med til at bremse nogle af de tiltag, som de ellers gerne vil implementere. Dette stemmer overens med det, som vi ved om lærerkognition: At der ikke findes et én-til-én forhold mellem tanker og praksis (Fives & Buehl, 2012), og at didaktisk fornyelse tager tid (Wagner, 1991).

### Sammenfatning

I denne artikel har jeg forsøgt at belyse, hvordan et semester på universitetet, som er specifikt målrettet til at give kandidatstuderende en didaktisk toning i deres studium, har påvirket en gruppe sprogstuderende. Undersøgelsen af deres tanker før, under og efter semestret samt deres evt. praksis efter overstået studium har vist, at semestret har været et vigtigt studieelement i deres bevidsthed, som de ikke synes, at de ville kunne undvære. De kom til semestret med enslydende tanker og usikkerheder, og de fik gennem forskellige arbejdsformer ny viden om almen- og fagdidaktik – to områder som ellers havde været helt fraværende fra deres studier. Især gymnasiepraktikken gjorde stort indtryk på dem og var med til at bestemme eller bekræfte deres fremtidige karrierevej. Ikke al den nye viden er blevet omsat til praksis i deres første tid som lærere, og det er især elementer som klasseledelse, de har haft mest overskud til at arbejde med. De opnåede resultater i denne undersøgelse stemmer overens med tidligere studier om lærerkognition hos lærerstuderende og viser, at en gymnasieprofil, eller andre former for didaktiske studieelementer, allerede på kandidatstudiet, og før den officielle teoretiske pædagogikumuddannelse for gymnasielærere, er et positivt tiltag, selvom ikke alt den tilegnede viden appliceres i klasseværelset med det samme.

### Referencer

- Aarhus Universitet (u.d. a). Profilmag. Lokaliseret den 6. marts 2018 fra <http://studerende.au.dk/studier/fagportaler/arts/undervisning/profilmag/>
- Aarhus Universitet (u.d. b). Undervisning i teori og praksis. Lokaliseret den 6. marts 2018 fra <http://kursuskatalog.au.dk/da/course/81062/Undervisning-i-teori-og-praksis>
- Basturkmen, H. Loewen, S. & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics*, 25(2), s. 243-272.
- Borg, S. (1998a). Talking about grammar in the foreign language classroom. *Language Awareness*, 7, 159-75.
- Borg, S. (1998b). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, 32 (1), 9-38.
- Borg, S. (2001). Self-perception and practice in teaching grammar. *ELT Journal*, 55(1), 21-29.
- Borg, S. (2003). Teacher cognition in language teaching: a review of research on what language teachers think, believe, know and do. *Language Teaching*, 36, 81-39.
- Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, 10, 3-31.

- Borg, S. 2006 (2015). *Teacher cognition and language education. Research and practice*. 2. edition. London: Bloomsbury.
- Borg, S. (2009). Teacher research in language teaching: A critical analysis (Cambridge Teachers' Conceptions of Research', *Applied Linguistics*, 30 (3) pp. 355-388.
- Borg, S. & Burns, A. (2008). "Integrating Grammar in Adult TESOL Classrooms." *Applied Linguistics*, 29, 456-482.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Bristol: Multilingual Matters.
- Børsting, L. D. (2018). *Danish Upper Secondary School English Teachers' Cognitions on Learning Strategies*. Upubliceret speciale. Aarhus Universitet.
- Cota Grijalva, S. & Ruiz-Esparza Barajas, E. (2013). Pre-service teachers' beliefs about language teaching and learning: A longitudinal study. *PROFILE* 15, 1, s. 81-95.
- Daryai-Hansen, P. & Henriksen, B. (2017). Lærerkognition som centralt udgangspunkt for en efteruddannelsesmodel til udvikling af lærernes undervisningspraksis. *Studier i Læreruddannelse og -profession*, 2, nr. 2.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Fernández, S. S. (2014). Fokus på form. I H. L. Andersen, S. S. Fernández, D. Fristrup, & B. Henriksen (red.), *Fremmedsprog i gymnasiet: teori, praksis og udsyn* (Kapitel 4, s. 43-54). København: Samfundslitteratur.
- Fernández, S. S. (2015). Concepciones de los profesores daneses acerca de la competencia intercultural en la clase de español como lengua extranjera. *Milli Mála - Journal of Language and Culture*, 7, 95-120.
- Fives, H. & Buehl, M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? I K. Harris, S. Graham & T. Urdan (eds.) *Educational Psychology Handbook*. Vol. 2, s. 471-499.
- Graden, E. C. (1996). How Language Teachers' Beliefs about Reading Instruction are Mediated by their Beliefs about Students. *Foreign Language Annals*, 29:3, s. 387-395.
- Graus, J. & Coppen, P. (2016). Student teacher beliefs on grammar instruction. *Language Teaching Research*, 20(5), s. 571-599.
- Hattie, J. (2003). Teachers Make a Difference: What is the research evidence? Presented at the *Building Teacher Quality: What does the research tell us ACER Research Conference*, Melbourne, Australia: Conference Archive at ACEReSearch. Retrieved from [http://research.acer.edu.au/research\\_conference\\_2003/4/](http://research.acer.edu.au/research_conference_2003/4/)
- Kjærgaard, H. W. (2018). *Technology-mediated written corrective feedback in the Danish lower secondary classroom*. Upubliceret phd-afhandling. Aarhus Universitet.
- Knudsen, M.H. (2017). *Den motiverende spanskundervisning – et empirisk studie af danske gymnasieelevers motivation for at lære spansk*. Upubliceret speciale. Aarhus Universitet.
- Kuzborska, I. (2011). Links between teachers' beliefs and practices and research on reading.

*Reading in a Foreign Language*. 23: 1, s. 102-128.

- Kvale, S. & Brinkmann, S. (2009). *Interview. Introduktion til et håndværk*. København: Hans Reitzels Forlag.
- Lightbown, P. & Spada, N., 2006- (2011). *How languages are learned*. (3. edition). Oxford: Oxford University Press.
- Lortie, D. C. (1975). *Schoolteacher: a sociological study*. Chicago: University of Chicago Press.
- Meijer, P., Verloop, N. & Beijaard, D. (1999). Exploring language teachers' practical knowledge about teaching reading comprehension. *Teaching and Teacher Education* 15, s. 59-84.
- Merriam, S. B. (2009). *Qualitative research: a guide to design and implementation*. San Francisco: Jossey-Bass.
- Merriam, S. B., & Tisdell, E. J., 1955-. (2016). *Qualitative research: a guide to design and implementation* (4. edition). San Francisco, CA: Jossey-Bass.
- Nishimuro, M. (2013) Teacher Cognition and Grammar Teaching in a Japanese High School. *JALT Journal*, 35, 1, s. 29-50.
- Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning up a Messy Construct. *Review of Educational Research*, 62(3), 307-332.
- Pedersen, S. (2017). *Motivation factors for learning English among Danish upper secondary school students*. Upubliceret special. Aarhus Universitet.
- Phipps, S. & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System* 37, s. 380-390.
- Schön, D. (2001). *Den reflekterende praktiker. Hvordan professionelle tænker når de arbejder*. København: Klim.
- Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(2), s. 4-14.
- Svendsen Pedersen, M. (2001). Task Force. Et bud på kommunikativ sprogundervisning. *Sprogforum* 20, 7, s. 7-19.
- Syddansk Universitet (u.d.), Teoretisk Pædagogikum. Lokaliseret den 6. marts 2018 fra [https://www.sdu.dk/da/Om\\_SDU/Institutter\\_centre/IKV/Uddannelse/Paedagogikum1](https://www.sdu.dk/da/Om_SDU/Institutter_centre/IKV/Uddannelse/Paedagogikum1)
- Undervisningsministeriet (u.d.) Pædagogikum ved de Gymnasiale Uddannelser. Uddannelsesguiden. Lokaliseret den 6. marts 2018 fra <https://www.ug.dk/uddannelser/andreerhvervsrettedeuddannelser/andreuddpaedogpsyk/paedagogikum-ved-de-gymnasiale-uddannelser>
- Vijayarathan, K. (2017) *Teacher cognition of grade 8 teachers on teaching speaking in English as a foreign language in The Faroe Islands and its impact on teachers' pedagogical praxis: seven case studies*. Upubliceret phd-afhandling. Aarhus Universitet.
- Wagner, J. (1991). Innovation in Foreign Language Teaching. I R. Phillipson et al. (eds.). *Foreign/Second Language Pedagogy Research*. Multilingual Matters, pp. 288-306.
- Woods, D. (1996). *Teacher cognition in language teaching: beliefs, decision-making, and class-*



*room practice*. Cambridge: Cambridge University Press.

Xiaohui, H. & Li, S. (2011) Teacher Cognition of Intercultural Communicative Competence in the Chinese ELT Context. *Intercultural Communication Studies* XX: 1, s. 175-192.

Young, T. & Sachdev, I. 2011. Intercultural communicative competence: exploring English language beliefs and practices. *Language awareness, Vol.20, 2*, 81-98.

# A flowchart as a tool to support student learning in a laboratory exercise

Dominik K. Grosskinsky<sup>a,1</sup>, Katrin Hammer úr Skúoy<sup>b</sup>, Kirsten Jørgensen<sup>c</sup>

<sup>a,c</sup>Department of Plant and Environmental Sciences, University of Copenhagen, <sup>b</sup>Department of Science Education, University of Copenhagen

Practice paper, peer-reviewed

Practical laboratory exercises are an essential part of university education in natural sciences. However, the multitude of positive aspects of this active and lively teaching approach is accompanied by some challenges, which have to be considered by the teacher(s) of practical exercises. In this project, a strategy was designed, employed, and evaluated to support teaching and student learning in laboratory exercises spread over several days by implementing a flowchart as a central component. Initial co-construction of the flowchart with the students gives ownership of the exercise to the students and forms a common basis for communication and interaction as well as a point of reference throughout the exercise. This approach supported student learning as evidenced by increased understanding of the content and the ability to connect individual parts of the exercise. In addition, it allows the teacher to easily track student progress.

## Introduction

In natural sciences, practical exercises, including laboratory work, are a central part of student education (Hofstein & Lunetta, 2004; Reid & Shah, 2007). Even though aspects such as level of education, e.g., BSc vs. MSc studies, and the specific discipline affect the design of meaningful practical laboratory exercises (Hofstein & Lunetta, 2004; Hofstein & Mamlok-Naaman, 2007), they are generally a suitable approach to combine various positive aspects. Students get the chance to connect theoretical knowledge with practical applications, and subsequently, may better internalize theoretical knowledge. In addition, students have the indispensable opportunity to gain hands-on experience in important methods they may need to address scientific questions later in their career. Finally, it is a good way to diversify teaching beyond lecturing and other theoretical teaching approaches, which appear to be beneficial *per se* for diverse learners (Tanner & Allen, 2004).

Planning and running practical laboratory exercises in natural sciences can be challenging, especially when working with biological systems, which often are relatively complex. Experimental work can be lengthy, which means that individual work steps have to be spread over several days. This increases the students' challenge of following and connecting different work steps and thus, the teachers' challenge of efficiently supporting student learning. The students' development in the context of the learning goals of the course, their awareness of the current status of their work, and their understanding of the dynamically

---

<sup>1</sup> Contact: dkg@plen.ku.dk

developing content have to be observed and supported over several days. Therefore, reliable and efficient tools to help a teacher cope with these challenges in class to maintain a positive and supportive learning environment for students are very valuable.

Flowcharts are tools to visually break down complex information into individual building blocks and how the blocks are connected. In practical exercises in natural sciences, complex experimental approaches could be represented by connected sequences of individual work steps and several parallel approaches may be interlinked at different levels of such a flowchart. The visual breakdown facilitates the accessibility of information and could, therefore, support student understanding of experimental procedures and link them to theoretical knowledge (Davidowitz & Rollnick, 2001). They can be used to prepare students for the laboratory work and encourage them to read instructions (Davidowitz & Rollnick, 2001). Following a suitable implementation strategy, which contains aspects of knowledge integration and monitoring, flowcharts can be a teacher's tool in laboratory teaching to enhance students' metacognitive learning (Davidowitz & Rollnick, 2003). In addition, flowcharts seem to be well appreciated by students in natural science laboratory exercises (Davidowitz & Rollnick, 2003). Thus, flowcharts can be useful to support students' understanding of the content of an exercise and their progress in the practical work.

The implementation of a flowchart for a practical laboratory exercise in plant science at BSc level is presented here. The outline of the flowchart is pre-designed and further construction is performed in class in cooperation with the students. This concept was specifically tailored to the respective teaching unit and was designed within the framework of a teaching development project as part of a university pedagogical education program. The implementation of the flowchart is aimed at preparing the students for the individual work steps, giving them a framework for connecting individual work steps and sequences, and being able to evaluate the results. For the teacher, different activities composed in an implementation plan should simultaneously serve the purpose of tracking the development of the students, providing regular feedback and flexibly reacting to the needs of the students, guiding them through the exercise. Furthermore, the flowchart should serve as a basis for teacher-student communication, in which different ways of using the flowchart should contribute to diversifying the classroom by giving changing roles to the teacher and students (Beck & Gottlieb, 2002). All these aspects should contribute to improving student learning in this practical exercise.

### **Setup and Methodology**

The presented teaching approach was implemented in the laboratory exercise "Tracking Gene Expression," which was part of the course Plant Genomics in 2017. This course is embedded in the BSc programs in Biology-Biotechnology and Natural Resources at the Department of Plant and Environmental Sciences at the University of Copenhagen. Eleven students in their 2<sup>nd</sup> or 3<sup>rd</sup> year of BSc studies participated. For the laboratory exercise, the students were divided into four groups (three groups of three students, one group of two students). The presented teaching approach was implemented as a project to improve student learning (Grosskinsky, 2018) as part of the teacher's pedagogical education (Teaching and Learning in Higher Education Program/Universitetspædagogikum; Department of Science Education, University of Copenhagen), supervised by the pedagogical supervisor K. Hammer úr Skúoy and the department supervisor K. Jørgensen.

### *The teaching scenario*

The program of the exercise was completely changed based on student feedback from the previous year and because a new teacher was responsible for the exercise. The experimental work was spread over three days with a gap of six days between the first two days and the last day of the exercise. The new exercise comprised three individual experimental approaches using the same biological material with specific work step sequences for each approach. Analyzing the same biological material allows correlation of results derived from these different approaches.

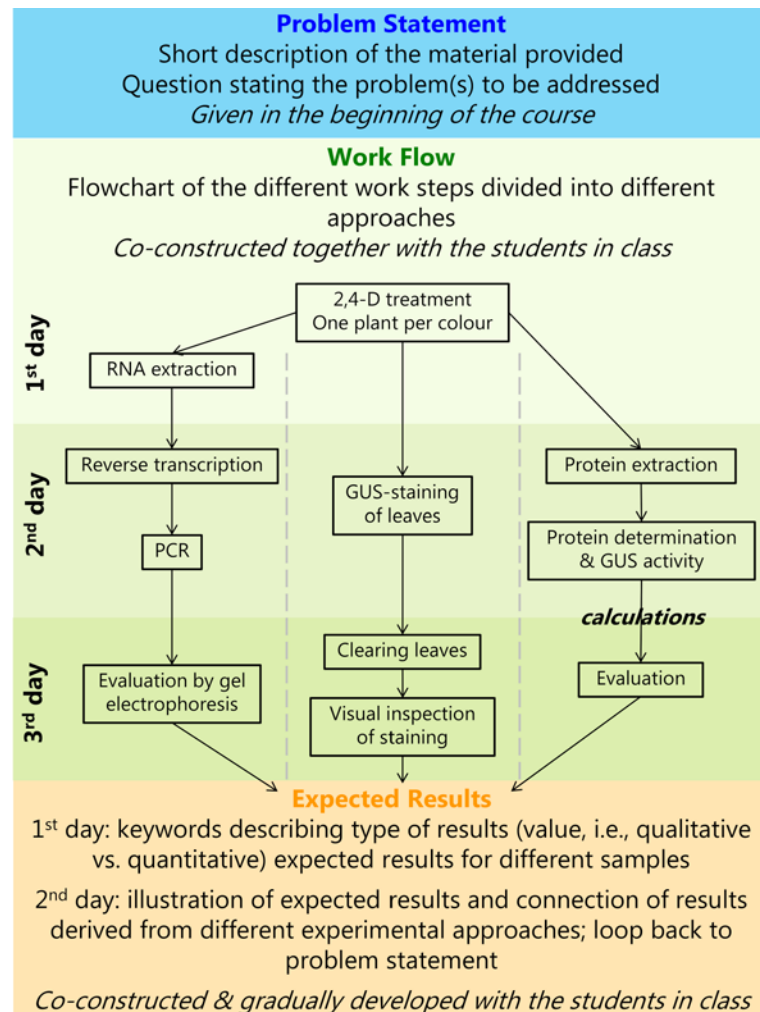
Based on the intended learning outcomes (ILOs) specified for the whole course, the teacher defined specific ILOs for this exercise. The basic aim of the exercise was that the students are able to perform and understand the methods used in the exercise and to connect the individual work steps, i.e. keeping track of experimental work over several days. Furthermore, they should understand the results, including a basic interpretation, and understand the theoretical background of the methods. Ideally, students should also be able to judge the value of the obtained results, correlate the obtained results with each other, and subsequently make an advanced interpretation of the results. Finally, they should be able to apply the newly gained knowledge to decide which methods to use to address scientific problems in a specific way.

### *A co-constructed flowchart as a tool*

In the preparation phase, the teacher and supervisors identified specific time-points in the exercise to be observed to facilitate student understanding and learning, which may be challenging due to the stretched structure of the exercise. To ensure a good learning environment and experience, it was crucial to initially determine the level of preparation of the students to ensure a minimum level of preparation in the beginning and on each individual day of the exercise. In addition, to keep track of the students' progress and to support them in an appropriate and dynamic way throughout the exercise, the practical work had to be managed (in time), and to understand the basic underlying principles in the context of the whole exercise. Facilitating this will subsequently support the students in achieving the aims of the exercise in the context of the whole course.

To support student understanding and learning, the implementation of a flowchart illustrating the whole exercise was intended, co-constructed in dialogue with the students on a whiteboard in class. Thus, the teacher developed a layout of a respective flowchart (Fig. 1) and an initial implementation plan (Table 1) for this exercise. The flowchart consists of three sections and includes the individual work steps of the exercise. The first section is given at the beginning of the exercise on the first day and presents the problem statement, i.e., the description of the source material and the scientific problem to be addressed. The second section describes the distribution of individual tasks over the course of the exercise and the connection of individual experimental approaches. According to the implementation plan, this part is co-constructed on the first day of the exercise and referred to during recapitulation and wrap-up sessions on the following days. The co-construction is done in a plenum discussion with the students about the content of the exercise, addressing the analyzed material, the methods used and the aim of the exercise. Within one week prior to the first day of the exercise, the students received written instructions containing all essential information on these aspects of the exercise, which allowed them to prepare for the exercise. By moderating and guiding the discussion, the teacher drew the flowchart on the whiteboard

step by step according to the students' input. Following the same approach, the third section is also co-constructed with the students on the first and second day of the exercise. It represents the expected results, their value, and connection with each other.



**Figure 1. Outline of the flowchart as a central part of a practical exercise spread over several days.** The flowchart consists of three sections and includes the individual work steps of the exercise, which are spread over several days. The first section is given at the beginning of the exercise and presents the problem statement (blue). The second section describes the distribution of individual tasks over the course of the exercise, exemplified by the work steps of the laboratory exercise “Tracking Gene Expression” in the BSc course “Plant Genomics” (green). In this example, the work is spread over three course days and split into three approaches, leading to results of different values, which are addressed in the third section, which shows the expected results (orange). The second and third sections are constructed with the students during the exercise.

Table 1. Implementation plan for the flowchart as a central element of this laboratory exercise.

Section	Action(s)	Intended effect
<b>Day 1</b>		
Problem Statement	<p>Questions and discussion on source material and scientific problem in plenum.</p> <p>If no answers are given, buzz groups are initiated followed by plenum discussion.</p>	<p>Familiarize students with the exercise basics → avoids losing students due to lack of initial understanding.</p> <p>Teacher can estimate the level of understanding and preparation of the students.</p>
Work Flow	<p>Teacher guides through the different approaches, thereby develops individual work step sequences in dialogue with the students.</p> <p>Includes questions targeting student understanding, i.e., question-answer sequences, buzz groups, and discussions in plenum.</p>	<p>Familiarize unprepared students with the content, recapitulate content for prepared students, and put the different work steps into context.</p> <p>Overview of the upcoming workload.</p>
Expected Results	<p>Questions on expectations regarding results of the individual approaches and their value; can be supported by buzz groups and plenum discussion.</p> <p>Adding keywords to the flowchart based on answers.</p>	<p>Teacher can estimate the students' understanding of the exercise concept and which parts need special attention.</p> <p>Short overview of the results before starting the practical work.</p>
<b>Start of day 2</b>		
Problem Statement	<p>Questions on the starting point of the exercise and scientific problem.</p>	<p>Recapitulation of the exercise framework to set the scene and the focus of the students.</p> <p>Teacher can check the understanding of the concept.</p>
Work Flow	<p>Questions on the different approaches and the current progress, i.e., What has been done so far? What will be done on that day?</p>	<p>Recapitulation of previous work and awareness/ preparation of the work to be performed on this day.</p> <p>Connection of the work steps conducted on different days.</p>

<b>End of day 2</b>		
Expected Results	Refining the expected results through illustrations depicting the specific results that will be obtain, e.g., color staining.	Supports understanding and interpretation of the results.  Preparation of students for the results from analyses they perform(ed) → connect results with practical procedures.  Looping back to scientific problem, i.e., why to expect certain results and what they would mean → connect results with scientific problem.
<b>Start of day 3</b>		
Problem Statement	Question whether there are indications for answering the problem based on primary results from day 2.	Setting the scene; initial connection of preliminary results and the scientific problem.  Teacher can check the understanding of the concept.
Work Flow	See day 2; extended by specifically addressing unclear or problematic steps/aspects retrospectively.	See day 2; extended by overall connecting the previous work with the last work steps and upcoming results.
<b>End of day 3</b>		
Expected Results	Guided cross-check of expected vs. obtained results.  Discussion of the results and their value.	Validation of results. Clarification of unexpected results.  Interpretation of results, connecting the information obtained by individual approaches.
All sections	Final plenum discussion and recapitulation.	Connect scientific problem, methods, and results with the aims of the exercise to facilitate a holistic understanding.  Highlight the need and value of the different approaches, i.e., which to choose to obtain what kind of information to solve a certain scientific problem.  Clarify problematic issues connected to the exercise.

<b>All through the exercise</b>		
All sections	Spontaneous discussions between teacher and students or between students based on the flowchart.  Individual students “consulting” the flowchart.	Reference point for orientation for students during the progress of the exercise and for exchange/discussions.

#### *Evaluation of the effects of implementing the co-constructed flowchart*

One important aspect is the perception of the flowchart by the students and their evaluation of the exercise. Considering the fact that this exercise was conducted the first time with this content, it was the intention to get a very open basic feedback (via anonymous online questionnaire or by email) on the exercise, for example, what the students experienced as (very) positive or what needs to be improved. Therefore, to avoid any bias in the general evaluation of the exercise, no specific evaluation of the flowchart and its implementation was requested.

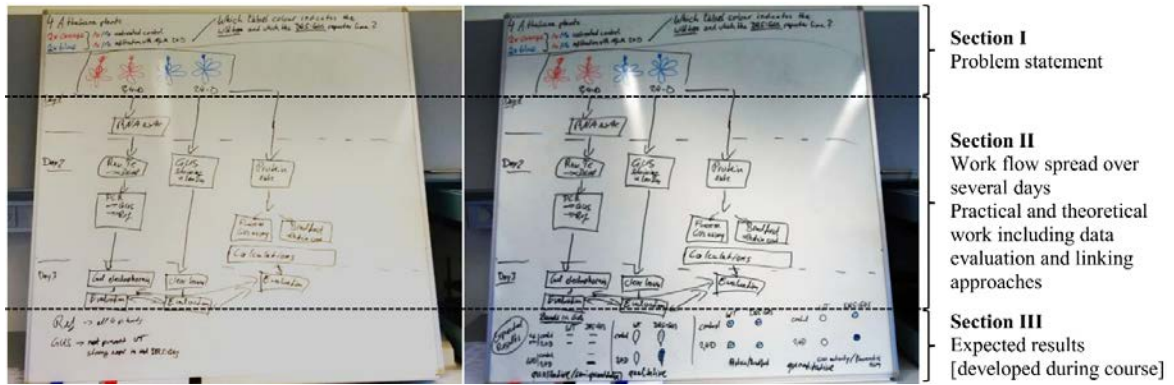
In addition to student feedback, the effect of implementing a co-constructed flowchart as a central element of the exercise was evaluated by the teacher based on experiences from previous teaching of similar exercises. Important aspects of this evaluation comprise the estimation of student understanding of the content based on the final discussion of obtained results and case studies, the quality of final reports submitted by the students, and the teacher’s role and perspective on the exercise compared to similar exercises not using this tool.

### **Results**

#### *The co-constructed flowchart facilitates active communication and orientation*

The concept of co-constructing a flowchart with the students turned out to also be very valuable in initiating dialogue with the students. The exercise started with a discussion on the basic scientific question and source material (Table 1). It was a natural process to allow students to discuss in buzz groups when the students appeared reluctant to answer the questions due to different reasons such as lack of preparation, limited initial understanding of the topic, and possibly initial reservations to speak in plenum. This initial phase of restraint was quickly overcome, and when jointly developing the work step sequences it became a relatively lively and interactive process between the teacher and the students. Thus, the co-constructed flowchart (Fig. 2) was a good initiation of the exercise by facilitating the interaction between the teacher and the students.

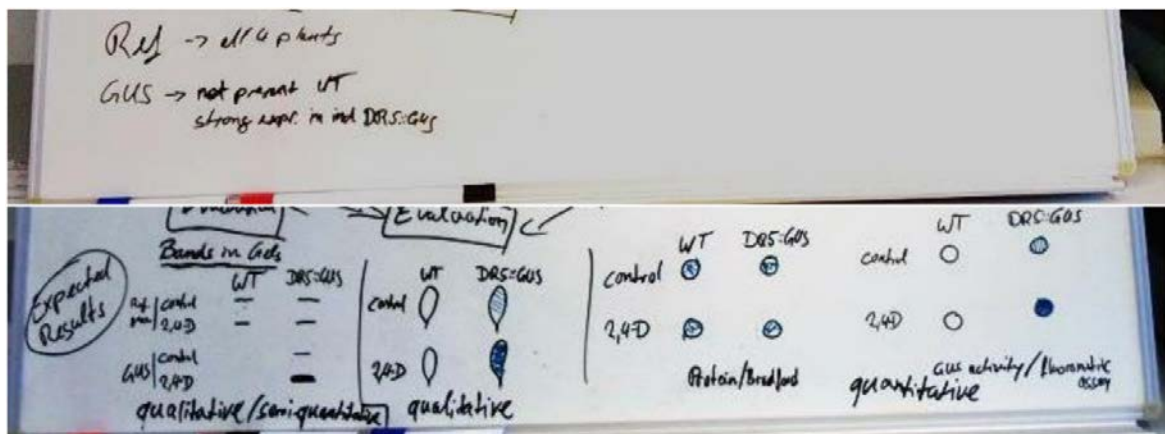




**Figure 2. Flowchart co-constructed with the students in class.** The picture shows the flowchart after day 1 (left panel), and the finalized flowchart after day 2 (right panel) with the three sections corresponding to the scheme depicted in Fig.1.

According to the implementation plan (Table 1), the flowchart served as a reference and communication tool throughout the exercise on the following days. Through recapitulation, discussion and developing the content of the exercise based on the flowchart, the students could easily connect individual work steps and follow their progress within the work sequence over the days of the exercise. As assumed, the overview of the “Expected Results” on the first day was somewhat limited as the students lacked a connection to the work steps, which they had not performed yet. When completing this section on the second day (Fig. 3), it was obvious that priming this part on the first day was very beneficial for the students’ understanding. The students understood the added information very well and could easily connect it to the previous information and the other parts of the flowchart.

In general, the flowchart served as a common basis for communication of the content of the exercise between the teacher and the students, but also between individual students. In addition, it was observed that students actively used the flowchart for their orientation. It also improved the teacher’s confidence as the work with the flowchart resulted in a good overview of the status and progress of the students. It created a feeling of control, but also flexibility in running the exercise. Thus, the flowchart on the whiteboard and its co-construction with the students facilitated a good “scientific” exchange between all participants in the exercise (including the teacher) as well as the understanding of the content and the discussion partner.



**Figure 3. Close-up of section III - “Expected Results” of the flowchart.** The upper panel shows the expected results as short keywords after day 1, the lower panel shows the depiction of the results after co-constructing this section with the students on day 2 after recapitulating the work progress up to that point. A more detailed description of the expected results was possible on day 2 of the exercise after the students were more familiar with the content of the exercise and had time to reflect on it and the expected results.

#### *Success of the flowchart-centered exercise*

A central part of estimating the success of the exercise in terms of the students achieving the learning goals was a closing discussion session. Their results and selected research publications, which are based on the methods performed in the exercise, were discussed in plenum. Overall, the students were very well aware of the work they performed, and the meaning and value of the results, as evidenced by the quality of their active contributions to the discussions and constructive and conscious questions, as well as the involvement in the discussion of the majority of the students. They were able to recapitulate the individual methods, what they are useful for, and how they may complement each other. Furthermore, they could explain how the different results are connected, and how they may be combined to answer scientific questions. When discussing the case studies, the students were able to apply their knowledge of the learned methods to the presented research. While the basic understanding of their own work was much better than experienced in similar courses, it was the first time the teacher could say that the students were able to apply their knowledge to other scientific contexts. In previous courses, this turned out to be a very critical aspect, which is probably connected to the observed lack of basic understanding. Similarly, the high level of understanding was also reflected in the final reports, which on average were of a better quality and showed more depth of understanding of the ILOs than experienced before. Importantly, no group failed to describe the work performed, to present the results and to provide valid interpretation, which regularly happened in similar exercises before. This means that generally all intended learning goals of the exercise had been achieved, which was not always the case in previous courses. Instead, the final discussion and the reports highlighted very specific details the students struggled with, which may have otherwise stayed hidden through the lack of a basic understanding of the exercise content.

#### *Students' appreciation of the flowchart*

During the exercise feedback, some students indicated that they appreciate this approach, having the flowchart developed together step by step on a whiteboard that stays in the laboratory. In addition, the flowchart was often highlighted as a very positive aspect in the feedback provided by the students after completing the exercise. They were initially simply asked for feedback on aspects they perceived as positive and those that could be improved. Eight out of eleven students provided feedback of which seven specifically mentioned the flowchart/whiteboard as the most (or one of the two most) positive aspects of the exercise, while one student referred to it indirectly (Table 2). This positive evaluation after class confirmed the very positive impression during the course of the exercise.

Table 2. Compiled original feedback referring to the used flowchart/whiteboard provided by students after class.

Responses by email	Student 1	Positive: Overview on whiteboard, nice structure of exercise
	Student 2	I really liked the flow chart and that you went through it so thoroughly. It gives a very good overview combined with the protocol.
	Student 3	Positive: - Great with introduction to the experiment with board overview, but don't expect the students to have read for the next many lab days, usually we only read the stuff we are supposed to do for the day.
Anonymous responses via online questionnaire	Student 4	I like the way you used the whiteboard to tell about the different results in the exercise. That made a good brush-up and was along a good learning point.
	Student 5	White board layout, nice structure
	Student 6	Great with board overview before exercise
	Student 7 [commented]	The exercise was well described in the lab by Dominik, so there was no confusion. He was also very good at asking questions, so you had to think about some aspect you might not have thought about.  <i>Comment by D. K. Grosskinsky (teacher): As these points were performed with the use of the flowchart/whiteboard, this feedback can be regarded as a positive evaluation of this tool.</i>
	Student 8	The flow-chart and the run-through created a nice overview of the exercise (very pedagogical)

Students were asked to provide feedback on positive aspects and points of potential improvement of the whole practical exercise in general. All original, unedited feedback related to the flowchart is included in the table.

#### *The flowchart from an external observer's view*

As part of the teacher's pedagogical education, the co-construction and implementation of the flowchart in this exercise were accompanied by two experienced teachers as external observers. Overall, the described flowchart-based approach contributed to a very well-planned teaching session which incorporated varied activities. While discussions and buzz

meetings were centered on the use of the flowchart, the flowchart turned out to be an elegant way to link the theoretical and practical teaching activities to ensure progress towards reaching the planned overall objectives in this part of the course. This has been set up to make a way for the students to internalize the theoretical part of their study program. In addition, the flowchart served as an anchor for the teacher to change his teacher role with effortless ease. The various activities incorporating the flowchart allowed for a real-time formative assessment of the students' progress and direct feedback to current questions, even supporting differentiated teaching based on the students' different starting points and progress during the session.

Thus, the external observations match the positive experience described by the teacher and the students. The co-construction and subsequent repeated use of the flowchart strongly supported the students' learning process and created a very fruitful and safe learning environment, motivating the students to actively engage in their own learning. It was observed that the students also visited the flowchart in between the laboratory exercise as a tool to keep track of the different activities and to discuss in the group. On asking the students, they commented that it had been a constructive way to get an overview of what they should do and had already done to keep track of the learning in this exercise. In summary, the developed flowchart and its implementation, a flowchart to which everyone contributes, strongly supported the students in structuring their laboratory work.

## Discussion

### *The co-constructed flowchart as a supporting tool for student learning*

The implementation of the co-constructed flowchart very much supported the students in achieving the primary aims of the exercise. While the joint construction in itself facilitated the preparation of the students and their active contribution to creating the illustration of the framework, the final flowchart served as a common reference point throughout the course of the exercise. Based on the fact that the flowchart was jointly constructed, the students could also claim ownership, i.e., they could relate to it more easily than if it were just presented by the teacher. Furthermore, this flowchart allowed the different experimental approaches, which were followed in parallel, to be interconnected. The illustration itself may have a beneficial effect, as individuals may grasp the concept more easily from this than just from text. In addition, the integrated repetitive elements in the implementation plan, i.e., that individual points are discussed and recapitulated several times from slightly changing perspectives, appears very beneficial in increasing the understanding of the students and internalizing the obtained knowledge.

Because achieving the primary aims of the exercise was not a problem in contrast to previous courses, this approach revealed the specific problems students are commonly struggling with. In cases where students are struggling with the basic primary aims of such exercises, these specifically problematic aspects may be hidden and not addressed in an appropriate way. These points were very specific, but apparently did not limit the understanding of the basic concepts. In contrast, the overall understanding of the students corresponded to a relational qualitative level of the SOLO-taxonomy of understanding (Mørcke & Rump, 2015), which therefore was on a higher level than in previous courses. Furthermore, the feedback provided by the students clearly demonstrated their appreciation, as it apparently supported their understanding and learning, which is in agreement with the use of student-generated flowcharts in chemistry laboratory exercises (Davidowitz & Rollnick,

2001, 2003). This seems more important than the use of modern high-tech tools, which is in agreement with studies on preferred teaching techniques of natural science and medical students (Novelli & Fernandes, 2007; Waheeda & Murthy, 2015).

#### *The co-constructed flowchart as a supporting tool for the teacher*

The flowchart also provided a good basis for the teacher to facilitate a supportive interaction with the students. It helped in monitoring and guiding the progress of the students, regularly providing feedback, and getting the exercise back on track, e.g., when time was running short, as no big explanations were needed when referring to an already known scheme. The flowchart was also a good tool for the teacher to shift between different roles to diversify the teaching-learning environment (Beck & Gottlieb, 2002). When constructing the flowchart or referring to it as a common basis during daily recapitulation and discussions, the teacher had the chance to move from his role as a coach or supervisor during the practical work to the role of a participant in the discussion or functions as a moderator (Beck & Gottlieb, 2002). This diversified teacher role appeared to be a suitable way to adjust the teaching-learning environment in relation to different needs and it also seemed to be beneficial for the resulting understanding of the students. As the flowchart eased the interaction and communication between teacher and students in general regarding the scientific content of the exercise, it can be a simple tool to spur a kind of research-tutored teaching (Healey & Jenkins, 2005). This allows the teacher to easily refer to the practical work of the exercise in the context of higher scientific concepts or research questions to spur students' ability to transfer the learned methods to different scenarios.

#### **Conclusion**

In this specific laboratory exercise, the use of a co-constructed flowchart appeared to be a valuable approach to guiding the students through the stretched structure of the program, keeping track of their progress, supporting their understanding and learning, and forming the basis for discussions. The implementation of the flowchart had a beneficial impact on different aspects of the teaching-learning environment and can thus be regarded as a success. However, it is important that a teacher defines the purpose of the flowchart for a specific teaching unit and subsequently, how an appropriate flowchart has to look and how it has to be implemented according to a suitable plan, taking the teaching situation into consideration. This means that a specific flowchart could simply be provided as a guide or could be created either through co-construction (like here) or by the students on their own, followed by feedback on their drafts. Certainly, the latter options have the advantage that students have to deal with the content of teaching and develop a feeling of ownership by contributing to the flowchart construction. With these considerations, flowcharts seem to be good tools for guiding students through exercises and content with a complex structure.

#### **Acknowledgement**

Valuable feedback during the preparation of the manuscript by Frederik V. Christiansen and Vibeke Langer (University of Copenhagen) is gratefully acknowledged.

#### **References**

Beck, S., & Gottlieb, B. (2002). Elev/student - en teoretisk og empirisk undersøgelse af begrebet studiekompetence. *Gymnasiepædagogik*, Nr. 31, Institut for

Kulturvidenskaber, Syddansk Universitet.

- Davidowitz, B., & Rollnick, M. (2001). Effectiveness of flow diagrams as a strategy for learning in laboratories. *The Australian Journal of Education in Chemistry* 57 (1), 18-24.
- Davidowitz, B., & Rollnick, M. (2003). Enabling metacognition in the laboratory: a case study of four second year university chemistry students. *Research in Science Education* 33 (1), 43-69.
- Grosskinsky, D. K. (2018). Tracking and supporting student learning in practical laboratory exercises spread over several days. In Christiansen F. V., & Prestholm, I. (Eds.), *Improving University Science Teaching and Learning - Pedagogical Projects 2018*, pp 47-61. Copenhagen: Department of Science Education, University of Copenhagen.
- Healey, M., & Jenkins, A. (2005). *Developing undergraduate research and inquiry*. York: The Higher Education Academy
- Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: foundations for the twenty-first century. *Science Education* 88 (1), 28-54.
- Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: the state of the art. *Chemistry Education Research and Practice* 8 (2), 105-107.
- Mørcke, A. M., & Rump, C. Ø. (2015). University teaching and learning – models and concepts. In Rienecker, L., Jørgensen, P. S., Dolin, J., & Ingerslev, G. H. (Eds.), *University Teaching and Learning*, pp. 93-104. Frederiksberg: Samfundslitteratur.
- Novelli, E. L. B., & Fernandes, A. A. H. (2007). Students' preferred teaching techniques for biochemistry in biomedicine and medicine courses. *Biochemistry and Molecular Biology Education* 39 (4), 263-266.
- Reid, N., & Shah, I. (2007). The role of laboratory work in university chemistry. *Chemistry Education Research and Practice* 8 (2), 172-185.
- Tanner, K., & Allen, D. (2004). Approaches to biology teaching and learning: learning styles and the problem of instructional selection – engaging all students in science courses. *Cell Biology Education* 3 (4), 197-201.
- Waheeda, S., & Murthy, K. S. (2015). A comparative study of blackboard teaching with PowerPoint teaching in 1 year medical students. *National Journal of Basic Medical Sciences* 6 (1), 11-13.

# Matematiktest som prædiktor af ingeniørstuderendes studiesucces

Henrik Skov Midtby<sup>a,1</sup>, Nadia Dyrberg Egemose<sup>b</sup>

<sup>a</sup>Mærsk Mc-Kinney Møller Institut, Syddansk Universitet, <sup>b</sup>SDU Universitetspædagogik, Syddansk Universitet

Faglig artikel, fagfællebedømt

Mangel på ingeniører og andre STEM-uddannede vækker bekymring i Europa. Og Danmark er blandt de lande, som er særligt udfordrede. Internationale studier tyder på, at matematiske evner har prædiktiv værdi i forhold til STEM-studerendes studiesucces – et begreb der omfatter både præstationer og progression på uddannelsen. I dette studie undersøges det, hvorvidt to matematiktests kan forudsige danske ingeniørstuderendes studieprogression; studiegennemførelse og beståelse af førsteårsprøven. De studerende havde ikke forberedt sig på test 1, mens de havde mulighed for at forberede sig på test 2. I forhold til den første test tyder studiets fund på, at det i højere grad er den frivillige deltagelse i undervisningen, hvori testen indgik end selve testscoren, der har sammenhæng med studiesuccesen. Pointscoren i den anden test viste sig at være en middelstærk prædiktor af både frafald og beståelse af førsteårsprøven. Studiet bidrager med input til diskussionen om optagelsestests på danske videregående uddannelser samt til ønsket om tidligt at kunne identificere frafaldstruede studerende.

## Introduktion

Forventet mangel på ingeniører og andre STEM-dimittender (science, technology, engineering and mathematics) har vakt bekymring i Europa siden årtusindskiftet (Business Europe, 2011; European Commission, 2004; Harrison, 2012). Alene i Danmark forventer Ministeriet for Børn, Uddannelse og Ligestilling en mangel på 500 kandidater i 2020. Dette gælder særligt inden for studieretningerne fysik, matematik og datalogi (UddannelsesGuiden, 2016). Ingeniørforeningen IDA har på vegne af Engineer the Future udarbejdet en fremskrivning, der forudsiger en endnu større mangel: Her viser analysen, at *der på trods af stigningen i optaget de seneste år samlet vil mangle 10.000 ingeniører og naturvidenskabelige kandidater i 2025. Heraf vil der mangle 6.500 civil- og diplomingeniører, mens der vil mangle 3.500 naturvidenskabelige kandidater*(Engineer the Future, 2018, s. 1).

Principielt kan antallet af STEM-uddannede øges på to måder: I) øget rekruttering til STEM-uddannelserne og II) øget fastholdelse og gennemførelse af STEM-studerende. Med øget rekruttering som mål har regeringen i begyndelsen af 2018 afsat 180 mio. kr. i forbindelse med en ny naturvidenskabsstrategi som i årene 2018-2024 skal udmønte sig i konkrete initiativer, der skal skabe større interesse for naturvidenskabelige uddannelser blandt børn og unge

---

<sup>1</sup> Kontakt: hemi@mmpi.sdu.dk

(Undervisningsministeriet, 2018). Genstandsområdet for dette studie er dog studerende, som allerede er rekrutteret og således allerede er blevet optaget på en STEM-uddannelse.

Optaget af STEM-studerende på danske universiteter er samlet set stigende. Mens det naturvidenskabelige bacheloroptyag har ligget på et stabilt niveau på omkring 5.000 optagne studerende de seneste fem år, så er optaget på civil- og diplomingeniøruddannelserne i samme periode steget stødt fra 4.620 studerende i 2013 til 5.929 studerende i 2017 (Uddannelses- og Forskningsministeriet, 2017a). Lave gennemførelsesprocenter udfordrer dog endnu det natur- og tekniskvidenskabelige hovedområde. I 2016 var førsteårsfrafaldet hhv. 20 % og 16 % på de naturfaglige og tekniske bacheloruddannelser (Uddannelses- og Forskningsministeriet, 2017b), og den samlede gennemførelse fire år efter påbegyndelse af en bacheloruddannelse var i 2017 mellem 51 % og 56 % på tværs af de danske universiteter. Eneste undtagelse herfor var IT-Universitetet (ITU) med en gennemførelsesprocent på 67 % (Danske Universiteter, 2018).

I en analyse af en række landes prioriteringer angående studiesucces, bl.a. forstået som studie gennemførelse, kategoriseres Danmark som et af de lande, der sætter studerendes succes særligt højt på den politiske dagsorden (Vossensteyn et al., 2015). Landets universiteter støtter op om denne prioritering og har igangsat både større og mindre frafaldsbegrænsende tiltag. Af større projekter kan bl.a. Københavns Universitet nævnes; her har der siden studieåret 2011/2012 været fokus på datalogistuderende i *Projekt fastholdelse*<sup>2</sup>. På Syddansk Universitet er projekterne *Datadrevet frafaldsminimering og øget studiekvalitet på humaniora*<sup>3</sup> og *STAY*<sup>4</sup> på hhv. Det Humanistiske og Det Naturvidenskabelige Fakultet igangsat. Der er således stor interesse for at implementere frafaldsbegrænsende tiltag for igangværende studerende samt at identificere frafaldstruede studerende tidligt i studiet eller forud for studiestart. I denne artikel præsenteres et studie med sidstnævnte fokus: identifikation af frafaldstruede ingeniørstuderende tidligt i studiet ved brug af en matematiktest. Den anlagte forståelse af studiesucces er her omfattet af studieprogression. Nærmere bestemt benyttes studiestatusserne 'afsluttet', 'aktiv' eller 'afbrudt (studieforløb)' og beståelse af førsteårsprøven som er en række studieelementer, der skal gennemføres inden for studiets første år. De studerendes præstationer i form af karakterer på studiet indgår således ikke i dette studies brug af begrebet studiesucces.

Årsagerne til frafald på STEM-uddannelserne er både mange og komplekse, og de udfordringer, som frafaldne studerende oplever, genkendes i høj grad også af ikke-frafaldne studerende (Holmegaard, Madsen og Ulriksen, 2014; Seymour og Hewitt, 1997). Af Tintos 30 år gamle frafaldsmodel (1987), der endnu hyppigt refereres til (Bonne og Nutt, 2016), fremgår det, at både den sociale og den akademiske (eller faglige) integration er essentielle elementer i studiemotivationen og dermed i beslutningen om enten at fortsætte eller forlade sit studie. Trods et omfattende og veludviklet teoriapparat og adskillige frafaldsstudier er igangværende studerendes motivation for studiet vanskelig at måle. Ofte er der tale om affektive faktorer, der evalueres gennem selvvurderingsspørgeskemaer, hvilket er forbundet med en række usikkerheder. Kognitive faktorer er imidlertid lettere at kvantificere, og en række studier har derfor fokuseret på studerendes kognitive evner til at forudsige studiesucces; bl.a. konklude-

<sup>2</sup> <http://www.science.ku.dk/projektfastholdelse/>

<sup>3</sup> [https://www.sdu.dk/da/om\\_sdu/institutter\\_centre/c\\_cifu/projekter/datadrevet+kvalitetsudvikling+og+frafaldsminimering+paa+humaniora](https://www.sdu.dk/da/om_sdu/institutter_centre/c_cifu/projekter/datadrevet+kvalitetsudvikling+og+frafaldsminimering+paa+humaniora)

<sup>4</sup> [https://www.sdu.dk/da/om\\_sdu/fakulteterne/naturvidenskab/nyhedsbrev/2017\\_03\\_30\\_stay](https://www.sdu.dk/da/om_sdu/fakulteterne/naturvidenskab/nyhedsbrev/2017_03_30_stay)



rede McKenzie og Schweitzer (2010) i et australsk studie, at blandt en række faktorer herunder psykologiske, kognitive og demografiske var tidligere faglige præstationer den bedste enkeltstående prædikator af førsteårsstuderendes præstationer på universitetet. Integration på studiet, self-efficacy og forpligtigelser fra studiejob havde dog også en vis forklarende eller forudsigende værdi. Ved integration på studiet henvises til Tintos begreber om akademisk og social integration (se Tinto, 1987 for en detaljeret gennemgang heraf). Trods nuancer i definitionen af begrebet self-efficacy på tværs af studier trækker dette begreb på Banduras arbejde (1986). I det omtalte australske studie forstås self-efficacy som tro på og forventning om akademisk succes (McKenzie og Schweitzer, 2010).

Inden for STEM-uddannelserne tyder studier på, at især tidligere matematikpræstationer er en prædikator for studiesucces i form af både præstationer og gennemførelse (Ackerman et al., 2013; De Winter & Dodou, 2011; Veenstra et al., 2008). Et amerikansk studie med fokus på ingeniørstuderende fandt, at køn og tidligere præstationer (high school-præstation og SAT-score) kunne forklare 18 % af variansen af de ingeniørstuderendes præstationer på universitetet (French et al., 2005). Motivation havde ingen yderligere signifikant forklaringsværdi på de studerendes præstationer eller deres fastholdelse på universitetet, men det havde indflydelse på, hvorvidt de studerende fortsatte med et hovedfag (major) inden for ingeniørfagene. I samme stil benyttede Moses et al. (2011) en matematiktest, Calculus Readiness Exam, som viste sig at være en endnu bedre prædikator end de studerendes tidligere og adgangsgivende præstationer, og i Flandern afprøves en frivillig matematiktest, *Engineering Positioning Test*, der tages af studerende forud for et studievalg inden for ingeniørfagene (Vanderroot et al., 2014). De foreløbige resultater af sidstnævnte studie indikerer, at testen kan forudsige de studerendes akademiske præstation i den første eksamensperiode, men der foreligger endnu ikke resultater omkring senere eksamener eller de studerendes gennemførelsesstatus.

Ingeniørstuderendes matematiske færdigheder og kompetencer er ikke kun i fokus i forhold til overordnet studiesucces. Ofte fremhæves utilstrækkelige færdigheder og kompetencer inden for matematik som en udfordring af undervisere, når studerende påbegynder et STEM-studie. I England har problematikken, kendt som 'the mathematics problem', ført til oprettelsen af adskillige støttecentre i matematik på de videregående uddannelsesinstitutioner. Og flere nationale rapporter har peget på netop matematiske evner som en nøglefaktor for STEM-studerendes progression og gennemførelse (Lawson, 2012).

Der er således international bevågenhed omkring STEM-studerendes matematiske evner samt evidens for, at matematiske evner til en vis grad kan forudsige, hvordan studerende vil klare sig på STEM-uddannelser. Dette studie har til mål at undersøge, hvorvidt sådanne fund også er gældende for danske ingeniørstuderende. Imidlertid er den generelle tendens i fralfaldsforskningen dog gået fra at betragte studerendes succes som et produkt af den enkelte studerendes evner til i højere grad at blive betragtet som et resultat af de studerendes møde med uddannelsen (Tinto, 2016). Derfor inkluderes der i dette studie også adfældsvariablen 'fremmøde ved første undervisningsgang' som en indikator for motivation. Fremmødet og dermed (frivillig) deltagelse i den første matematiktest, som foregik i forbindelse med undervisningen, betragtes således som en motiveret adfærd i forhold til studiet.

### Studiets fokus

Studiet undersøger sammenhængen mellem danske ingeniørstuderendes succes i form af gennemførelse på studiet og beståelse af førsteårsprøven og deres præstation i to matematiktests. Ydermere sammenholdes undervisningsdeltagelsen med præstationen i de to tests.

Den første matematiktest har tidligere været benyttet til evaluering af læringseffekten af det digitale læringsmiddel *Khan Academy* i et præ-/postdesign (Midtby og Ahrenkiel, 2015). Testen viste sig her nyttig til at synliggøre de studerendes læring og kunne benyttes som et led i underviserens kursusplanlægning. I dette studie undersøges det, hvorvidt testen har prædiktiv værdi i forhold til studerendes overordnede studiesucces. Studiesucces kan operationaliseres på forskellig vis (Vossensteyn et al., 2015), men som beskrevet tidligere er nedenstående indikatorer for studiesucces valgt i dette studie (tabel 1).

*Tabel 1. Studiets afhængige variable: valgte indikatorer for studiesucces.*

Indikator	Mulige udfald
Studiestatus	Afsluttet/aktiv eller afbrudt
Førsteårsprøven	Bestået eller ikke bestået

Grundet tendensen til ikke udelukkende at betragte studerendes succes som personafhængig (dvs. afhængig af personlige karakteristika og forudgående oplevelser/læring) men også kontekstafhængig (hvad den studerende oplever på studiet) undersøges det, hvorvidt tre forskellige variable kan forudsige studiesucces: to kognitive variable, hvoraf den ene er et udtryk for forudgående matematiske evner, mens den anden er modereret af en måneds studieaktivitet samt én adfærdsvariabel. De kognitive variable udgøres af I) forudgående matematikkompetencer vurderet gennem en test udleveret den første undervisningsgang (kaldet test 1) og II) matematikkompetencer efter en måneds studium, vurderet gennem en tilsvarende test, der udgjorde kursets første tællende aktivitet (kaldet test 2). Adfærdsvariablen udgøres af fremmøde ved første undervisningsgang og dermed deltagelse i test 1. Til sammenligning undersøges også den prædiktive værdi af de studerendes adgangsgivende karaktersnit, dvs. de studerendes gymnasiale karaktergennemsnit.

Opsummerende udgøres studiets uafhængige variable således af gymnasialt karaktersnit, pointscoren i 'test 1' og 'test 2' samt 'deltagelse i test 1', mens de afhængige variable udgøres af 'status på studiet' og 'førsteårsprøven'. En førsteårsprøve består af en række prøver, som bachelorstuderende skal bestå inden udgangen af første eller andet studieår, afhængig af det enkelte universitets udmøntning af bekendtgørelsen. De konkrete kurselementer forbundet med førsteårsprøven varierer mellem studieretningerne. Studerende, der ikke har bestået førsteårsprøven inden for den fastsatte tid, kan med henvisning til særlige forhold søge om udsættelse for at fortsætte på studiet.

## Metode

### *Kontekst og udformning af test 1 og test 2*

Matematiktestene test 1 og test 2 blev gennemført som en del af to førsteårskurser i matematik (her kaldet SDS og KBM) i efteråret 2014. De studerende på yderligere to førsteårs matematikkurser (her kaldet GSR og ITI) deltog ligeledes i test 1. Test 1 blev gennemført af de studerende i første undervisningsuge, og test 2, der var SDS- og KBM-kursernes første tællende aktivitet, blev gennemført efter en måneds undervisning. På alle fire hold blev test 1 gennemført som følge af underviserens ønske om indsigt i de studerendes matematiske færdigheder ved kursusstart. På SDS og KBM benyttedes testen også som forberedelse til

kursernes tællende aktiviteter. En grundlæggende forskel på de to tests er, at de studerende ikke havde mulighed for at forberede sig til test 1, mens de havde mulighed for at forberede sig forud for test 2. Hvorvidt denne mulighed blev udnyttet, var op til de studerende og behandles ikke videre her, men indgår i Midtby og Ahrenkiel (2015). En anden forskel på test 1 og test 2 er, at pointgivning blev foretaget af hhv. medstuderende og underviseren. Den maksimale pointscore i de to tests var 36 point for test 1 og 40 point for test 2. I Midtby og Ahrenkiel (2015) udgjorde test 1 og test 2 hhv. præ- og posttest til måling af læring efter en måneds forløb med det digitale læringsmiddel Khan Academy med henblik på at teste effekten af læremidlet.

Testene bestod af korte matematikopgaver på elementært niveau (heraf nogle på grundskoleniveau), som studerende erfaringsmæssigt har udfordringer med. Indhold, form og distribution af test 1 er beskrevet i detaljer i Midtby og Ahrenkiel (2015). Af opgavetyper kan nævnes: 'gange og dividere med negative tal', 'ligninger, der kan løses i to trin', 'gange parenteser sammen' og 'division af brøker'. De studerende forventes som udgangspunkt at kunne besvare opgaverne, da matematik på A-niveau er et adgangskrav på uddannelserne, og det derfor som beskrevet af Midtby og Ahrenkiel må *anses for rimeligt at forvente at de studerende kan løse disse opgavetyper såfremt de som minimum har matematik på B-niveau på gymnasiet* (2015, s. 33).

#### *Studiets deltagere*

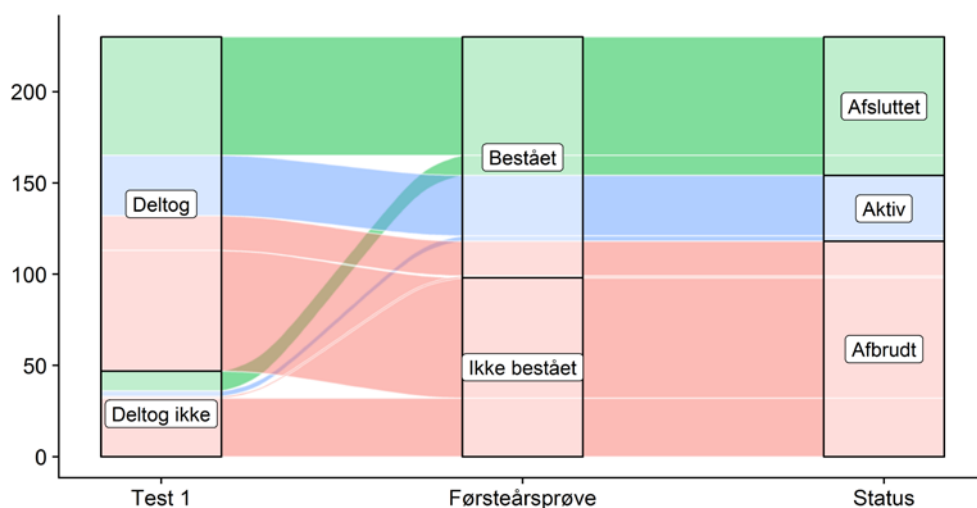
De studerende var indskrevet på syv forskellige ingeniørretninger på Syddansk Universitet (deres kursus angives i parentes): diplomingeniør i kemi- og bioteknologi (KBM), civilingeniør i kemi- og bioteknologi (KBM), diplomingeniør i stærkstrømsteknologi (SDS), diplomingeniør i elektronik og datateknik (SDS), civilingeniør i fysik og teknologi (SDS), civilingeniør i robotteknologi (GSR) og civilingeniør i energiteknologi (ITI). KBM- og SDS-holdene bestod således af kommende diplom- og civilingeniører, mens GSR- og ITI-holdene udelukkende udgjordes af civilingeniørstuderende.

Studerende med studiestart i efteråret 2014 er valgt, idet en stor del af disse studerende, såfremt de har fulgt det anbefalede studieforløb, opnåede deres bachelorgrad i sommeren 2017. Undtagelsen herfor er de diplomingeniørstuderende, som har en normeret studietid på 3,5 år. Disse studerende bør afslutte deres studier i vinteren 2017/2018. Der er således en unik mulighed for netop nu (opgjort i efteråret 2017) at undersøge en årgang, hvor en gruppe studerende vil have gennemført deres bacheloruddannelse eller endnu være aktive studerende, mens en anden gruppe vil være droppet ud af studiet. De KBM-studerendes adgangsgivende karakterer er brugt til sammenligning mellem årgang 2014 og 2015 for at sikre, at den undersøgte årgang ikke skiller sig signifikant ud fra andre årgange.

Tabel 2. Studerende inkluderet i dette studie fordelt på studieretning. Endvidere angives antallet af studerende, der hhv. deltog i test 1, deltog i test 2, har bestået førsteårsprøven og har afbrudt deres studie. \*angiver samlet deltagelsesprocent blandt KBM- og SDS-studerende.

Studieretning (antal)	Deltagelse i test 1		Deltagelse i test 2		Bestået førsteårsprøve		Afbrudt studiet	
	n	%	n	%	n	%	n	%
GSR (n=57)	45	79 %	-	-	28	62 %	30	53 %
ITI (n=53)	40	75 %	-	-	43	81 %	15	28 %
KBM (n=60)	53	88 %	50	83 %	32	53 %	37	62 %
SDS (n=60)	45	75 %	47	78 %	29	48 %	36	60 %
I alt (n=230)	183	80 %	97	81 %*	132	72 %	118	51 %

Data angående de studerendes studieprogression er opnået fra Syddansk Universitets afdeling for 'Analyse & Kvalitet'. Tabel 2 giver et overblik over antallet af studerende, der er studiedata på, og hvor mange af disse der deltog i hhv. test 1 og test 2, mens figur 1 giver et visuelt overblik over de studerendes studieforløb. Den grønne farve repræsenterer studerende, som har afsluttet deres studier, blå repræsenterer endnu aktive studerende, og rød repræsenterer studerende, som har afbrudt deres studier. Det bemærkes bl.a., at mens ca. halvdelen af de afbrudte studerende deltog i kurssets første undervisningsgang og dermed i test 1, så er det kun få af disse studerende, der har bestået førsteårsprøven. Dette indikerer, at størstedelen af de afbrudte studerende enten selv har valgt at droppe ud i løbet af første studieår eller har haft faglige vanskeligheder, der har umuliggjort beståelse af førsteårsprøven.



Figur 1. Visuelt overblik over studerendes studieforløb i forhold til deltagelse i test 1, beståelsesstatus på førsteårsprøven og studiestatus.

### Statistiske analyser

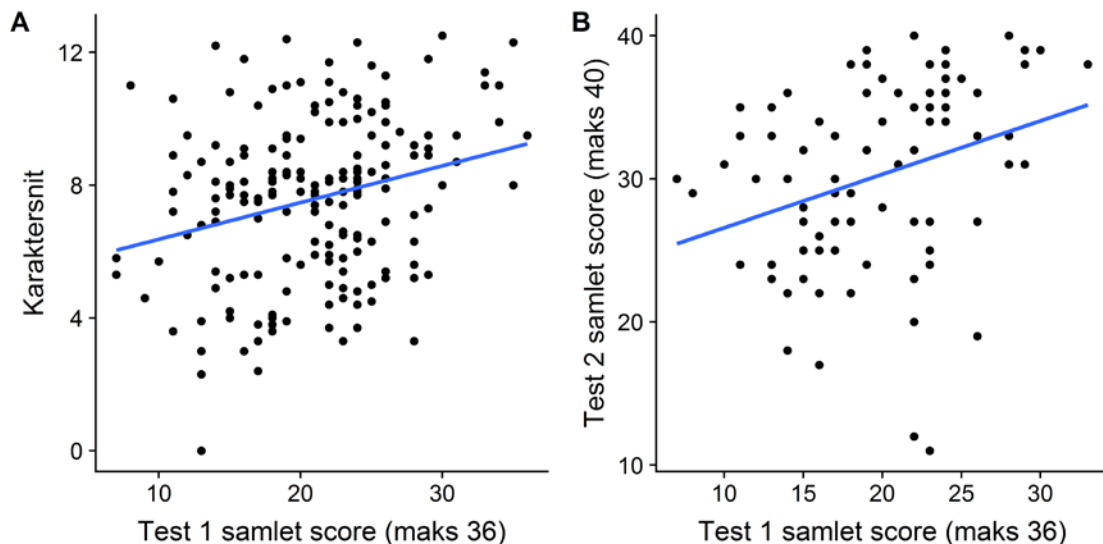
Data blev analyseret i statistikprogrammet R. Følgende statistiske tests blev benyttet: tosidet uafhængig  $t$ -test og ANOVA til sammenligning af middelværdier mellem grupper af studerende, beregning af korrelationskoefficient til test af sammenhænge mellem kontinuerte variable og  $\chi^2$ -test til test af fordeling af grupper på nominale variable. Da statistisk signifikans ikke indikerer, hvor betydningsfuld en påvist forskel eller sammenhæng er, benyttes Cohen's  $d$  som mål for effektstørrelse (Field, 2013).

### Resultater

Med en gennemsnitlig adgangsgivende karakter for studerende fra årgang 2014 og 2015 på hhv. 7,63 (SD=2,213) og 7,52 (SD=2,368) er der ingen statistisk signifikant forskel på de to årganges faglige præstation forud for optagelse på ingeniørstudiet ( $t=0,249$ ,  $p=0,804$ ). Data omkring årgang 2015 er udelukkende benyttet til denne sammenligning, og de følgende statistiske analyser foretages alene for studerende på årgang 2014 - studerende med studiestart i september 2014. Årgang 2015 er ikke medtaget i analyserne, da disse studerende ikke kan forventes af have gennemført studiet på tidspunktet for dataudtrækket.

#### Adgangsgivende karakterer og testscorer

Der er en svag positiv lineær sammenhæng mellem de studerendes adgangsgivende karakterer og scoren i test 1 ( $r=0,069$ ,  $p<0,001$ ), men der kunne ikke påvises en signifikant sammenhæng mellem adgangsgivende karakterer og test 2-scoren ( $r=0,036$ ,  $p=0,085$ ). Der er en svag positiv lineær sammenhæng mellem pointscoren i test 1 og test 2 ( $r=0,191$ ,  $p<0,001$ ).

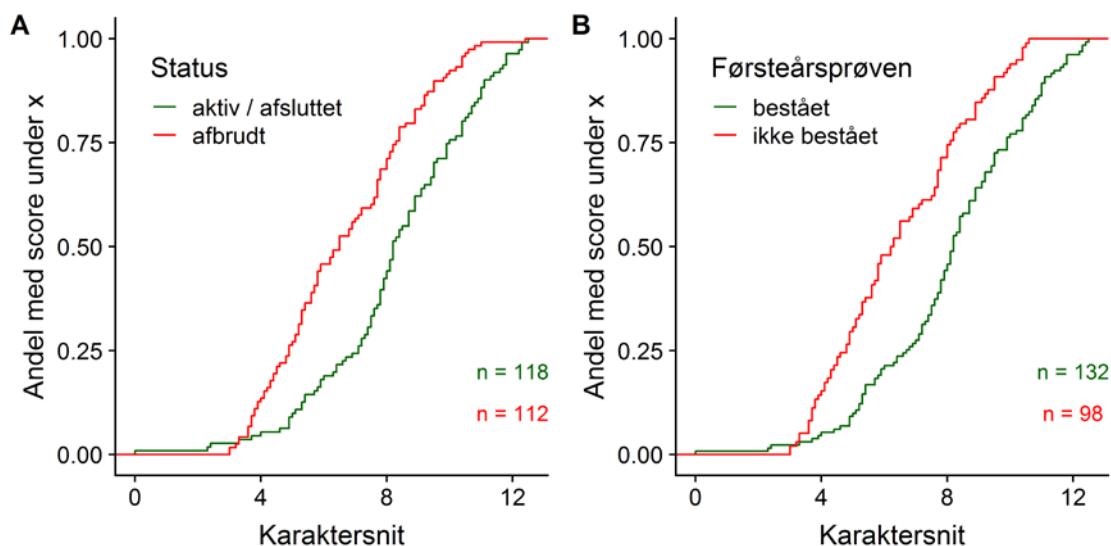


Figur 2 A. Scatterplot med test 1-score på X-aksen og adgangsgivende karaktersnit på Y-aksen ( $r=0,069$ ). B: Scatterplot med test 1-score på X-aksen og test 2-score på Y-aksen ( $r=0,191$ ).

Figur 2 illustrerer forholdet mellem de studerendes adgangsgivende karaktersnit og scoren i test 1 samt mellem scoren i test 1 og test 2. Der er ikke påvist en effekt af typen af adgangsgivende eksamen på testscoren af hverken test 1 ( $F=1,975$ ,  $p=0,085$ ) eller test 2 ( $F=0,364$ ,  $p=0,833$ ).

### Studiesucces og adgangsgivende karakterer

De studerendes adgangsgivende karaktergennemsnit har prædiktiv værdi i forhold til både studiestatus ( $p < 0,001$ ) og beståelse af førsteårsprøven ( $p < 0,001$ ). Effektstørrelserne på hhv.  $d = 0,695$  og  $d = 0,719$  svarer til en moderat til stor effekt. Figur 3 illustrerer sammenhængene mellem karaktersnit og studiesucces.



Figur 3. Gymnasiale karaktergennemsnit for studerende, som A) har gennemført deres uddannelse eller stadig er aktive [aktiv/afsluttet], eller som er droppet ud [afbrudt] og B) studerende som har bestået [bestået] eller dumpet deres førsteårsprøve [ikke bestået].

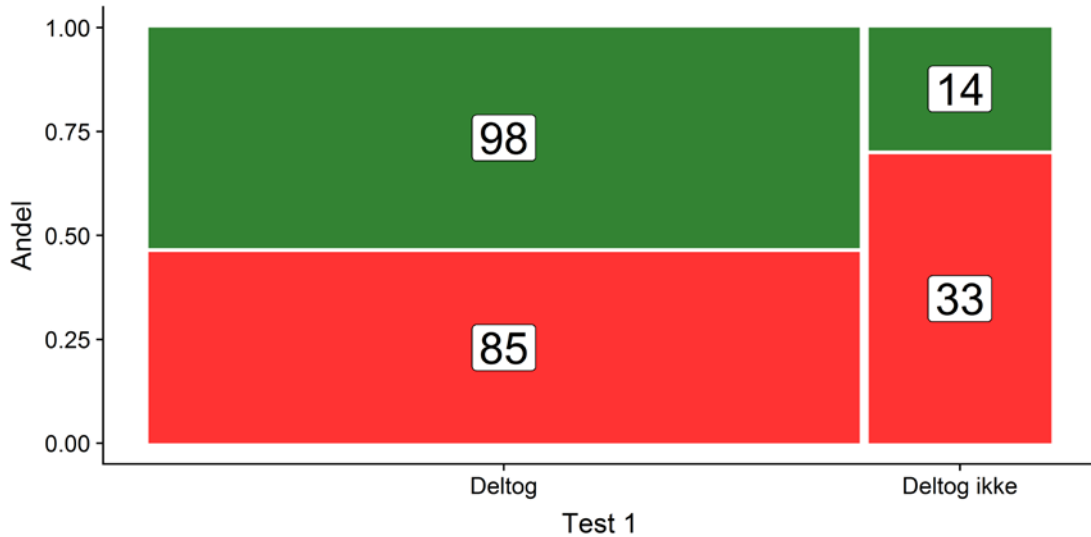
I figuren angives karaktersnittet på X-aksen, mens andelen af studerende, der har opnået dette karaktersnit under en given X-værdi aflæses på Y-aksen. En graf, der ligger længere mod højre, illustrerer således højere gymnasiale karakterer for den repræsenterede gruppe.

### Studiestatus, testdeltagelse og testscore

Studerende, der deltog i test 1 ( $n = 183$ ), har generelt klaret sig bedre i deres studie end de studerende, som ikke deltog i testen ( $n = 47$ ) (tabel 3). Der er således signifikant forskel på andelen af deltagende og ikke-deltagende studerende, der har gennemført deres uddannelse/endnu er aktive studerende eller er droppet ud ( $\chi^2 = 7,529$ ,  $p = 0,006$ ). Figur 4 visualiserer denne sammenhæng.

Tabel 3. Studiestatus på studerende som hhv. deltog og ikke deltog i test 1.

Studiestatus	Deltagende	Ikke-deltagende
	( $n = 183$ )	( $n = 47$ )
Afsluttet/Aktiv	53,6 %	29,8 %
Afbrudt	46,4 %	70,2 %



Figur 4. Studiestatus på studerende som hhv. deltog og ikke deltog i test 1. Grøn repræsenterer studerende, som har afsluttet deres studie eller endnu er aktive, mens rød repræsenterer studerende, som har afbrudt deres studie.

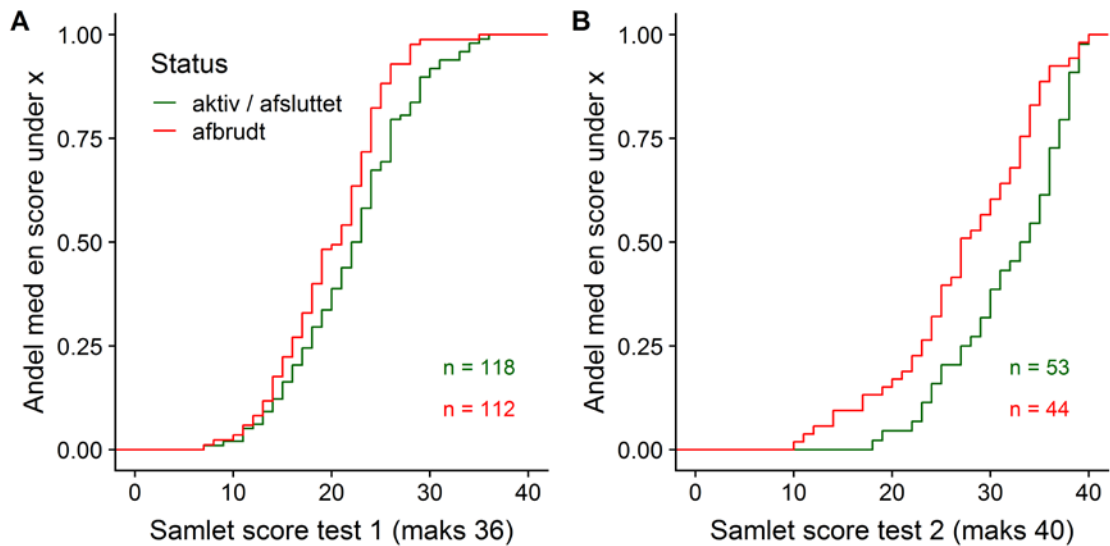
De følgende analyser er foretaget på studerende, som deltog i de undersøgte tests.

### Test 1

Studerende, der har afsluttet deres bacheloruddannelse eller endnu er aktive studerende, opnåede højere test 1-scorer ( $M=22,03$ ,  $SD=6,18$ ) end studerende, der har afbrudt deres uddannelse ( $M=19,91$ ,  $SD=5,28$ ) ( $t=2,507$ ,  $p=0,013$ ). Denne forskel svarer til en lille til middel effekt ( $d=0,368$ ).

### Test 2

For test 2 gælder det ligeledes, at studerende, der har afsluttet deres studier eller stadig er aktive, opnåede højere test-scorer ( $M=31,91$ ,  $SD=5,97$ ) end studerende, der har afbrudt deres studier ( $M=27,38$ ,  $SD=7,54$ ). Her var p-værdien=0,001 ( $t=3,302$ ) og effektstørrelsen  $d=0,659$ , hvilket indikerer til en middel til stor effekt.

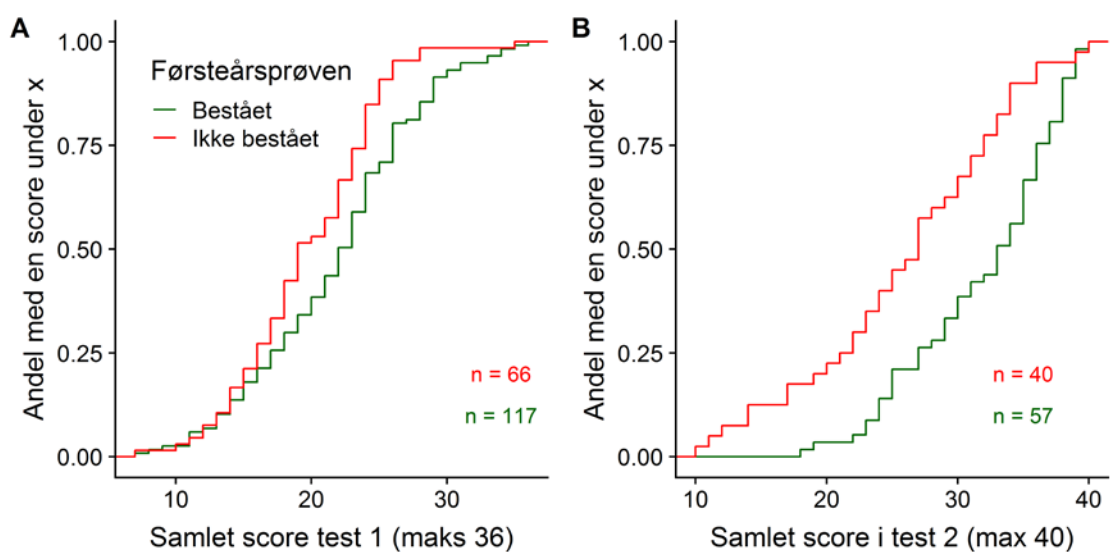


Figur 5. Score i test 1 (A) og test 2 (B) fordelt på studerende, som har gennemført deres uddannelse eller stadig er aktive [aktiv/afsluttet], eller som er droppet ud [afbrudt].

Figur 5 illustrerer de opnåede testscorer i test 1 og test 2 for de to grupper studerende: 'aktiv/afsluttet' og 'afbrudt'. I figuren angives matematiktestscoren på X-aksen, mens andelen af studerende, der har opnået en testscore under en given X-værdi aflæses på Y-aksen. En graf, der ligger længere mod højre, illustrerer således højere testscorer for den repræsenterede gruppe.

#### Førsteårsprøven og testscorer

Studerende, der har bestået førsteårsprøven, opnåede højere scorer i test 1 end studerende, der ikke har bestået førsteårsprøven ( $t=2,496$ ,  $p=0,014$ ). Denne forskel er illustreret i figur 6. Forskellen i den gennemsnitlige test 1-score repræsenterer dog kun en lille til middel effekt ( $d=0,364$ ). Samme tendens med højere pointscorer for beståede studerende ses også for test 2 ( $t=4,092$ ,  $p>0,001$ ), hvor effektstørrelsen ( $d=0,888$ ) indikerer en stor effekt.



Figur 6. Score i test 1 (A) og test 2 (B) fordelt på studerende, der har bestået eller ikke bestået førsteårsprøven.



## Diskussion

Et stort behov for flere STEM-dimitterende i fremtiden sætter fokus på både rekruttering og fastholdelse af studerende til de natur- og ingeniørvidenskabelige uddannelser. En del af dette fokus består i et ønske om at kunne forudsige hvilke studerende, der har øget risiko for at droppe ud af deres påbegyndte studier. En sådan identifikation af igangværende studerende vil muliggøre tidlige frafaldsbegrænsende interventioner, mens identifikation af risikofaktorer forud for studiestart vil kunne give anledning til en ekstra refleksion om, hvorvidt ansøgning til netop dette studie er det rigtige for den kommende studerende eller give input til optagelsesproceduren fra uddannelsesinstitutionens side. Internationale studier peger på bl.a. matematikkompetencer som afgørende for studerendes succes på STEM-studier, og studiet her undersøger, hvorvidt dette også gælder for danske ingeniørstuderendes studie-succes i form af deres progression; gennemførelse og beståelse af førsteårsprøven. I nærværende studie er studerendes karaktermæssige præstationer ikke medtaget i forståelsen af og fokus på studiesucces, selvom dette element bestemt er relevant i en generel diskussion af studiesucces.

Resultaterne viser, at ingeniørstuderende, som falder fra studierne inden udgangen af deres tredje studieår, opnår signifikant lavere karaktersnit i deres gymnasiale uddannelse og signifikant lavere testscorer i både den første uforberedte og anden forberedte matematiktest end medstuderende, som enten endnu er aktive eller har gennemført uddannelsen. Den mest markante forskel mellem de to grupper studerendes testscorer ses ved test 2, som var de studerendes første tællende aktivitet og var den test, hvor de studerende havde mulighed for at forberede sig. Samme mønster ses i forhold til beståelse af førsteårsprøven. Der må da også formodes at være et større overlap af studerende, som er faldet fra studiet og ikke har bestået førsteårsprøven, idet studerende, der falder fra i løbet af første studieår, ikke kan have bestået førsteårsprøven, ligesom ikke-beståelse af studieprøven automatisk vil føre til udskrivning af studiet, medmindre den studerende opnår dispensation. I dette studie består overlappet af 98 personer, svarende til, at 83 % af de frafaldne studerende ikke har bestået førsteårsprøven.

Adfærdsvariablen, fremmøde ved første undervisningsgang, viste sig også signifikant. En markant større andel af de ikke-fremmødte studerende har afbrudt deres studie (70,2 %) end de fremmødte studerende (46,4 %). Idet fremmøde til undervisningen i det kursus, hvor matematiktesten indgik, var frivillig, og denne blev udført i en undervisningsgang, var deltagelse i matematiktesten også principielt frivillig. Derfor kan deltagelse i testen ses som et udtryk for den enkelte studerendes villighed til at deltage i kursusaktiviteter og et indirekte mål for engagement. Forfatterne er bevidste om, at flere indikatorer sammen vil give et mere retvisende og samlet billede af de studerendes studieengagement og motivation. Den enkelte studerende har ikke nødvendigvis bevidst fravalgt matematiktesten, men kan have været ufrivilligt forhindret af fx sygdom på testdagen. Til en vurdering af engagement på gruppeniveau, vurderedes det dog, at et enkelt målepunkt, deltagelse i testen, var tilstrækkelig. Det bemærkes dog også, at enkelte studerende, som ikke mødte op til matematiktesten sandsynligvis aldrig har været mødt op på studiet.

I studiet valgte forfatterne at udelukke studerende, som ikke havde deltaget i testene, fra analyser, hvori testscorerne indgik. Hvis ikke-deltagende studerende i stedet var blevet givet en score på nul i de pågældende tests, ville et tydeligere billede kunne tegnes. I så fald ville studerende, som allerede var droppet ud af studiet på testtidspunktet og derfor tillagt en

testscore på nul, sænke gennemsnittet betydeligt. Velvidende at det kunne betyde et mindre tydeligt billede, mente forfatterne det dog mere retvisende kun at medtage reelt opnåede testscorer.

Nærværende studie har ikke fokuseret på forsinkede studerende. Blandt gruppen af endnu aktive studerende vil nogle være forsinkede, mens andre vil være diplomingeniørstuderende, som ved at følge det normerede studieforløb endnu er aktive uden af være forsinkede. Opdelingen afsluttet/aktiv versus afbrudt blev valgt ud fra den betragtning, at de endnu ikke-afsluttede studerende ikke nødvendigvis er forsinkede i deres studieforløb. Fremtidige studier vil med fordel kunne inkludere 'afsluttet', 'forsinket' og 'afbrudt' som tre forskellige mulige udfald af en studiestatusvariabel. Det vil være interessant at undersøge, om gruppen af forsinkede studerende skiller sig ud fra enten de afsluttede, de afbrudte eller begge grupper studerende. Ved et længerevarende studie kan graden af forsinkelse i forhold til normeret studietid for de afsluttede studerende også inkluderes. Ligeledes er karaktersnit på universitetet en relevant studiesuccesindikator, som kan inkluderes.

Studiet inkluderede syddanske ingeniørstuderende fra fire hold i introducerende matematikkurser. Der vil være brug for yderligere studier for at klarlægge studiets resultaters overførbarehed fx til naturvidenskabelige uddannelser og til andre danske geografiske regioner. Der argumenteres dog for en vis gyldighed i andre kontekster, idet andre europæiske og amerikanske studier har fundet lignende indikationer af matematikkompetencer som prædikator for studiesucces (Ackerman et al., 2013; De Winter & Dodou, 2011; Veenstra et al., 2008). Dette studie fandt imidlertid, at forudgående matematikkompetencer (målt ved test 1) havde en mindre effekt på studiesuccesen end de opøvede matematikkompetencer efter en måneds studieaktivitet (målt ved test 2). Her kan muligheden for forberedelse forud for test 2 have været udslagsgivende og have kompenseret for nogle studerendes 'manglende' præstation i test 1. Det er dog stadig bemærkelsesværdigt, at der kun er en svag sammenhæng mellem pointscoren i test 1 og test 2. At have deltaget i test 1 har givet de studerende en forventning om og forsmag på kursets første tællende aktivitet, test 2, og har dermed reelt fungeret som en form for træning, som de ikke-deltagende studerende har været foruden. Udover de forskellige vilkår i forhold til forberedelse var der også forskel i hvem der rettede test 1 (rettet af andre studerende) og test 2 (rettet af underviseren), som kan have påvirket pointgivningen. Dette er dog ikke undersøgt nærmere her, men må betragtes som en mulig bias. Fremtidige studier bør sikre ensartet tildeling af pointscore på tværs af tests.

De matematiktests som benyttedes af Moses et al. (2011) og Vanderoost et al. (2014) er fagligt bredere end den matematiktest, der indgik i dette studie, og som testede basale matematikfærdigheder. Testen i Moses et al. (2011) vurderer, hvorvidt de kommende studerende er klar til matematik på college-niveau – hvis ikke, kan de studerende ikke tage disse kurser. Testen i Vanderoost et al. (2014) er frivillig og benyttes forud for studievalg. Testen betragtes som en hjælp til kommende studerendes afklaring af studievalg og evt. behov for opkvalificering i matematiske kompetencer. Testen omfatter flere forskellige matematiske kompetencer, bl.a. ræsonnementskompetence, rumlig visualiseringskompetence og matematiske færdigheder. Kun sidstnævnte er inkluderet i dette studie, som alligevel identificerer signifikante forskelle i de studerende testscorer. Dette peger på, at der er potentiale for at udvikle en bredere test specifik til den danske kontekst med en betydelig prædiktiv værdi i forhold til studiesucces. Det kan diskuteres om en sådan test bør være obligatorisk eller frivillig, og om testen bør have konsekvenser i forhold til optagelse og adgang til kurser, som i Moses et al. (2011), eller være en guide for kommende studerende, som i Vanderoost et al. (2014). En

tredje mulighed er, at igangværende studerende benytter en matematiktest til at afklare evt. behov for at opsøge faglig støtte. Denne mulighed ligger tæt op ad den måde, testen i praksis blev benyttet i forbindelse med dette studie. En sådan brug af matematiktesten af de deltagende studerende kan være en del af forklaringen på, hvorfor test 2-scoren bedre forudsiger frafald end test 1-scoren. For nogle studerende vil testscoren i test 1 have givet anledning til at gøre en ekstra faglig indsats forud for test 2, og dermed har disse studerende også udvist en særlig vilje til at forblive på studiet.

I diskussionen om frafald hører også en skelnen mellem studieskifte og frafald fra uddannelsessystemet. I dette studie omfatter kategorien 'afbrudt' alle studerende, som er faldet fra den uddannelse, de var indskrevet på, da de tog det undersøgte kursus. En del frafaldne studerende vil sandsynligvis være begyndt på en anden uddannelse sidenhen og således være blevet i uddannelsessystemet som studieskiftede. Uddannelsesinstitutioner vil naturligvis være mest interesserede i, at studerende optages på deres rette uddannelse ved første optagelse på en uddannelse, men en matematiktest, som er aktuell i dette studie, vil ikke kunne skelne 'faglig egnethed' mellem uddannelser, der fagligt har store overlap. Dette være sig fagligt overlappende studier både internt på uddannelsesinstitutionen og på tværs af institutioner. I forbindelse med studieskift vil det være relevant at inddrage faktorer, som ikke er medtaget i dette studie, fx studie- og læringsmiljøet på den pågældende uddannelse. Et eksempel herpå er en overvejende gruppebaseret undervisningstilgang versus en mere individualiseret tilgang til læring. Sådanne forskelle kan mindske generaliserbarheden af dette studies fund. Sådanne forskelle vil også være tæt relateret til mere affektive faktorer for den enkelte studerende, faktorer som blev fravalgt i dette studie grundet deres vanskelige kvantificerbarhed. Studiets fokus har været identifikation af let målbare variable med betydning for ingeniørstuderendes studiesucces.

Nærværende studie bidrager med en konkret handlemulighed i forhold til uddannelsesinstitutioners ønske om tidligt at kunne identificere studerende, som er i risiko for at droppe ud. Endvidere indikerer studiet, at matematiktests kan inkluderes i optagelsesproceduren på relevante studier eller i afklaringen af et STEM-studievalg. De foreliggende resultater er dog endnu ikke klare nok til at udmønte sig i konkrete anbefalinger omkring udformning af en sådan test eller at konkludere på testens prædiktive værdi og overførbare til andre kontekster. Yderligere studier er nødvendige for at afklare den optimale udformning af en matematiktest, og hvorvidt en sådan test vil kunne stå alene, ligesom yderligere variable i forhold til forsinkelse, studieskift m.m. vil være relevante at undersøge nærmere.

### **Konklusion**

Internationale forskere har peget på, at matematikkompetencer er centrale for studiesucces på STEM-uddannelser, dette studie undersøgte, om samme konklusion kan drages omkring danske ingeniørstuderende. Fundene indikerer, at der er en signifikant sammenhæng mellem både forudgående (forud for studiestart) og opøvede (efter en måneds studieaktivitet) matematikkompetencer og studerendes studiesucces (fracfald og beståelse af førsteårsprøven) – den største effekt identificeres dog af de opøvede kompetencer. Der er således basis for at arbejde videre med perspektivet for brug af faglige tests i optagelse af STEM-studerende, og studiet lægger op til en diskussion heraf samt til yderligere studier herom i dansk kontekst. Endvidere inkluderede studiet en adfærdsvariabel, fremmøde ved første undervisningsgang, som også viste sig er være forbundet med lavere frafald. Hvis sådanne enkle adfærdsmarkører kan identificere frafaldstruede studerende tidligt i studiet, skaber

dette også mulighed for at sætte ind overfor den enkelte studerende *før det er for sent*. Begrænsning af frafald ved at optage de *rigtige* studerende og ved at fastholde igangværende studerende er højt på både uddannelsesinstitutionernes og den politiske dagsorden, og gør studiet højaktuelt.

## Referencer

- Ackerman, P.L., Kanfer, R. & Beier, M.E. (2013). Trait complex, cognitive ability, and domain-knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, 105: 3, 911-927.
- Bandura, A. (1986). *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice-Hall.
- Bonne, P. & Nutt, D. (2016) (Eds.). *Ten Times the First Year: Reflections on ten years of the European First Year Experience conference*. Leuven: Lannoo Campus.
- Business Europe (2011). *Plugging the Skills Gap: The clock is ticking*. Hentet fra <https://www.buinessurope.eu/sites/buseur/files/media/imported/2011-00855-E.pdf>
- Danske Universiteter (2018). *Universiteternes Statistiske Beredskab*. Tabel G: Gennemførelse. Hentet fra <https://dkuni.dk/tal-og-fakta/beredskab/>
- De Winter, J.C.F. & Dodou, D. (2011). Predicting academic performances in engineering using high school exam scores. *International Journal of Engineering Education*, 27: 6, 1343-1351.
- European Commission (2004). *Increasing Human Resources for Science and Technology in Europe: Europe needs more scientists* (Report of the High Level Group on Human Resources for Science and Technology in Europe). Luxembourg: Office for Official Publications of the European Communities. Hentet fra European Commission [http://ec.europa.eu/research/conferences/2004/sciprof/pdf/final\\_en.pdf](http://ec.europa.eu/research/conferences/2004/sciprof/pdf/final_en.pdf)
- Engineer the future (2018). *Prognose for STEM-mangel 2015*. Hentet fra [https://engineerthefuture.dk/sites/default/files/prognose\\_for\\_stem-mangel\\_2025\\_endelig\\_med\\_forside.pdf](https://engineerthefuture.dk/sites/default/files/prognose_for_stem-mangel_2025_endelig_med_forside.pdf)
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (fourth edition). London: SAGE Publications.
- French, B. F, Immekus, J. C. & Oakes, W. C. (2005). An Examination of Indicators of Engineering Students' Success and Persistence. *Journal of Engineering Education*, 94: 4, 419-425.
- Harrison, M. (2012). *Jobs and Growth: The Importance of engineering skills to the UK economy* (Royal Academy of Engineering econometrics of engineering skills project). London: Royal Academy of Engineering. Hentet fra <http://www.raeng.org.uk/publications/reports/jobs-and-growth>
- Holmegaard, H.T., Madsen, L.M. & Ulriksen, L. (2014). A journey of negotiation and belonging: understanding students' transition to science and engineering in higher education. *Cultural Studies of Science Education*, 9, 755-786. doi:10.1007/s11422-013-9542-3

- Lawson, D., Croft, T. & Waller, D. (2012). Mathematics support past, present and future. Paper præsenteret ved *International Conference on Innovation, Practice and Research in Engineering Education*, September 18–20, Coventry University, UK.
- McKenzie, K. & Schweitzer, R. (2010). Who succeed at University? Factors predicting academic performance in first year Australian university students. *Higher Education Research & Development*, 20:1, 21-33.
- Midtby, H.S. & Ahrenkiel, L. (2015). Digitale læremidlers potentiale til at støtte udviklingen af matematiske kompetencer. *Matematik- og Naturfagsdidaktik (MONA)*, 3, 29-42.
- Moses, L., Hall, C., Wuensch, K. et al. (2011). Are Math Readiness and Personality Predictive of First-Year Retention in Engineering? *The Journal of Psychology*, 145:3, 229-245, doi:10.1080/00223980.2011.557749
- Seymour, E., & Hewitt, N.M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder: Westview Press.
- Tinto, V. (1987). *Leaving College: Rethinking the causes and cures of student attrition* (first edition). Chicago: University of Chicago Press.
- Tinto, V. (2016). Student success does not arise by change. In P. Bonne & D. Nutt (eds.). *Ten times the first year – Reflections on ten years of the European First Year Experience Conference* (pp. 151-160). Leuven: Lannoo Campus.
- Uddannelses- og Forskningsministeriet (2017a). *Optag 2017: Oversigt, uddannelsesgrupper og institutioner*. Optaget 2017, nr. 1. Notat hentet fra <https://ufm.dk/uddannelse/statistik-og-analyser/sogning-og-optag-pa-videregaende-uddannelser/2017/notat-1-overblik-udvikling-uddannelsesgrupper-og-institutioner.pdf>
- Uddannelses- og Forskningsministeriet (2017b). *Frafald, Studietid og Fuldførelse: Tal om studerende på ordinære videregående uddannelser*. Hentet fra <https://ufm.dk/uddannelse/statistik-og-analyser/frafald-studietid-fuldforelse>
- Undervisningsministeriet (2018). 180 millioner kroner til ny naturvidenskabsstrategi. Hentet fra <https://uvm.dk/aktuelt/nyheder/uvm/2018/mar/180313-180-millioner-kroner-til-ny-naturvidenskabsstrategi>
- Vanderoost, J., Callens, R., Vandewalla, J.P.L. & De Laet, T. (2014). Engineering Positioning Test in Flanders: A powerful predictor for study success? *SEFI2014: 42<sup>nd</sup> Annual Conference*, Birmingham, UK.
- Veenstra, C.P., Dey, E. L. & Herrin, G. D. (2008). Is modeling of freshman engineering success different from modeling non-engineering success? *Journal of Engineering Education*, 97: 4, 467-479.
- Vossensteyn, H., Kottmann, A., Jongbloed, B., Kaiser, F., Cremonini, L., Stensaker, B. et al. (2015). *Dropout and Completion in Higher Education in Europe: Main report*. European Commission: Education and Culture. Luxembourg: Publications Office of the European Union. doi:10.2766/82696

# Screening for technical flaws in multiple-choice items. A generalizability study.

Lotte Dyhrberg O'Neill<sup>a,1</sup>, Sara Mathilde Radl Mortensen<sup>b</sup>, Cita Nørgård<sup>c</sup>, Anne Lindebo Holm Øvrehus<sup>d</sup>, Ulla Glenert Friis<sup>e</sup>

<sup>a</sup>SDU Universitetspædagogik, Syddansk Universitet, <sup>b</sup>Klinisk Institut, Aalborg Universitet, <sup>c</sup>SDU Universitetspædagogik, Syddansk Universitet <sup>d</sup>Klinisk Institut, Syddansk Universitet, <sup>e</sup>Sundhedsvidenskabeligt Fakultetssekretariat, Syddansk Universitet

Research article, peer-reviewed

Construction errors in multiple-choice items are quite prevalent and constitute threats to test validity of multiple-choice tests. Currently very little research on the usefulness of systematic item screening by local review committees before test administration seem to exist. The aim of this study was therefore to examine validity and feasibility aspects of review committee screening for item flaws. We examined the reliability of item reviewers' independent judgments of the presence/absence of item flaws with a generalizability study design and found only moderate reliability using five reviewers. Statistical analyses of actual exam scores could be a more efficient way of identifying flaws and improving average item discrimination of tests in local contexts. The question of validity of human judgments of item flaws is important - not just for sufficiently sound quality assurance procedures of tests in local test contexts - but also for the global research on item flaws.

## Introduction

Multiple-choice tests are particularly useful and effective test formats in test situations where there is a need to test knowledge (factual or applied) across a wide range of different content topics in larger groups of students (Downing & Yudkowsky, 2009). This potential for broad sampling of topics is often pivotal for 'Constructive Alignment' (Biggs & Tang, 2007), and for the validity of the exam or test in question (Swanson, Norcini, & Grosso, 1987). Another distinctly positive feature of multiple-choice tests is that excellent evidence-based guidelines for constructing good quality items exist (Case & Swanson, 2002; Haladyna, Downing, & Rodriguez, 2002; Paniagua & Swygert, 2016). However, little is known about the usefulness of this body of recommendations as a basis for systematic item screening *before* test administration by local review committees.

Quality assurance measures can be applied both before and after multiple-choice exams. *Before* the exam, item writers can themselves attempt to review their own item drafts according to guideline criteria. Such reviews involve critical reading which is qualitative in nature. Subsequently, other reviewers (e.g. colleagues or external examiners) may repeat this process of qualitative reviewing and suggest item corrections or removal (Downing, 2004, 2006; Malau-Aduli & Zimitat, 2012). *After* the exam, quality assurance typically involves quantitative analyses of test results (quantitative review), followed by yet another round of qualitative reviewing, this

---

<sup>1</sup> Kontakt: ldo@sdu.dk

time focused on discounting selected test items which performed aberrantly because of identifiable flaws missed in the initial phases (Case & Swanson, 2002). In other words: decisions based on qualitative reviews of items are fundamental to the overall quality assurance of multiple-choice tests or exams in all stages (drafting, editing and grading).

Undetected item flaws may infuse exam scores with construct-irrelevant variance, which undermines the validity of the exam (Downing, 2002, 2003, 2005; Kane, 2006). Reliable flaw detection is therefore a sine qua non for optimal test validity. Studies have shown that in practice, writing flawed items is a very common event in local educational contexts (Downing, 2002, 2005; Downing & Yudkowsky, 2009; Hansen & Dexter, 1997; Jozefowicz et al., 2002; Masters et al., 2001; Palmer & Devitt, 2007; Rodríguez-Díez, Alegre, Díez, Arbea, & Ferrer, 2016; Tarrant, Knierim, Hayes, & Ware, 2006; Tarrant & Ware, 2008; Vahalia, Subramaniam, Marks, & De Souza, 1995), and it has been suggested that item writing is 'as much art as science' (Downing, 2005; Downing & Yudkowsky, 2009; Ebel, 1951; Haladyna et al., 2002). By logic extension, it would seem a reasonable hypothesis that a qualitative flaw detection process may also be 'as much art as science', since such a process relies on human reading, evaluation and decision skills. Even though guidelines quite clearly outline a number of straightforward flaw types to look out for, screening for item flaws very often involves both interpretation and subjectivity in our experience. However, very little research on reliability aspects of qualitative item screening by committees seems to exist to support or refute this hypothesis (Engelhard Jr, Davis, & Hansche, 1999). As a consequence, the number of board members necessary for reliable item screening processes in a review committee approach does not appear to be well documented either. The human resources necessary for securing exam validity may also have important consequences for the feasibility and acceptability of a test, particularly in local, smaller educational contexts. Apart from being valid, assessments must also be reasonably feasible for test administrators in terms of time and human resources necessary (Dent, Harden, & Hunt, 2017).

#### *Validity assumptions*

The current Standards for Educational Psychological Testing describes validity in these words:

*Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. Validity is therefore, the most fundamental consideration in developing and evaluating tests. The process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations (American Educational Research Association, 2014).*

Test scores are typically used to support claims beyond the observed performance (Kane, 2013), and in such cases either implicit or explicit score interpretations exist. According to modern Validity Theory, construct validity evidence for such score interpretations may arise from all stages of testing, i.e., from the initial development of test content to subsequent implications of decisions based on test scores (American Educational Research Association, 2014; Kane, 2006). Kane suggested that score interpretations (or 'interpretative arguments') may be categorized to relate to the following four stages: scoring, generalization, extrapolation and decision (Kane, 2006).

Proposed score interpretations related to *scoring* in a multiple-choice exam could for instance be: the recording of students' electronic responses represents students' intended answers, the answer key for the test items is appropriate, and the answer key is applied accurately and consistently. Score interpretations relating to *generalization* in a multiple-choice exam could

be: the items in the test are representative of the universe of all possible items which are congruent with the learning outcomes and the teaching and learning activities of the course, and the sample of items is large enough to control for sampling error. Score interpretations relating to *extrapolation* could for example be: the test tasks require the competencies developed in the course, we may extrapolate expertise levels from the test scores, and there are no skill irrelevant sources of variability that seriously bias the interpretation of scores as measures of students' subject knowledge. The fourth category of assumptions about scores relates to *decision* and represents going from conclusions about test takers' competences to making a decision with implications. Score interpretations relating to decision in a multiple-choice test could for example be: students with no or low level of subject knowledge are unlikely to pass the test and progress in the program.

Such a chain of validity assumptions about multiple-choice exam scores may be challenged if test items are flawed. Technical flaws in multiple-choice items tend to 'pollute' the test with competing test constructs, so that it is no longer just subject knowledge which is measured. 'Testwiseness' and 'irrelevant difficulty' have been identified as competing test constructs in items with technical flaws (Case & Swanson, 2002). Testwiseness refers to test takers ability to answer items based on logic rather than on subject expertise, whereas irrelevant difficulty is item difficulty caused by a confusing presentation of item content rather than actual subject difficulty. If competing test constructs (like testwiseness and irrelevant difficulty) are sufficiently influential in the test situation, assumptions or score interpretations such as those outlined above relating to extrapolation and decisions may be compromised. The validity argument for test scores is no stronger than the evidence for the weakest links in the chain of inferences made from scoring to decision. In other words: Technical item flaws have the capacity to be a threat to the validity of multiple-choice tests, and valid technical flaw detection processes are therefore important for the overall test validity of multiple-choice tests. *Flaw detection is also an assessment process* with its own set of validity assumptions relating to scoring, generalization, extrapolation and decision. With a modern validity framework as a backdrop (American Educational Research Association, 2014; Kane, 2006), we suggest that some important validity assumptions for a review committee's qualitative screening for technical item flaws could be: 1) the recording of reviewers' item assessments represents reviewers' intended item assessments, 2) the list of technical flaws to screen for is appropriate, 3) items are checked accurately and consistently for all technical flaw types on the list, 4) the items reviewed actually represent all the items in the proposed tests, 5) the sample of reviewers is large enough to control for reviewer bias, 6) we may extrapolate item quality from the results of the flaw detection, 7) there are no irrelevant sources of variance which bias the interpretation of the results of the flaw assessment as a measure of item quality, 8) items of lower quality are less likely to be included in the test based on the results of the flaw assessment. Assumptions 1-3 relate to scoring, 4 & 5 to generalization, 6 & 7 to extrapolation and 8 to decision.

The aim of this study was to examine aspects of validity and feasibility of a review committee screening approach for the identification of technical item flaws. The objectives were to examine the generalizability of reviewers' item screening for technical flaws, and to exemplify and discuss the potential impact of these results on item quality and feasibility in a local context. These objectives examine validity assumptions 5) and 6) outlined above.



## Methods

### Design

The reliability study was designed as a generalizability study rooted in Generalizability Theory (GT) (Brennan, 2001). Multiple-choice items ( $n=160$ ) were checked independently by raters ( $n=5$ ), who checked for the presence/absence of flaws based on a list of 19 predefined constructions errors. In other words, the items ( $i$ ) were the object of measurement and raters ( $r$ ) were the facet of differentiation. The generalizability design used was 'items crossed with raters', also described as the 'i x r' design in GT (Brennan, 2001). The general formula for the Index of Dependability (the  $\Phi$  or phi coefficient) for reviewers' item screening process is outlined in equation 1, where  $\sigma^2$  is the variance.

$$\Phi = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_r^2 + \sigma_{ir,e}^2} \quad \text{Eq. 1}$$

This coefficient (eq. 1) is not the dependability of any final consensus decisions reached about an item. It describes the extent to which independent reviewers agreed on the categorization of the item as either flawless or flawed. The formula for decision studies for alternative numbers of item reviewers is described in equation 2.

$$\Phi_{Decision} = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_r^2/n_r + \sigma_{ir,e}^2/n_r} \quad \text{Eq. 2}$$

### Participants

Raters were all educationalists with knowledge of the National Board of Medical Examiners (NBME) guidelines on item writing (Case & Swanson, 2002).

Rater 1 had 3 years of experience in teaching item construction for university teachers and a background in the natural sciences. Rater 2 had a medical background and around 13 years of experience with teaching item construction to medical item writers, quality assurance of items and statistical item analyses and evaluations. Rater 3 had a biomedical background and many years of practical experience as an educationalist but was a novice to item screening. Rater 4 was a medical doctor with a PhD in medical education and around 3 years of experience with screening items qualitatively for flaws. Rater 5 had a health science background, a master's degree and a PhD in medical education, and 3 years of experience with teaching item construction to medical item writers, qualitative screening for item flaws, and statistical item analyses and evaluations. None of the raters were experts in all the medical specialist content tested in the exams. We considered this sample of raters a random sample from a universe of educationalists in the national higher education context with a variety of prior practical and theoretical experience with item construction.

### Review Criteria

Item reviewers searched for technical item flaws in items based on criteria rooted in the item writing guidelines of the National Board of Medical Examiners (NBME) (Case & Swanson, 2002). Since the reliability of judgments of item relevance has already been examined by others previously (Norcini & Grosso, 1998), we limited the flaw types that reviewers screened for to be the technical flaws described in table 1, which are related to the undue influence of 'testwise-ness' and 'irrelevant difficulty'.

Table 1. Types of construction errors (n=19) the reviewers screened for.

Error type	Explanation
Unfocused lead-in	The stem of an item should be focused. It should pose a clear question, and it should be possible to arrive at an answer with the options covered
Grammatical cue	One or more distractors do not follow grammatically from the stem
Logical cue	A subset of options is collectively exhaustive
Absolute term in option	Terms such as 'always' or 'never' are in some options
Long correct answer	The correct answer is longer, more specific, or more complete than other options
Word repeats	A word or a phrase is included in the stem and in the correct answer
Convergence	The correct answer includes the most elements in common with the other options
Long/complicated/double options	Options are long, complicated, or double.
Inconsistent use of numeric data	Numeric data in options, such as intervals, are not stated consistently
Vague term in option	Vague terms such as 'rarely' or 'usually' or 'frequently' etc. are used in the options
Non-parallel language in options	The language in the options is not parallel
Illogical order of options	Options are in nonlogical order
AOTA/NOTA option used	'All of the above' or 'None of the above' is used as an option
A tricky/complicated stem	The stem is unnecessarily tricky, complicated or verbose
Inter-dependent items	The answer to an item is 'hinged' to the answer of a related item
Overlapping options	The answer choices should be independent and non-overlapping
Negations in lead-in	Negatively phrased lead-ins containing words such as 'except' or 'not' etc. should be avoided.
Options not in same domain	The choices must stem from the same content dimension or domain (e.g. all diagnoses, tests, treatments, prognoses, disposition alternatives etc.)
Implausible distractor	Distractors which are blatantly absurd or ridiculous should be avoided

#### *Items reviewed*

The items reviewed originated from two multiple-choice exam papers administered in a medical master's degree program at a Danish University. The items were in the One-Best-Answer format (Case & Swanson, 2002), with three answer options (A, B, or C) per item. There is quite solid research-based evidence showing that test items seldom contain more than three useful options anyway (Rodriguez, 2005), and that local test developers may be better off using 3

options as opposed to 4 or 5 options (Haladyna & Downing, 1993), as long as the test contains more than 35-40 items (Downing & Yudkowsky, 2009). The two tests from which the items for this study were drawn contained 80 items each. The overall purpose of the two exams was to test clinical knowledge across a number of medical specialties. A variety of regional medical doctors representing different fields of medical expertise (medical specialties) constructed test items for these exams. Item writers had undergone courses in item construction based on the item writing guidelines of the National Board of Medical Examiners (NBME) in the US (Case & Swanson, 2002). The guidelines and instructions delivered to item writers included thorough explanations and exemplification of all of the technical flaws listed in table 1.

### *Data collection*

Two randomly selected exam papers were drawn from the pool of all existing multiple-choice exam papers developed in the context ( $n_{\text{items}}=160$ ). Each exam paper was independently reviewed by the raters who searched for 19 specific types of item flaws (table 1). Each rater independently filled out a prepared evaluation sheet, indicating which type or types of errors were detected in each item. We subsequently dichotomized the data for the purpose of dependability analyses, so that raters' interpretation of the presence of one or more errors in an item was coded 1, while their interpretation of an absence of errors for an item was coded 0.

### *Analysis*

GENOVA for PC was used to estimate variance components and calculate dependability/phi coefficients for test situations with alternative numbers of reviewers as outlined in equation 2 (Brennan, 2003).

In order to *exemplify and illustrate the potential impact* of the dependability results on item quality and feasibility in the context, we compared three different screening approaches and the effects they would have had on selected test parameters and use of time. The three screening approaches compared were: No screening at all, a qualitative screen, and a quantitative screen. The qualitative screen example was based on the results of our 5 reviewers' evaluations and the inclusion criteria for a test item was that none of the reviewers had found any flaws in the item. In the quantitative screen example, only items with item-rest correlations of 0.15 or above were included in a test, as this has been recommended as minimum levels of acceptable item discrimination by experts (Haladyna, 2012). These three examples of approaches to quality assurance were compared on the following parameters: The number of items that ended up in the test ( $n_{\text{items}}$  included), the mean item difficulty (DIF) of included items, the mean item discrimination (DI) of included items, the proportion of non-functioning distractors of included items, and the total number of screening hours spent.

The item difficulty index (DIF) was the percentage of examinees answering an item correctly (Case & Swanson, 2002). We used the item-rest correlations as the item discrimination index (DI). Non-functioning distractors were defined as the incorrect answer options (distractors) which less than 5% of examinees had chosen (Haladyna & Downing, 1993; Malau-Aduli & Zimitat, 2012). The number of non-functioning distractors were counted and converted to a proportion of the total number of distractor options in the included test items. Differences in DIF, DI, and non-functioning distractors between the three screening examples were analyzed either with t-tests or Fisher's Exact Tests. In addition, flawless items (items judged flawless by all 5 raters simultaneously) were compared to flawed items with respect to DIF, DI and non-

functioning distractors. All item analyses were performed with IC STATA 15. The alpha command in STATA was used to obtain DI values for items.

The calculations of DIF, DI and non-functioning distractors were based on real students' (n=128 and n=166) exam performances in the two authentic exams in which the items had been administered. Individual student cases in the exam data were identified by unique numbers, and researchers in this study did not have access to any keys which could break students' anonymity. This project was exempt from ethics review by the regional ethics committee as surveys, database studies, and quality assurance studies do not require their permission. Permission from the Danish Data Protection Agency was not required either, because the data is not considered sensitive data. Data was stored according to current laws on data protection.

## Results

Item reviewers found that 19-50% (or 31-80) of the 160 items contained at least one flaw, with an average item flaw rate across all five raters of 39% (or 62/160 items). Only 21% (or 34/160) of the items reviewed were categorized as flawless by all five raters simultaneously. In the remaining items (n=126), reviewers found between 1-7 different types of flaws with a median of 2 flaws per item. The most common flaw detected by reviewers was an unfocused lead-in (table 2).

Table 2. Types of construction errors detected in the items (n=160) by reviewers.

Error type	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	SUM
Unfocused lead-in	5	34	15	25	12	91
Grammatical cue	1	0	5	2	0	8
Logical cue	8	5	38	0	1	52
Absolute term in option	1	0	3	0	0	4
Long correct answer	5	12	9	4	3	33
Word repeats	8	1	8	6	0	23
Convergence	13	13	6	4	2	38
Long/complicated/ double options	7	2	1	3	4	17
Inconsistent use of numeric data	1	0	1	1	0	3
Vague term in option	1	3	2	2	0	8
Non-parallel language in options	3	0	1	23	0	27
Illogical order of options	3	1	2	0	0	6
AOTA/NOTA option used	2	0	0	0	0	2
A tricky/complicated stem	5	2	1	2	1	11
Inter-dependent items	0	0	2	0	0	2
Overlapping options	2	1	4	9	1	17
Negations in lead-in	3	1	1	5	2	12
Options not in same domain	8	4	11	1	6	30
Implausible distractor	11	5	7	5	3	31

The estimated variance component values ( $\sigma^2$ ) were: 0.055 (SE=0.010) for items, 0.012 (SE=0.008) for raters, and 0.173 (SE=0.010) for items by raters.

Using five reviewers for item screening yielded a dependability coefficient of 0.60 (table 3).

Table 3. Dependability coefficients ( $\Phi$ ) for alternative numbers of item reviewers.

N reviewers used	$\Phi$
1	0.23
2	0.37
3	0.47
4	0.54
5	0.60
14	0.81
31	0.90

For very high levels of dependability ( $\Phi > 0.90$ ) of a qualitative item screening process, more than 31 reviewers should be used according to these estimates (table 3).

The mean item discrimination (DI) of the 34 items considered flawless by all five reviewers simultaneously was 0.18 compared to 0.14 for the 126 items categorized as flawed by one or more reviewers, and this difference was statistically significant ( $p = 0.017$ ). In contrast, there was no statistically significant difference in item difficulty or in the proportion of non-functioning distractors between items judged flawed and flawless respectively by the five raters.

The qualitative screening approach example could potentially have resulted in discussions of a very large proportion (79%) of the test items reviewed if consensus deliberations among the five reviewers been pursued (table 4).

Table 4. Different screening approach examples and their effects on selected test parameters.

Screening method	Test inclusion criteria	N Items Included (%)	Mean DI (SD)	Mean DIF (SD)	% non-functioning distractors (n/N)	Total screening hours spent
No screening	None	160 (100)	0.15 <sup>a</sup> (0.10)	80 <sup>a</sup> (18)	53% <sup>a</sup> (168/320)	0
Qualitative review (n <sub>r</sub> =5)	Flawless items	34 (21)	0.18 <sup>a</sup> (0.11)	84 <sup>a</sup> (20)	57% <sup>a</sup> (39/68)	30
Quantitative review (n <sub>r</sub> =1)	Items with DI ≥0.15	75 (47)	0.24 <sup>b</sup> (0.06)	0.82 <sup>a</sup> (16)	48% <sup>a</sup> (72/150)	2

DI= discrimination index, DIF= item difficulty, n<sub>r</sub>=number of reviewers used. Values with different superscripts in the same column are significantly different ( $p < 0.05$ ).

## Discussion

Qualitative screening of multiple-choice items for flaws required many reviewers for higher levels of dependability in a local test context. Post-exam statistical screening appeared to be a more feasible way of improving the ability of the tests to discriminate between examinees.

### Flaw rates and types

Previous studies have reported that as many as 20-75% of test items produced displayed at least one flaw (Downing, 2002, 2005; Haladyna & Downing, 1993; Hansen & Dexter, 1997; Jozefowicz et al., 2002; Masters et al., 2001; Palmer & Devitt, 2007; Rodríguez-Díez et al., 2016; Tarrant et al., 2006; Tarrant & Ware, 2008), which seems to suggest that the average flaw rates in our sample of 39% were somewhere in the middle range in comparison. We found 'unfocused lead-ins' to be the most common flaw type in our setting (table 2). Other studies have found similar results, but it is generally difficult to compare the precise composition of flaw types between different studies directly because flaw lists are not completely congruent between studies (Downing 2005; Tarrant and Ware 2008).

### Dependability of qualitative technical flaw detection

We found estimated variance components and corresponding standard errors of 0.055 (SE=0.010) for items, 0.012 (SE=0.008) for raters, and 0.173 (SE=0.010) for items by raters. In comparison, when Norcini and Grosso examined the generalizability of 37 reviewers' judgments of item relevance, they found estimated variance components and corresponding standard errors of 0.135 (SE=0.063) for items, 0.122 (SE=0.073) for raters, and 0.641 (SE=0.074)

for items by raters (Norcini & Grosso, 1998). They concluded that their standard errors were small relative to the size of the variance components and indicative of reasonably good estimation, so a similar judgment of our results seems fair.

The resultant moderate dependability coefficient ( $\Phi=0.60$ ) for an item review using five reviewers confirmed our initial assumption that flaw detection (like item writing) is challenging. It also means that the evidence for validity assumption 5 described above was weak to modest. For very high levels of dependability in technical flaw detection, we would have needed the independent judgements of at least 31 reviewers with comparable backgrounds instead of just 5 (table 3), which would most likely not be a feasible solution in the context. Norcini and Grosso found comparable results when they examined the dependability of judging item relevance with a design similar to ours (Norcini & Grosso, 1998). Thirty-seven medical practitioners independently rated one-best-answer items for relevance for professional practice in general internal medicine on a five-point scale. Results of their variance components analysis indicate that for ratings of item relevance with high ( $\Phi>0.90$ ) levels of dependability, 51 raters should assess the items (Norcini & Grosso, 1998). In contrast to our results and the results of Norcini and Grosso (1998), we found three studies reporting higher levels of reliability in flaw detection. One study examined whether 39 experienced reviewers on an item review committee could accurately identify test items constructed or selected to exhibit 16 different cultural or technical flaws accurately. The results indicated that the mean accuracy rates amongst reviewers were high following directly on from a 60-minute tailored training session (Engelhard Jr et al., 1999). However, the authors also mentioned that 'the specific training provided' in the tailored training session immediately before the review and the 'obviousness of flaws' could have accounted for the substantive results reported. Two studies reported reliability coefficients of 0.89 when using three expert NBME item reviewers. The expert reviewers independently categorized items into just five broader quality categories relating to: the type of knowledge tested in the item (applied or factual knowledge), the item format (one-best-answer versus true-false), and the presence/absence of technical flaws (Jozefowicz et al., 2002; Wallach, Crespo, Holtzman, Galbraith, & Swanson, 2006). The superiority of the expert reviewers used is a plausible explanation for the high levels of reliability presented in these studies. On the other hand, keeping the quality criteria fewer and broader could also have secured higher levels of reliability. Recognizing whether an item contains a vignette or not and whether it is of the one-best-answer format or not would appear to be easier to spot than technical flaws, because the latter comes in so many shapes and sizes (table 1). It is therefore also possible that our flaw focus (technical flaws only) may have contributed to a relatively larger influence of cognitive overload and subsequent errors in the flaw assessment. Our results cannot be interpreted to mean that awareness of technical item flaws is not warranted, as existing evidence indicates that training item writers may improve the quality of test items developed in local contexts for in-house exams (Jozefowicz et al., 2002; Naeem, van der Vleuten, & Alfaris, 2012; Wallach et al., 2006). Likewise, one study reported significant improvements in the quality of multiple-choice exams after the introduction of peer review workshops for item writers in which the focus was on three issues relating to the relevance of items (Malau-Aduli and Zimitat 2012). While longer lists of technical flaws to avoid may be very useful in the general training of item writers, and perhaps as check lists for individual item writers and their peer reviewers reviewing only a limited number of items, our results seem to question their use in systematic judgements of whole tests (many items) by a local review committee.

Apart from the reflections our results may trigger in local educational test contexts about the efficiency of quality assurance processes, the dependability of human judgments of item flaws is also important in the global research on the effects of item flaws on student performances. In research studies where only a few item reviewers have been used and no reliability/generalizability coefficient of the flaw assessment in question has been reported, a concern about the correct categorization of items as flawed/not-flawed is warranted. As we have shown in this study, this categorization is not necessarily straightforward even amongst experienced reviewers.

#### *Impact on validity*

The potential impact of the dependability results reported above on test validity aspects and feasibility in a local context is exemplified in table 4.

In the three examples of approaches illustrated here, there were no significant differences in mean item difficulty (DIF) between the three screening approach examples compared (table 4). However, others have reported that flawed items tended to be more difficult for test takers and to increase fail-rates (Downing, 2002, 2005). In one such study, items with and without flaws intended to test basic science were compared. Results showed that flawed items were 7% or nearly  $\frac{1}{2}$  SD more difficult for students than items without flaws, and the flawed items failed nearly one fourth more students than the flawless items (Downing, 2002). Similar results were also found later in a larger study of four basic science examinations administered to year-one and year-two medical students. Across examinations, only 47% of students passed the flawed items while 53% students passed the flawless items (Downing, 2005). Studies in other settings have since corroborated these results, finding flawed items to be 7.4-12.3% more difficult for students than flawless items (Caldwell & Pate, 2013; Pate & Caldwell, 2014). In addition, one study also found high-achieving students were more likely than borderline students to be penalized by flawed items (Tarrant & Ware, 2008). The results of these studies are evidence of flawed items' potentially negative effects on test validity, and they appear to be in line with the existence of 'irrelevant difficulty' as a competing test construct.

The evidence for the assumption that we could extrapolate item quality from our reviewers' flaw detection (validity assumption 6) was not particularly convincing. Although we found the mean item discrimination was significantly higher (0.18) in the items judged flawless by all reviewers simultaneously compared to the rest of the items (0.14), there was no significant difference in average item discrimination between the 'no screening' approach and the 'qualitative review' approach (table 4). The lack of dependability of the 'qualitative review' example presented here was associated with too much noise and too little useful signal. In contrast, using a 'quantitative review' approach post-exam (based on minimum recommendations of item discrimination) would have been a more efficient way of improving the average item discrimination in the test context (table 4). This ability of a test to discriminate between high and low-ability test takers has been described as 'a fundamental principle of all educational measurement and a basic validity principle' (Downing & Yudkowsky, 2009), and curbing construct-irrelevant variance arising from poorly crafted items is considered one of the important ways of improving item discrimination and overall test validity (Downing, 2002).

We found no statistically significant difference in the proportion of non-functioning distractors in items judged flawed versus flawless by the five reviewers, and no statistically significant differences in the proportions of non-functioning options in the three screening approach examples illustrated in table 4. In contrast, others have reported a significant decrease in the



number of non-functioning item options after the introduction of peer review workshops for item writers where the focus was on item relevance (Malau-Aduli & Zimitat, 2012). Item relevance was not evaluated by our reviewers (table 1).

### *Impact on feasibility*

Multiple-choice tests are often hailed for their efficiency in the scoring process compared to other written test formats, but resources must instead be spent on the training of item writers and on writing items in sufficient numbers. Reviewing items qualitatively will further detract from the feasibility of this test format in local contexts. The five raters in this study spent approximately 6 hours each on a review of the 160 items, and others have reported similar rates for other experienced reviewers (Engelhard Jr et al., 1999). Based on the dependability results described above, we would have needed to spend at least 186 hours screening qualitatively and have had access to 31 experienced reviewers for high levels of dependability, or perhaps even more if we had also included judgments of item relevance, as indicated by the results of a previous study (Norcini & Grosso, 1998). To maintain a core of that many reviewers on an exam committee at course level is most likely an unacceptable challenge in many local educational contexts. In addition, subsequent consensus work about the final fate of each of the many items (n=126 in our case) perceived to be flawed by at least one reviewer would further detract from the feasibility. The question is whether it would be practically possible in very large exam committees. As this example illustrates, insufficient dependability of flaw detection has the power to undermine both exam validity as well as the feasibility of the quality assurance process.

In contrast, using the 'quantitative review' screening approach instead, which is based on students' actual exam performance as opposed to reviewers' judgments, would in our experience probably require around 2-3 hours of work for one person (table 4).

### *Limitations*

The global evidence on the reliability of qualitative human judgments of item flaws appears to be quite sparse. In this study, we concentrated solely on the detection of a range of technical flaws that may favor testwise students or infuse the test with construct-irrelevant difficulty, and thereby threaten test validity. Other fundamental issues, such as whether the rules of the item format are respected, correct keying, and whether the amassed content and taxonomic level of test items is congruent with the course learning goals, the teaching and learning activities in the course as well as with subsequent professional practice etc. are of course also extremely important for the overall test validity.

### *Future research*

Influential assessment literature often originates from large-scale professional test agencies in the English-speaking world, but the resulting *guidelines may represent infeasible and non-transferable paragons of perfection in smaller-scale educational settings*. We believe there is a need to critically and openly examine if and how multiple-choice tests can be valid and feasible in local/small scale test situations around the globe. It would be very helpful if guidelines on how to optimize quality assurance processes in smaller-scale/lower resource contexts were available. Finally, systematic reviews of the effects on test validity of item writer training and of violating proposed item writing principles - in local educational contexts across the globe - also seem to be missing.

## Conclusion

Collectively, we found only weaker evidence for validity assumptions relating to generalization and extrapolation in this study of qualitative item screening by a review committee. Furthermore, review committee quality assessments seemed to have the capacity to detract from the feasibility of multiple-choice exams.

## Acknowledgements

The authors would like to thank the item reviewers who participated in this study.

## Declaration of Interest

The authors declare no conflict of interest.

## References

- American Educational Research Association, A. P. A., National Council of Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington (DC): American Educational Research Association.
- Biggs, J., & Tang, C. (2007). *Teaching for Quality Learning at University* 3rd edition Open university Press: Maidenhead.
- Brennan, R. L. (2001). In *Generalizability Theory*. New York: Springer.
- Brennan, R. L. (2003). GENOVA for PC. Retrieved from <https://education.uiowa.edu/centers/center-advanced-studies-measurement-and-assessment/computer-programs#GENOVA>
- Caldwell, D. J., & Pate, A. N. (2013). Effects of question formats on student and item performance. *American Journal of Pharmaceutical Education*, 77(4), 71.
- Case, S., & Swanson, D. B. (2002). *Constructing written test questions for the basic and clinical sciences* (3rd ed.). Philadelphia (PA): National Board of Medical Examiners.
- Dent, J., Harden, R. M., & Hunt, D. (2017). *A practical guide for medical teachers*: Elsevier Health Sciences.
- Downing, S. M. (2002). Threats to the validity of locally developed multiple-choice tests in medical education: construct-irrelevant variance and construct underrepresentation. *Advances in Health Sciences Education*, 7(3), 235-241.
- Downing, S. M. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37(9), 830-837.
- Downing, S. M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education*, 38(9), 1006-1012.
- Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education*, 10(2), 133-143.

- Downing, S. M. (2006). Twelve steps for effective test development. *Handbook of test development*, 3-25.
- Downing, S. M., & Yudkowsky, R. (2009). *Assessment in health professions education*: Routledge.
- Ebel, R. L. (1951). Writing the test item. *Educational measurement*, 185-249.
- Engelhard Jr, G., Davis, M., & Hansche, L. (1999). Evaluating the accuracy of judgments obtained from item review committees. *Applied Measurement in Education*, 12(2), 199-210.
- Haladyna, T. M. (2012). *Developing and validating multiple-choice test items*: Routledge.
- Haladyna, T. M., & Downing, S. M. (1993). How many options is enough for a multiple-choice test item? *Educational and Psychological Measurement*, 53(4), 999-1010.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309-333.
- Hansen, J. D., & Dexter, L. (1997). Quality multiple-choice test questions: Item-writing guidelines and an analysis of auditing testbanks. *Journal of Education for Business*, 73(2), 94-97.
- Jozefowicz, R. F., Koeppen, B. M., Case, S., Galbraith, R., Swanson, D., & Glew, R. H. (2002). The quality of in-house medical school examinations. *Academic Medicine*, 77(2), 156-161.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (pp. 17-64): ACE/Praeger.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73.
- Malau-Aduli, B. S., & Zimitat, C. (2012). Peer review improves the quality of MCQ examinations. *Assessment & Evaluation in Higher Education*, 37(8), 919-931. doi:10.1080/02602938.2011.586991
- Masters, J. C., Hulsmeyer, B. S., Pike, M. E., Leichty, K., Miller, M. T., & Verst, A. L. (2001). Assessment of multiple-choice questions in selected test banks accompanying text books used in nursing education. *Journal of Nursing Education*, 40(1), 25-32.
- Naeem, N., van der Vleuten, C., & Alfaris, E. A. (2012). Faculty development on item writing substantially improves item quality. *Advances in Health Sciences Education*, 17(3), 369-376.
- Norcini, J., & Grosso, L. (1998). The generalizability of ratings of item relevance. *Applied Measurement in Education*, 11(4), 301-309.
- Palmer, E. J., & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Medical Education*, 7(1), 49.
- Paniagua, M. A., & Swygert, K. A. (2016). *Constructing Written Test Questions For the Basic and Clinical Sciences* (4th ed.). Philadelphia, PA: National Board of Medical Examiners.

- Pate, A., & Caldwell, D. J. (2014). Effects of multiple-choice item-writing guideline utilization on item and student performance. *Currents in Pharmacy Teaching and Learning*, 6(1), 130-134.
- Rodríguez-Díez, M. C., Alegre, M., Díez, N., Arbea, L., & Ferrer, M. (2016). Technical flaws in multiple-choice questions in the access exam to medical specialties ("examen MIR") in Spain (2009–2013). *BMC Medical Education*, 16(1), 47.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3-13.
- Swanson, D. B., Norcini, J. J., & Grosso, L. J. (1987). Assessment of clinical competence: written and computer-based simulations. *Assessment & Evaluation in Higher Education*, 12(3), 220-246. doi:10.1080/0260293870120307
- Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Education in Practice*, 6(6), 354-363.
- Tarrant, M., & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical Education*, 42(2), 198-206.
- Vahalia, K., Subramaniam, K., Marks, S., & De Souza, E. (1995). The use of multiple-choice tests in anatomy: Common pitfalls and how to avoid them. *Clinical Anatomy*, 8(1), 61-65.
- Wallach, P. M., Crespo, L. M., Holtzman, K. Z., Galbraith, R. M., & Swanson, D. B. (2006). Use of a Committee Review Process to Improve the Quality of Course Examinations. *Advances in Health Sciences Education*, 11(1), 61-68. doi:10.1007/s10459-004-7515-8

# Intended and unintended test constructs in a Multiple-Mini admission Interview. A validity study.

Lotte Dyhrberg O'Neill<sup>a,1</sup>, Eva Lykkegaard<sup>b</sup>, Kulamakan Kulasageram<sup>c</sup>

<sup>a</sup>SDU Universitetspædagogik, Syddansk Universitet, <sup>b</sup>Institut for Kulturvidenskaber, Syddansk Universitet, <sup>c</sup>The Wilson Centre for Research in Education, University of Toronto

Research article, peer-reviewed

Admission interviews in higher education may be developed with the intention to select applicants with specific personal competences not captured by traditional grade-based admission. In this study, we examined whether the data structure of multiple-mini admission interview scores supported the presence of communication, empathy, collaboration, and resilience as independent test dimensions. In addition, the associations between the interview scores and unintended test constructs (station format, pre-university grades, age, gender) were examined. Confirmatory and exploratory factor analyses and regression analyses were used to examine interview data from a cohort of Danish medical school applicants. The proposed multi-dimensionality was not supported by the data structure. The influence of the unintended constructs examined was limited or non-existing. These results are in line with the scarce existing literature. This situation makes a priori claims that the multiple-mini interview can measure multi-dimensional personal competences inadvisable, and care should be taken about what is communicated to stakeholders.

## Introduction

Every year, thousands of medical school applicants around the world are selected or rejected based on their performances in admission interviews. Medical schools may aim to select and reject applicants based on specific, predefined, non-academic, personal competencies, although very little evidence exists in the literature to support that such specificity is possible (Albanese, Snow, Skochelak, Huggett, & Farrell, 2003; Eva et al., 2009; Knorr & Hissbach, 2014; Patterson et al., 2016). Leaving society and rejected applicants with the impression that specific personal competencies (such as communication skills or empathy, etc.) were assessed seems unfortunate if judgments are actually much more general in nature.

The last decades have seen the emergence of an interest in admission testing in Danish health science education (O'Neill, Christensen, Vonsild, & Wallstedt, 2014; O'Neill, Vonsild, & Wallstedt, 2013; O'Neill, Hartvigsen, Wallstedt, Korsholm, & Eika, 2011; Wallstedt, 2004), and a subsequent political focus on examining the traditional admission system in Denmark (Danmarks Evalueringsinstitut, 2015). Most recently, the interest in admission testing has spread across faculties and to several large institutions in Danish higher education (Dinesen, 2018). A number of programs now use the Multiple Mini-Interview (MMI) as a selection tool, but only a

---

<sup>1</sup> Contact: ldo@sdu.dk

limited number of Danish effect studies have been published to date (Vonsild, Schibler, & Wallstedt, 2016; Danmarks Evalueringsinstitut, 2017).

The Multiple Mini-Interview (MMI) is the latest major development in selection interviews for medical education. The MMI consists of a series of independent interview stations (typically 8-12), each with different tasks to be solved, and each manned with its own rater. The test time on each station is fixed and relatively short (typically <10 minutes per station), and test takers proceed from station to station at the sound of a bell until all interview stations in the MMI have been completed. Figure 1 gives a birds-eye view of an ongoing MMI in which individual stations are separated with partitions.

*Figure 1. An MMI in operation.*



The MMI was originally described as being useful for testing non-academic competences believed to be important for high quality health care, such as communicative and collaborative competences etc. (Eva, Rosenfeld, Reiter, & Norman, 2004), but few published studies have examined whether such specific assumptions are warranted. There seems to be emerging evidence that MMI scores can predict subsequent pre-graduate and post-graduate performances (Knorr & Hissbach, 2014; Patterson et al., 2016; Pau et al., 2013), although it has been pointed out that presently most of this evidence arises from a single institution (Eva, Reiter, Rosenfeld, & Norman, 2004a; Eva et al., 2012; Eva et al., 2009; Knorr & Hissbach, 2014). In addition, there is evidence for stakeholder acceptability (Patterson et al., 2016; Pau et al., 2013).

When it comes to validity evidence for the 'internal structure' of MMIs (American Educational Research Association, 2014), the results are somewhat equivocal. Evidence of internal structure rooted in generalizability studies tends to display adequate generalizability when sufficient stations and raters are sampled (Knorr & Hissbach, 2014; Patterson et al., 2016; Pau et al., 2013). However, the evidence of internal structure relating to the dimensionality of MMIs remains problematic (Knorr & Hissbach, 2014; Patterson et al., 2016). Patterson et al. reviewed the literature and concluded that 'it is critically important that schools better understand what they are seeking to measure with this approach' (Patterson et al., 2016). We found only a few published studies which examine the intended multi-dimensionality of MMI scores (Hecker et al., 2009; Lemay, Lockyer, Collin, & Brownell, 2007; Oliver, Hecker, Hausdorf, & Conlon, 2014). At the same time, a few studies based on Item Response Theory (IRT) approaches indicate that MMIs perhaps measure a much broader and unidimensional con-

struct (Jones & Forister, 2011; Knorr & Hissbach, 2014; Roberts, Zoanetti, & Rothnie, 2009; Sebok, Luu, & Klinger, 2014).

Because of this general lack of global evidence for the multi-dimensionality of MMIs (Knorr & Hissbach, 2014; Patterson et al., 2016), we felt it was important to examine and publish evidence on the internal structure of MMI scores for global as well as for national and local purposes. Examining the relationships between MMI scores and potential competing constructs such as prior grades was also particularly important in a Danish context, as the test-based admission track (kvote-2 admission) is supposed to offer a real alternative to the grade-based admission track (kvote-1 admission).

### *Aim*

The aim of this study was to examine aspects of validity in an MMI relating to the internal structure and to the relationships to other variables that could be competing constructs. The objectives were to 1) examine whether the factor structure of the MMI scores supported the presence of test domains corresponding to content themes (communication, empathy, collaboration, and resilience), and to 2) examine the association between the MMI scores and unintended competing test constructs such as the station format, age, gender and pre-university grades. The hypotheses were that i) the internal structure of the scores would reflect the four content themes as test domains, and ii) that the MMI scores were not just a reflection of demographic differences, different station formats or general academic abilities.

## **Method**

### *Design*

The study is an observational validation study of the scoring process in an MMI for a cohort of medical school applicants. It has both a cross-sectional element (factor analyses of MMI scores from one occasion), and a retrospective element (regression analyses of the association between competing explanatory variables and MMI scores). The theoretical framework behind the research questions examined is unified Validity Theory (Messick, 1989), where evidence of the 'internal structure' of scores and the 'relationships with other variables' count as sources of construct validity evidence (American Educational Research Association, 2014). With that framework as the backdrop, the hypotheses (i and ii above) are the validity assumptions examined (Kane, 2006).

### *Test takers*

The intended study population was all medical school applicants and raters who participated in the MMI at Aarhus University in June 2016. MMI test takers were selected from a larger pool of approx. 1,800 medical school applicants based on best performance in a written reasoning test (the uniTEST).

### *Raters*

The raters on the MMI stations were a mixture of biomedical lecturers, clinical teachers (practicing regional doctors), and health science doctoral students. Raters participated in a two-

hour training session which involved general information, scoring practice, reflections and discussions before participating as raters.

### *MMI scores*

In 2016, eighty percent of medical students at Aarhus University were admitted via traditional grade-based admission, and the remaining twenty percent via test-based admission. A double-track admission system has been in place in Denmark since the mid 70's, because changing governments wanted to secure broader access to higher education with an alternative to purely grade-based admission. The test-based admission at Aarhus University Medical School consisted of three steps. Step 1 was a requirement of minimally acceptable grades from an upper secondary education. Step 2 was participation in a 95-item multiple choice test of generic reasoning and thinking across the two broad domains of mathematics/science and humanities/social sciences known as the uniTEST, which is developed by the Australian Council for Educational Research. Step 3 was participation in an MMI developed to test non-academic personal competences. It is the latter step we examine in this study.

The MMI was developed and pilot tested by the Centre for Health Sciences Education, Aarhus University in cooperation with invited stakeholders (regional clinicians, lecturers, patient organizations, the national medical association). Four relevant test themes (communication, empathy, collaboration, and resilience) were identified with the Nominal Group Technique by a Delphi group of stakeholders. Eight different MMI stations (2 per test domain) were subsequently developed and pilot tested in November 2015. Preliminary validity evidence has been published relating to content, internal structure, relations with other variables, and consequences of testing based on the pilot test results (Andreassen et al., 2016). A generalizability study (internal structure validity evidence) revealed acceptably high generalizability coefficients ( $G=0.85$  using two raters per station). However, if test results were to be generalized to any other set of MMI stations with alternative station content (station as a random facet), the MMI would have had to consist of at least 25 stations in order to reach a G-coefficient of 0.75 when using two raters per station. In addition, a disattenuated correlation analysis showed only low-moderate correlation coefficients between station pairs with the same theme but different content. In some instances, stations with different themes but similar station format (e.g. 'the situational judgment' format) displayed even stronger correlations (Andreassen et al., 2016; Eva & Macala, 2014), which could suggest a potential station format effect. Examination of the pilot MMI scores' relations to other variables found no significant correlations with selected competing constructs, such as prior academic competences (pre-university grade point averages), gender, or age (Andreassen et al., 2016). However, the pilot test results were based on a small sample ( $n=26$ ) of simulated applicants (volunteering junior medical students) and not necessarily representative of the broader applicant pool. After minor adjustments, the MMI was re-administered in June 2016, this time to real medical school applicants participating in the test-based ('kvote-2') admission track at Aarhus University medical school. The development and pilot testing of the stations has been described in a published report (Andreassen et al., 2016). Table 1 gives an overview of the scales scored and the station format used on each of the eight stations. In addition to the four themes (intended test domains), raters were also asked to rate perceived participant suitability for medicine. This scale gave raters the opportunity to convert important (positive or negative) observations outside the realm of the station theme to an independent score. Raters were explained that they



were allowed to use their subjectivity on the suitability scale, and that the scores on this scale were allowed to deviate from the scores given on the other two scales.

*Table 1. Distribution of scales scored and station formats in the MMI stations.*

Station	Scales	Station format
1	Resilience, communication, suitability for medicine	Behavioral Interview
2	Resilience, communication, suitability for medicine	Situational Judgment
3	Communication, empathy, suitability for medicine	Situational Judgment
4	Communication, empathy, suitability for medicine	Practical/Social task
5	Empathy, communication, suitability for medicine	Situational Judgment
6	Empathy, communication, suitability for medicine	Situational Judgment
7	Collaboration, communication, suitability for medicine	Practical/Social task
8	Collaboration, communication, suitability for medicine	Practical/Social task

*Note:* The first scale listed for each station designates the theme for which station content was developed.

In the station format 'Behavioral Interview' (table 1), participants were asked to reflect on a previous personal experience of the station theme. In the 'Situational Judgment' station format, on the other hand, participants were asked to discuss with the interviewer either a written or a video case reflecting a relevant challenging situation. The 'Practical/Social task' station format involved solving a given practical problem in collaboration with an actor or a fellow participant. On each station, participants received a score on three scales, i.e., in 2 themes as well as a score for general suitability for medicine. This yielded 24 individual scores or data points per participant in total. Each data point was scored on a 7-point Likert scale ranging from 1 (Completely unsatisfactory) to 7 (Excellent). The scores were recorded on a paper scoresheet by raters during the MMI and subsequently transferred to an electronic admission database, and the electronic recordings were double-checked for errors. A generalizability study of the pilot test data showed a phi coefficient of 0.74 with one rater per station, and 0.85 using two raters per station (Andreassen et al., 2016). On the real admission test day, the stations were therefore manned with either one or two raters. For admission process purposes, a station score was calculated for each participant as the average of all scores given on a station across scales and raters (i.e. a number between 1 and 7). The total MMI score for a participant was the sum of the 8 individual station scores (i.e. a number between 8 and 56).

### *Demographic variables*

Applicants were registered in the admission database with their unique Personal Identification Numbers (PIN), which all citizens have. Birthday/age and gender variables were extracted from this 10-digit PIN number. Applicants' pre-university grade-point averages were also registered in the admission database and extracted for the purpose of this study. Examining age and gender as important competing and unintended test constructs is relevant, because Danish universities are not allowed to discriminate applicants on age and gender in admissions. Similarly, if applicants' upper-secondary school GPAs were found to be an influential predictor of MMI performance, it would be problematic not just from a validity perspective as an unintended competing construct, but also for the practical purpose of having a double-track admission system in operation.

### *Data collection*

All data was collected by the administrative leader of the admission process. The researchers received all data from the administrative leader of admissions in June 2016.

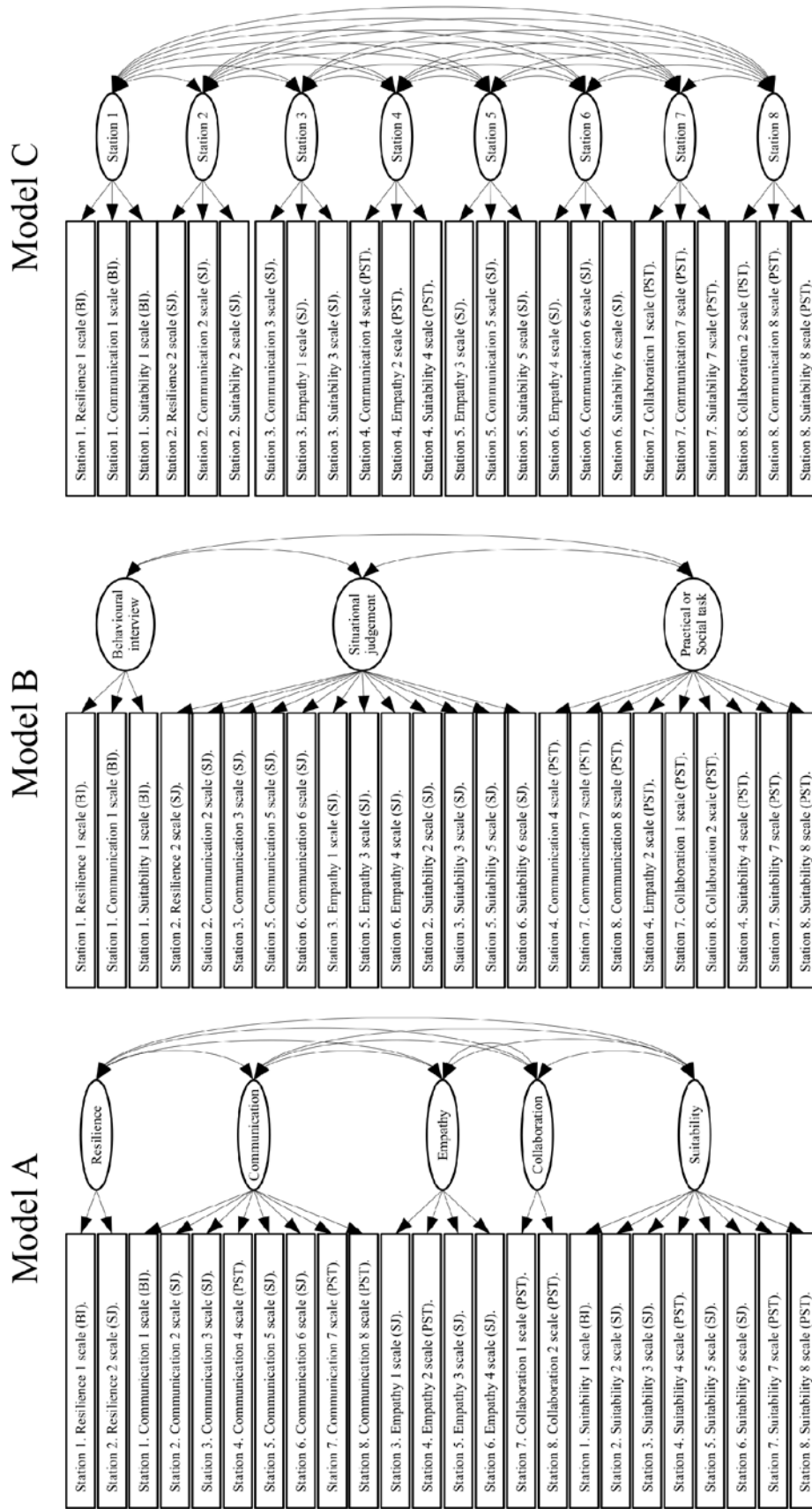
### *Ethics*

The project was exempt from ethics review by the regional (biomedical) ethics committee as database studies and quality assurance studies do not require their permission. We obtained permission to conduct the research study from the internal legal review board at Aarhus University. The researchers adhered to the ordinary rules in the Personal Data Act regarding appropriate data handling in research projects.

### *Analyses*

To determine whether factor analysis was at all appropriate for the data, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was examined. If the KMO was  $>0.5$  a Confirmatory Factor Analyses (CFA) would be used to test whether the intended test domains (Model A in figure 2) or the station formats (Model B in figure 2) best explained the data observed (Field, 2009).

Figure 2. Models examined in the factor analyses.



If neither model was found suitable, it was decided that Exploratory Factor Analysis (EFA) was to be conducted in order to propose a better model. Exploratory factor analysis was conducted in SPSS (IBM, 2012) on the set of 24 items. A promax oblique rotation was chosen and the number of factors was decided quantitatively based on Kaiser's criterion (eigenvalues above 1) and qualitatively based on the interpretability of the factors, meaning whether each factor could be sensibly described and named separately from each other (Field, 2009). Internal consistencies of factors were examined with Cronbach's alpha. The suitability of all models examined was tested by CFA with SPSS-AMOS (Arbuckle, 2012).

Linear regression analyses were used to examine the association between potential competing constructs (age, gender, and GPA) and the MMI scores. All univariate predictors with  $p < 0.10$  were included in a multivariate explanatory model. We used IC STATA 14 for regression analyses and model checking procedures.

## Results

175 applicants (93 males and 82 females) participated in the MMI on the test day, and their characteristics in terms of the other variables examined are summarized in table 2.

*Table 2. MMI participants' characteristics (n=175).*

	n	Median	5% percentile	95% percentile	Min-max
Age (years)	175	21.34	19.33	28.52	18.68-49.20
Pu-GPA	170 <sup>a</sup>	9.20	6.70	10.90	6.00-12.40
MMI scores	175	39.33	29.33	46.83	21.50-50.00

*Note:* n= number of observations, SD= standard deviation, Pu-GPA=pre-university grade point average, MMI=multiple mini interview.

<sup>a</sup>Five MMI participants were foreign applicants without a Danish pu-GPA.

Both models A and B resulted in poor model fit measures (see Table 3), and model re-specifications of both models allowing for co-variations between items did not help sufficiently.

Table 3. Model Fit measures for models A, B, and C.

	CMIN		RMR.GFI		Parsimony-Adjusted Measures		RMSEA	
	p	CMIN/DF	GFI	AGFI	PCFI		RMSEA	PCLOSE
Model A	<0.001	15.49	0.42	0.28	0.20		0.21	<0.001
Model B	<0.001	12.03	0.45	0.34	0.36		0.25	<0.001
Model C	<0.001	1.41	0.87	0.83	0.80		0.05	0.549

*Note:* CMIN = The Chi-square, CMIN/DF = the minimum discrepancy, RMR = the root mean square residual, GFI = the goodness of fit, AGFI = the adjusted goodness of fit, PCFI = the parsimonious comparative fit index, RMSEA = the root mean square error of approximation and PCLOSE is a test for the null hypothesis for RMSEA. The selected threshold values were: p-value > 0.05, CMIN/DF < 5.00, GFI  $\geq$  0.80, AGFI  $\geq$  0.80, PCFI  $\geq$  0.80. RMSEA < 0.08. PCLOSE > 0.05.

Therefore, an exploratory factor analysis was conducted. This resulted in a convincing eight-factor structure (model C in figure 2) accounting for approximately 92% of the total variance and with a high interpretability, as the eight factors each represented exactly one MMI station's three domain scores. The eight factors only correlated weakly (<0.45) with each other. In addition, the internal consistencies of each of the eight factors were all excellent (above 0.9). The quality of the proposed model C (figure 2) was subsequently tested using a confirmatory factor analysis. All threshold values except the p-value indicated good model fit for model C (table 3). However, since the p-value is rather sensitive to sample size, it is defensible to disregard it in a situation like ours, where there is otherwise evidence of good model fit.

Prior grades did not predict MMI performance, whereas a combination of gender and biological maturity was able to explain an estimated 12.6% of the variance in overall MMI scores (table 4).

Table 4. Background characteristics as predictors of MMI performance.

Predictor	Univariate analyses			Final model			
	n	$\beta$ [CI <sub>95%</sub> ]	p	R <sup>2</sup>	$\beta$ [CI <sub>95%</sub> ]	p	R <sup>2</sup>
Gender (female)	175	1.60 [0.05-3.14]	0.043	0.023	2.04[0.53-3.54]	0.008	0.126
Age (years)	175	0.42 [0.24-0.59]	<0.001	0.089	0.45[0.28-0.63]	0.000	
Pu-GPA	170	0.28 [-0.30-0.85]	0.347	0.005	-	-	-

Note:  $\beta$  = regression coefficient, CI<sub>95%</sub> = 95% confidence interval, R<sup>2</sup> = the coefficient of determination.

As seen in the final model (table 4), a female participant in the MMI tended to score 2.04 points higher on average than a male participant of the same age (table 4), corresponding to a difference of <5% of the possible range (7-56 points) of applicants' MMI scores. Likewise, an age difference of 1 year between two participants of the same gender tended to result in an MMI score of 0.45 in favor of the older participant.

## Discussion

This study aimed to examine aspects of validity in an MMI relating to the internal structure and relationships with other variables which could be competing constructs. Our results show that MMI performance appeared to be overwhelmingly station-specific. The proposed multi-dimensionality was not supported by the data structure. Neither station format nor pre-university grades appeared to be competing constructs and the influence of gender and age on test scores was limited.

### *Dimensionality*

The test domains examined in this MMI have also been in use elsewhere (Cleland, Dowell, McLachlan, Nicholson, & Patterson, 2012; Dowell, Lynch, Till, Kumwenda, & Husbands, 2012; K. W. Eva et al., 2004; Gafni, Moshinsky, Eisenberg, Zeigler, & Ziv, 2012; Harris & Owen, 2007; Lemay et al., 2007), but little evidence in support of their actual existence has been published so far. The question of dimensionality of MMIs therefore remains largely unresolved and the existing evidence is sparse, and results are equivocal (Hecker et al., 2009; Lemay et al., 2007; Oliver et al., 2014). The results of this study did not support the presence of four test domains (communication, empathy, collaboration, and resilience) corresponding to the themes used in the blueprint for the development of station content (Andreassen et al., 2016). On the contrary, we found very high correlations (>0.90) between different subscales used on the same stations, which indicates raters' problems with differentiating between multiple test dimensions in a situation. The results generally fit well with what is known about context and context specificity (*situation* or *state*) as the dominant influence on human behavior (Ross & Nisbett, 1991). Findings in social and cognitive psychology have suggested that competences such as problem solving, professionalism, communication, team performance, etc. are highly

context-dependent rather than generic (Eva, 2003; C. P. M. van der Vleuten, 2014). This does not necessarily mean that these personal qualities and the traits/attributes cannot be measured in meaningful ways. For example, models of personality theory have recently begun to engage with the relevance of context in shaping behaviors that are affected by behavioral traits, which are more akin to 'mean norms' with a variability affected by contextual and other interacting factors (Ferguson & Lievens, 2017). Detecting 'stable measurable traits' in admissions tools such as MMIs must address this complex interplay between contextual factors and underlying traits to truly establish construct validity.

This complexity is also reflected in the heterogeneous and conflicting conclusions in previous research on MMI construct validity (Hecker et al., 2009; Lemay et al., 2007; Oliver et al., 2014). Lemay and colleagues evaluated a 10-station MMI which was assumed to test different domains on each station. Their exploratory factor analysis confirmed a 10-factor solution which supported the proposed multi-dimensionality (Lemay et al., 2007). However, the 10 dimensions examined were nested in stations, and so could not be disentangled from situation or context. A second study examined the factor structure of a 5-station MMI used in admission to veterinary school (Hecker et al., 2009). While the intention was to examine the intended five non-academic domains in the MMI, the factor analysis revealed only two factors. Domains such as empathy and moral/ethical reasoning were not discrete and independent as initially assumed (Hecker et al., 2009). In the third study, researchers examined MMI scores for the presence of two conceptually distinct a priori identified, non-academic attributes to be tested. While the data was best explained by a two-factor model rather than a one-factor model, the two proposed domains ('Oral Communication' and 'Problem Evaluation' respectively) correlated very highly (0.87). According to the authors themselves, this limited the ability to conclude that two independent factors were assessed (Oliver et al., 2014). At this point, therefore, there seems to be conflicting evidence from factor analytic studies (including ours) as to the ability of the evaluated MMIs to test multiple dimensions. Similarly, evaluations of the factor structures of Objective Structured Clinical Examinations (OSCE) - the station-based assessment format which inspired the development of the MMI - have also shown heterogeneity with some aligning well with broad factors underlying multiple stations, while others reflect station-specific factors (Volkan, Simon, Baker, & Todres, 2004).

The equivocal results from the published factor analytic studies are further buttressed by a smaller number of studies based on multi-facet Rasch and item-response theory approaches, which have suggested that some MMIs measure a much broader and unidimensional construct (Knorr & Hissbach, 2014). These studies have suggested unidimensional constructs entitled 'entry-level reasoning skills in professionalism', 'latent professional potential', 'suitability for medical school/'professionalism' (Jones & Forister, 2011; Roberts et al., 2009; Sebok et al., 2014). In some contexts, 'general suitability' scales have also been used purposely (Dore et al., 2010; Eva, Reiter, Rosenfeld, & Norman, 2004b; Knorr & Hissbach, 2014). In its basic form, an MMI can be viewed as a serial oral examination, a test format which has previously been shown to be influenced by examinees 'verbal style', their 'capacity to formulate ideas' and their 'communication skills' (Davis & Karunathilake, 2005). It is perhaps also possible that oral communication could be a quite influential underlying unidimensional construct being measured in MMIs. The study by Oliver et al. (2014) described above would certainly seem to indicate difficulties in disentangling oral communication from another test construct (problem evaluation).

### *Competing constructs*

We also examined the potential influence of unintended competing test constructs. Neither station format nor pre-university grades appeared to predict MMI scores, and the influence of sex and age on test scores was limited (table 4). Finding no/weak correlations between these variables and the MMI scores is acceptable, because it indicates that both confounding and sex discrimination were probably not serious validity issues. Others have also found the correlation between MMI scores and prior grades to be insignificant (Eva et al., 2012). The limited influence of interviewees' gender has also been shown in other contexts (Eva et al., 2012; K. W. Eva et al., 2004). But overall, there appears to be relatively little focus on examining the influence of potential confounders and bias in the literature (Eva et al., 2012; K. W. Eva et al., 2004; Moreau, Reiter, & Eva, 2006), which makes it difficult for us to compare our results with the literature to any greater extent. Examination of potential confounders is important both in future validity research (American Educational Research Association, 2014; Kane, 2006; Knorr & Hissbach, 2014), and for ethical and legal purposes in local contexts.

### *Future research*

Further work is necessary to understand the conditions under which MMIs relate to a few underlying factors and subsequent constructs, versus when they represent high levels of multi-dimensionality related to stations. Moreover, underlying relationships between stations and other theoretical sources of validity (e.g., modern personality theory) may be worth considering. Hidden in the debate on measurement is also the role of raters. Rater judgement and decision-making in assessment is receiving renewed attention in medical education (Gauthier, St-Onge, & Tavares, 2016; Tavares, Ginsburg, & Eva, 2016). Raters' idiosyncrasies, cognitive limitations, and cultural attitudes play a role in informing judgement above and sometimes far beyond the raters' abilities to judge candidate performance (Sebok & Syer, 2015; Gauthier et al., 2016). A recently published qualitative study of 12 of the raters in the MMI examined in our study found that raters spontaneously applied subjectivity criteria (their 'taste') to the assessment of applicants. They seemed to share a taste for certain qualities in the candidates (e.g., reflectivity, resilience, empathy, contact, likeness, 'the good colleague') (Christensen, Lykkegaard, Lund, & O'Neill, 2017). Which roles rater idiosyncrasy, subjectivity, and decision-making have in affecting the measured dimensions of MMIs is still an open question worth further exploration. Raters typically infuse assessments with both error variance (bias) and with necessary quality (their expertise) simultaneously. This is a Gordian knot which has yet to be solved in any assessment of complex competences (Govaerts & Vleuten, 2013; Hodges, 2013; C. P. van der Vleuten et al., 2012; C. P. M. van der Vleuten, 2014).

### *Implications for practice*

On a practical note, based on our evaluation of the current literature as discussed above and our own results, as well as current thoughts on validity (American Educational Research Association, 2014; Kane, 2006), programs should probably be quite cautious about claiming to test specific traits without local evidence in support of such claims. The challenge is to remember to distinguish between station themes and content used and theoretical constructs, and to be careful about the communication with stakeholders – particularly the rejected applicants - when explaining what MMI results and decisions mean and do not mean. In the



meantime, the purpose of the MMI should probably be described in much broader terms, for example, along these lines: The MMI aims to collect information concerning personal qualities in the broadest sense, and across a range of different situations perceived to be of relevance for the culture and the society in which admitted applicants will be practicing.

### *Limitations*

Measurement will necessarily depend heavily not only on station content and situation, but also on many other aspects such as the rater population and the local medical education culture which creates value judgements around appropriate and inappropriate displays of personal qualities. This may well affect the transferability of our conclusions to other contexts. This study may also be limited by its use of a smaller sample of test takers representing only one admission cohort in one medical school. Finally, it was also a limitation that only smaller sample sizes of stations assessed each of the four domains examined, although this shortcoming is probably quite representative of the situation in many contexts.

### **Conclusion**

This study adds to the growing body of literature on the assessment of personal qualities in admissions and specifically on the construct validity evidence for the use of MMIs. Our evaluation of the MMI at Aarhus University medical school suggests that factors such as age and gender only contributed to MMI score variance to lesser extents, and the dimensions measured in our MMI were likely highly context-specific manifestations of personal qualities within each MMI station. This result - and existing international evidence on MMI dimensionality - strongly suggests *against* an unquestioning acceptance of claims of stable multi-dimensional personal qualities being assessed in MMIs. Local evaluations of MMIs should carefully consider the contextual elements of implementation, the communication about what is being measured, and empirical verification of measurements.

### **Declaration of interest**

The authors declare no conflict of interest.

### **References**

- Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine*, 78(3), 313-321.
- American Educational Research Association, A. P. A., National Council of Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington (DC): American Educational Research Association.
- Andreassen, P., Pedersen, K., Jensen, R. D., Møller, J. E., Carlsen, C. G., & O'Neill, L. (2016). Optagelsessamtaler på medicin-studiet ved Aarhus Universitet.
- Arbuckle, J. L. (2012). *IBM® SPSS® AMOS 21. User's guide*. Chicago: IBM.
- Christensen, M. K., Lykkegaard, E., Lund, O., & O'Neill, L. D. (2018). Qualitative analysis of MMI raters' scorings of medical school candidates: A matter of taste? *Advances in Health Sciences Education*, 23(2), 289-310.

- Cleland, J., Dowell, J., McLachlan, J., Nicholson, S., & Patterson, F. (2012). Research report: Identifying best practice in the selection of medical students (literature review and interview survey). *London: General Medical Council.*
- Danmarks Evalueringsinstitut. (2015). *Universiteternes organisering af optag til bacheloruddannelserne*. København: Danmarks Evalueringsinstitut.
- Danmarks Evalueringsinstitut. (2017). *Effekten af optagelsessamtaler på læreruddannelsen*. København: Danmarks Evalueringsinstitut.
- Davis, M. H., & Karunathilake, I. (2005). The place of the oral examination in today's assessment systems. *Medical Teacher*, 27(4), 294-297.
- Dinesen, T. (2018). Testbaseret optag vinder frem. *Magisterbladet*, 3, 7.
- Dore, K. L., Kreuger, S., Ladhani, M., Rolfson, D., Kurtz, D., Kulasegaram, K., Cullimore, A. J., Norman, G. R., Eva, K. W., Bates, S., Reiter, H. I. (2010). The Reliability and Acceptability of the Multiple Mini-Interview as a Selection Instrument for Postgraduate Admissions. *Academic Medicine*, 85(10), S60-S63.  
doi:10.1097/ACM.0b013e3181ed442b
- Dowell, J., Lynch, B., Till, H., Kumwenda, B., & Husbands, A. (2012). The multiple mini-interview in the UK context: 3 years of experience at Dundee. *Medical Teacher*, 34(4), 297-304.  
doi:10.3109/0142159X.2012.652706
- Eva, K. W. (2003). On the generality of specificity. *Medical Education*, 37(7), 587-588.  
doi:10.1046/j.1365-2923.2003.01563.x
- Eva, K. W., & Macala, C. (2014). Multiple mini-interview test characteristics: 'tis better to ask candidates to recall than to imagine. *Medical Education*, 48(6), 604-613.  
doi:10.1111/medu.12402
- Eva, K. W., Reiter, H. I., Rosenfeld, J., & Norman, G. R. (2004a). The Ability of the Multiple Mini-Interview to Predict Preclerkship Performance in Medical School. *Academic Medicine*, 79(10), S40-S42.
- Eva, K. W., Reiter, H. I., Rosenfeld, J., & Norman, G. R. (2004b). The Relationship between Interviewers' Characteristics and Ratings Assigned during a Multiple Mini-Interview. *Academic Medicine*, 79(6), 602-609.
- Eva, K. W., Reiter, H. I., Rosenfeld, J., Trinh, K., Wood, T. J., & Norman, G. R. (2012). Association between a medical school admission process using the multiple mini-interview and national licensing examination scores. *JAMA*, 308(21), 2233-2240.  
doi:10.1001/jama.2012.36914
- Eva, K. W., Reiter, H. I., Trinh, K., Wasi, P., Rosenfeld, J., & Norman, G. R. (2009). Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education*, 43(8), 767-775. doi:10.1111/j.1365-2923.2009.03407.x
- Eva, K. W., Rosenfeld, J., Reiter, H. I., & Norman, G. R. (2004). An admissions OSCE: the multiple mini-interview. *Med Educ*, 38(3), 314-326.

- Ferguson, E., & Lievens, F. (2017). Future directions in personality, occupational and medical selection: myths, misunderstandings, measurement, and suggestions. *Advances in Health Sciences Education*, 22(2), 387-399. doi:10.1007/s10459-016-9751-0
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.
- Gafni, N., Moshinsky, A., Eisenberg, O., Zeigler, D., & Ziv, A. (2012). Reliability estimates: behavioural stations and questionnaires in medical school admissions. *Medical Education*, 46(3), 277-288.
- Gauthier, G., St-Onge, C., & Tavares, W. (2016). Rater cognition: review and integration of research findings. *Medical Education*, 50(5), 511-522. doi:10.1111/medu.12973
- Govaerts, M., & Vleuten, C. P. (2013). Validity in work-based assessment: expanding our horizons. *Medical Education*, 47(12), 1164-1174.
- Harris, S., & Owen, C. (2007). Discerning quality: using the multiple mini-interview in student selection for the Australian National University Medical School. *Medical Education*, 41(3), 234-241.
- Hecker, K., Donnon, T., Fuentealba, C., Hall, D., Illanes, O., Morck, D. W., & Muelling, C. (2009). Assessment of applicants to the veterinary curriculum using a multiple mini-interview method. *Journal of Veterinary Medical Education*, 36(2), 166-173.
- Hodges, B. (2013). Assessment in the post-psychometric era: Learning to love the subjective and collective. *Medical Teacher*, 35(7), 564-568. doi:10.3109/0142159X.2013.789134
- Jones, P. E., & Forister, J. G. (2011). A comparison of behavioral and multiple mini-interview formats in physician assistant program admissions. *The Journal of Physician Assistant Education*, 22(1), 36-40.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (pp. 17-64). Westport: ACE/Praeger.
- Knorr, M., & Hissbach, J. (2014). Multiple mini-interviews: same concept, different approaches. *Medical Education*, 48(12), 1157-1175. doi:10.1111/medu.12535
- Lemay, J. F., Lockyer, J. M., Collin, V. T., & Brownell, A. K. W. (2007). Assessment of non-cognitive traits through the admissions multiple mini-interview. *Medical Education*, 41(6), 573-579.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-103). New York: American Council on Education/Macmillan.
- Moreau, K., Reiter, H., & Eva, K. W. (2006). Comparison of aboriginal and nonaboriginal applicants for admissions on the Multiple Mini-Interview using aboriginal and nonaboriginal interviewers. *Teaching and Learning in Medicine*, 18(1), 58-61.
- Oliver, T., Hecker, K., Hausdorf, P. A., & Conlon, P. (2014). Validating MMI scores: are we measuring multiple attributes? *Advances in Health Sciences Education*, 19(3), 379-392.
- O'Neill, L. D., Christensen, M. K., Vonsild, M. C., & Wallstedt, B. (2014). Program specific admission testing and dropout for sports science students: a prospective cohort study. *Dansk Universitetspædagogisk Tidsskrift*, 9(17), 55-70.

- O'Neill, L., Hartvigsen, J., Wallstedt, B., Korsholm, L., & Eika, B. (2011). Medical school dropout-testing at admission versus selection by highest grades as predictors. *Medical Education*, 45(11), 1111-1120.
- ONEILL, L. D., VONSILD, M. C., & WALLSTEDT, B. (2013). Kvote 2 optagelse og akademiske præstationer: Hvor stor betydning har det adgangsgivende eksamenssnit? *Dansk Universitetspædagogisk Tidsskrift*, 8(14), 86-99.
- Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education*, 50(1), 36-60. doi:10.1111/medu.12817
- Pau, A., Jeevaratnam, K., Chen, Y. S., Fall, A. A., Khoo, C., & Nadarajah, V. D. (2013). The Multiple Mini-Interview (MMI) for student selection in health professions training – A systematic review. *Medical Teacher*, 35(12), 1027-1041. doi:10.3109/0142159X.2013.829912
- Roberts, C., Zoanetti, N., & Rothnie, I. (2009). Validating a multiple mini-interview question bank assessing entry-level reasoning skills in candidates for graduate-entry medicine and dentistry programmes. *Medical Education*, 43(4), 350-359.
- Ross, L., & Nisbett, R. (1991). *The person and the situation: Perspectives of Social Psychology* McGraw-Hill. New York.
- Sebok, S. S., Luu, K., & Klinger, D. A. (2014). Psychometric properties of the multiple mini-interview used for medical admissions: findings from generalizability and Rasch analyses. *Advances in Health Sciences Education*, 19(1), 71-84.
- Sebok, S. S., & Syer, M. D. (2015). Seeing things differently or seeing different things? Exploring raters' associations of noncognitive attributes. *Academic Medicine*, 90(11), S50-S55.
- Tavares, W., Ginsburg, S., & Eva, K. W. (2016). Selecting and simplifying: Rater performance and behavior when considering multiple competencies. *Teaching and Learning in Medicine*, 28(1), 41-51.
- van der Vleuten, C. P. M., Schuwirth, L., Driessen, E., Dijkstra, J., Tigelaar, D., Baartman, L., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, 34(3), 205-214.
- van der Vleuten, C. P. M. (2014). When I say ... context specificity. *Medical Education*, 48(3), 234-235. doi:10.1111/medu.12263
- Volkan, K., Simon, S. R., Baker, H., & Todres, I. D. (2004). Psychometric structure of a comprehensive objective structured clinical examination: a factor analytic approach. *Advances in Health Sciences Education*, 9(2), 83-92.
- Vonsild, M. C., Schibler, A. H., & Wallstedt, B. (2016). Testbaseret optag til videregående uddannelser. *Dansk Universitetspaedagogisk Tidsskrift*, 11(20), 130-143.
- Wallstedt, B. (2004). Optagelse af studerende til lægeuddannelsen. *Ugeskrift for læger*, 166, 1980-1983.

# Kan videregående uddannelser mindske frafald ved hjælp af læringsteknologi? Et systematisk review

Maria Hvid Stenalt<sup>a,1</sup>, Berit Lassesen<sup>b</sup>, Dorte Sidelmann Rossen<sup>c</sup>, Anna Bager-Elsborg<sup>d</sup>

<sup>a,b,c,d</sup>Center for Undervisning og Læring, Aarhus Universitet

Faglig artikel, fagfællebedømt

Denne artikel undersøger to centrale forhold for uddannelseskvalitet, hhv. frafald og læringsteknologi, hvis sammenhæng danske universiteter i stigende grad interesserer sig for. Formålet med undersøgelsen er at afdække, hvorvidt brugen af læringsteknologi kan bidrage til at reducere de studerendes frafald på videregående uddannelser, og under hvilke betingelser? Det sker ved hjælp af et systematisk review af nyere international, peer-reviewed litteratur, der beskriver sammenhængen mellem frafald og læringsteknologi på de videregående uddannelser. Ud af 1.483 fremsøgte referencer levede 10 studier op til kriterierne for inklusion. Resultaterne af denne undersøgelse indikerer, at man ved hjælp af læringsteknologi og en forholdsvis lille indsats i nogle sammenhænge kan fastholde og motivere de studerende til at engagere sig i faget. Samtidig finder vi også, at det ikke er muligt at løse frafaldsproblematikker ved blot at tilføje mere læringsteknologi i uddannelserne.

## Introduktion

Mindskelse af frafald blandt studerende er et centralt universitetspolitisk mål (f.eks. Uddannelses- og Forskningsministeriet, 2018). Udover at den enkelte studerende kan opleve det at forlade en uddannelse som et nederlag, medfører frafald også et økonomisk tab for uddannelsesinstitutionerne, ligesom det udfordrer politiske målsætninger for befolkningens uddannelsesniveau og for studerendes samlede gennemførelsestid. Ifølge en undersøgelse fra Styrelsen for Forskning og Uddannelse er førsteårs-frafaldet på 16 % på de videregående uddannelser for højt (Styrelsen for Forskning og Uddannelse, 2018). I forlængelse af dette kan man spørge, hvad uddannelsesinstitutionerne kan gøre for at mindske frafald.

Af styrelsens rapport fremgår det, at årsagerne til frafald skyldes flere forhold: Den væsentligste faktor omhandler uddannelsens vægtning af faglige aktiviteter, herunder det overvejende fokus på teoretiske fremfor praktiske aspekter (52 %). Mange studerende angiver endvidere, at kravene til det faglige niveau er for høje (45 %). Derudover fremhæves oplevelsen af ensomhed på studiet (43 %) og et ringe socialt miljø (38 %) som forhold, der har betydning for de studerendes beslutning om at forlade uddannelsen. Andre studier af frafald har identificeret tilsvarende faktorer, der har betydning for frafald på de videregående uddannelser. Review-studier af frafald af hhv. Troelsen (Troelsen, 2011) og Ulriksen m.fl. (Ulriksen, Madsen, & Holmegaard, 2010) peger på, at flere faktorer har betydning for frafald: studerendes sociale og kulturelle baggrund, social og faglig integration (f.eks. at kunne indfri kravene, og at faget opleves som interessant og relevant for den studerende selv), undervisningen og den

---

<sup>1</sup> Kontakt: mstenalt@au.dk

uformelle kontakt til undervisere og andet universitetspersonale samt det opnåede karaktergennemsnit fra den adgangsgivende uddannelse (højere karakter reducerer risikoen for frafald). Studierne indikerer også, at årsager til frafald både kan ligge forud for studiestart, uden for uddannelsen og i selve uddannelsen. Ikke desto mindre er der i undersøgelser af frafald en tendens til at fokusere på de faktorer, der ligger forud for uddannelserne eller omkring studiestart, f.eks. de færdigheder som studerende har med sig, eller de forventninger som studerende har til uddannelsen (Ulriksen et al., 2010). Man kunne måske i stedet for ønske sig, at flere undersøgelser fokuserede på uddannelseskulturen på universitetet (Ulriksen et al., 2010).

På den baggrund er der god grund til at interessere sig for, om forløbet undervejs i uddannelserne (det der sker i undervisningen, strukturer omkring undervisningen og de studerendes læreprocesser i løbet af studiet) også kan indvirke på frafaldet.

I det følgende afsnit vil vi rette fokus mod en anden universitetspolitisk indsats, som nu også bliver koblet til frafald.

### *Læringsteknologi og frafald*

I disse år investerer danske universiteter betydelige summer i læringsteknologi (Aarhus Universitet, 2018; Copenhagen Business School, 2017; Roskilde Universitet, 2017). Formålet med mange af disse tiltag er at bruge teknologien til at skabe fleksibilitet, variation og struktur i undervisningen og herigennem give mulighed for at fremme de studerendes faglighed. Brugen af læringsteknologi kobles også til fastholdelse af studerende. Aarhus Universitets strategiske rammekontrakt med ministeriet for 2018–2021 beskriver eksempelvis, hvordan nytænkning af underviseres didaktiske praksis koblet med teknologiens muligheder skal have bedre fastholdelse og faglig trivsel som ønskede sideeffekter (Aarhus Universitet & Uddannelses- og Forskningsministeriet, 2018, p. 9). Tilsvarende beskriver Syddansk Universitet i deres strategiske rammekontrakt for 2018–2021, at målsætningen for at eksperimentere med pædagogik, didaktik og digitale teknologier på de enkelte uddannelser blandt andet er et højere læringsudbytte blandt studerende, mindre frafald og hurtigere gennemførelse (Syddansk Universitet & Uddannelses- og Forskningsministeriet, 2018, p. 12).

*Læringsteknologi* har i praksis været anvendt i mere end 15 år på de danske universiteter (Dørup, Gomme, Hansen, & Heiberg, 2005). Ikke desto mindre kan det være svært at vide, hvad begrebet reelt dækker over. Tænk tanken DEA beskriver læringsteknologi som en samlet betegnelse for "undervisningsredskaber- og tilgange, der hviler på en digital infrastruktur" (Jensen, Lindorf, Norn, & Aller, 2015, s. 10). Med afsæt i den definition omfatter læringsteknologi et bredt spænd af selvstændige digitale teknologier såvel som samlinger af teknologier. Selvstændige teknologier kan være student response-systemer, blogs eller videokonference-systemer, der kan anvendes til understøttelse af forskellige undervisningsaktiviteter. Læringsteknologi kan også beskrive learning management-systemer (LMS), der samler et helt universitets eller fakultets behov for flere læringsteknologier. Et LMS kan understøtte gennemførelse af kurser i forskellige leveringsformater, fordi det rummer flere funktioner, såsom kursusadministration, distribution af kursusmaterialer og online kursuskommunikation, -aktiviteter og -afleveringer. Kendetegnende for teknologierne er således også, at de anvendes i en sammenhæng, hvor de understøtter undervisnings- eller læringsaktiviteter. For at beskrive læringsteknologi i brug eller dens forankring i undervisningen anvender flere betegnelsen *blended learning* (Willis et al., 2017). Vi vil senere præcisere, hvordan *blended learning* skal forstås i nærværende undersøgelse.

Inden for forskning såvel som i praksis er der stor interesse for effekten af læringsteknologi og scenarier for god brug af læringsteknologi. Via en gennemgang af empiriske studier, der angår undervisning og læring ved hjælp af teknologi, identificerer Price og Kirkwood (Price & Kirkwood, 2011) en række temaer, som teknologi har vist sig velegnet til at understøtte. Temaerne er bl.a. (Price & Kirkwood, 2011, s. 3-4):

- Mere fleksibilitet og let adgang til materialer
- Højere engagement blandt studerende
- Bedømmelse og feedback
- Udvikling af akademiske færdigheder
- Analyse af egen læring og personlig udvikling
- Samarbejde om indholdet og interaktion mellem medstuderende
- Kobling af teoretiske og praktiske aspekter
- Forberedelse af de studerende på arbejdslivet og eget liv

I nyere studier af studerendes brug af institutionaliseret læringsteknologi og studerendes oplevelse af fordelene herved, er det særligt strukturelle forhold, studerende vægter højt, herunder muligheden for at organisere deres uddannelse og kurser samt muligheden for at have adgang til undervisningsmateriale fra flere platforme (Caviglia, Dalsgaard, Davidsen & Ryberg, 2018; Henderson, Selwyn & Aston, 2017). Fra undervisningseksperimenter i en dansk universitetskontekst er der ikke desto mindre eksempler på, at læringsteknologi kan understøtte aktiviteter og læringsmål, der er relateret til studerendes faglige og sociale integration. Styrkelsen af den faglige integration kan blandt andet foregå ved at tilrettelægge undervisningsforløb, hvor studerende vha. online diskussionsfora diskuterer faglige problemstillinger (Møller, 2016), giver eller modtager feedback via et online peer-feedback-system (Bertel & Bentzen, 2018) eller i fællesskab reflekterer over stoffet vha. blogs (Danbjørg, Clemensen, Hansen, Thrysoe & Rothmann, 2018). Den sociale integration kan støttes ved at understøtte samarbejdet de studerende imellem, f.eks. i form af digitale kommunikationsaktiviteter (Davidsen & Ryberg, 2016). De hidtidige erfaringer med læringsteknologi peger således på, at læringsteknologi kan inddrages i uddannelser og herigennem indvirke på nogle af de faktorer, der er fokus på i frafaldsteori (Tinto, 1975).

### **Forskningsspørgsmål**

Der er os bekendt mangel på studier, der afdækker eller kan bekræfte, hvorvidt der er en sammenhængen mellem frafald og inddragelse af læringsteknologi i en universitetskontekst. Der er derfor god grund til at undersøge dette – ikke mindst på grund af en stadigt stigende interesse for læringsteknologi som middel til at øge kvaliteten af de studerendes læring på de videregående uddannelser og det fortsatte uddannelsespolitiske fokus på lavt frafald som mål for uddannelseskvalitet. Vi ønsker derfor at opnå en bedre forståelse af dette, såvel som en bedre forståelse af forhold, der kan bidrage til at reducere (eller øge) frafald i sammenhæng med brug af læringsteknologi.

Undersøgelsens overordnede mål er at besvare følgende spørgsmål:

- I hvilken grad kan brugen af læringsteknologi på videregående uddannelser bidrage til at mindske studerendes frafald, og under hvilke betingelser?

Med afsæt i forskningsspørgsmålet har vi gennemført et systematisk review af empiriske studier, der undersøger frafald på de videregående uddannelser i sammenhænge, hvor læringsteknologi indgår. Vi har valgt denne metode, fordi vi vil afdække publiceret viden inden for området og integrere de forskelligartede studiers resultater inden for en samlet universitetspædagogisk ramme. Til gennemførelse af det systematiske review læner vi os op ad systematikken, som den ses beskrevet hos Littell et al. (Littell, Corcoran & Pillai, 2008).

## Metode

### *Kriterier for studierne*

For at sikre, at studier, der indgår i reviewet: (1) angik universiteters traditionelle udbud af kurser og uddannelser (2) undersøgte studerendes frafald i et kursus eller på en uddannelse, hvor læringsteknologi anvendes og (3) undersøgte årsager til forskelle i frafald eller tilbyder forklaringer, der tager afsæt i kursets karakteristika (i modsætning til studerendes karakteristika) opstillede vi en række inklusions- og eksklusionskriterier (se Tabel 1).

*Tabel 1. Inklusions- og eksklusionskriterier*

Kriterier	Inklusion	Eksklusion
Middel	Studier, der angår brugen af blended learning/læringsteknologi	Studier, der ikke angår brugen af blended learning/læringsteknologi
Kontekst	Studier, der angår traditionel undervisning på videregående uddannelser. dvs. bachelor-, master- og kandidatuddannelser	Studier, der ikke angår traditionel undervisning på videregående uddannelser. F.eks. massive open online courses (MOOCs), ph.d.-uddannelser og erhvervsskoler
Tidsafgrænsning	Nyere studier (2007 - 2017)	Studier, der ligger før 2007
Mål	Studier, der angår udbyttet af teknologi i forhold til fastholdelse af studerende eller frafald blandt studerende.	Studier, der ikke angår udbyttet af teknologi i forhold til fastholdelse af studerende eller studerendes frafald  Studier, der beskæftiger sig med fastholdelse af viden (genkaldte/erindre) - frem for fastholdelse af studerende
	Studier, der operationaliserer frafald.	Studier, der ikke operationaliserer frafald.



Publikationstype	Peer-reviewed tidsskriftspublikationer	Ikke-peer-reviewed tidsskriftsstudier, rapporter og lign.
Sprog	Engelsk	Andre sprog end engelsk
Analyseobjekt	Studier, der angår variation på kursus- eller uddannelsesniveau	Studier, der ikke angår variation på kursus- eller uddannelsesniveau, f.eks., studier, der angår variation vedr. studenterkarakteristika

Forskning i læringsteknologi er et relativt nyt felt, og udviklingen af læringsteknologier sker med stor hastighed. For at sikre en vis aktualitet i forhold til de anvendte teknologier i studierne opstillede vi en tidsbegrænsning på de medtagne studier. Således afgrænsede studierne sig til de sidste 10 år (publiceret mellem 2007 – 2017). Endvidere indgår udelukkende studier der er udgivet i engelsksprogede peer-reviewet tidsskrifter.

#### *Søgestrategi*

Søgestrategien bygger på undersøgelsens inklusions- og eksklusionskriterier. Identifikation af databaser såvel som søgestreng blev foretaget med afsæt i review guidelines inden for læringsteknologi (Henrie, Halverson & Graham, 2015; Price & Kirkwood, 2011) suppleret med en identifikation af databaser, der indeholder anerkendte universitetspædagogiske tidsskrifter. Derudover gennemførte vi prøvesøgninger for at sikre, at søgestrategien hverken blev for inkluderende (bred) eller for ekskluderende (smal). Søgestrategien indebar, at vi alene identificerede de studier, der i enten titel, abstract eller keywords beskrev, at de angik læringsteknologi, frafald og 'higher education'. Den endelige søgestrategi (illustreret i Tabel 2) blev anvendt i det omfang, det var muligt i de enkelte databaser.

*Tabel 2. Søgetermer og -strategi*

Kategorier	Termer	Søgefelter
Mål (frafald)	retention OR attrition OR drop-out OR turnover OR "at risk"	Title or abstract Or keywords/subject terms
Middel (læringsteknologi)	"blended learning" OR "computer assisted instruction" OR "computer managed instruction" OR "courseware" OR "distance education" OR "electronic learning" OR "integrated learning systems" OR "intelligent tutoring systems" OR "online courses" OR "mobile learning" OR "virtual classrooms" OR "web based Instruction" OR "technology mediated" OR "online learning" OR "Educational IT" OR "technology enhanced learning" OR "technology supported learning" OR "Hybrid learning" OR "technology"	Title or abstract OR subject term

Kontekst	"Higher education"	Title or abstract OR subject term
Årstal	2007 - 2017	Search limitation
Publikationstype	Peer-reviewed journals	Search limitation

Følgende databaser blev anvendt til fremsøgning af studier:

- EBSCO HOST: ERIC (The Education Resource Information Center), CINAHL (Cumulative Index to Nursing and Allied Health Literature), British Education Index, Education Research Complete, Communication & Mass Media Complete
- ProQuest: Australian Education Index, Education Database, PsycInfo
- Web of Science: Social Science Citation Index
- Scopus

#### *Screeningsprocessen*

Review-processen bestod af fem faser:

1. Databasescreening: Identifikation af større mængde referencer, der med afsæt i søgestrategien potentielt kan inkluderes i undersøgelsen.
2. Screening af referencer for relevans på abstract- og titel-niveau.
3. Screening af referencerne for relevans på fuldtekst-niveau.
4. Dataekstraktion: beskrivelse af de inkluderede studier (ophav og udvalgte karakteristika).
5. Kvalitetsbedømmelse: vurdering af studiernes troværdighed og kvalitet.

Alle screeninger og forhandlinger skete med afsæt i de opstillede inklusions- og eksklusionskriterier for nærværende undersøgelse. Fase 2 - 5 blev gennemført ved brug af review-systemet Covidence. Alle referencer blev vurderet af artiklens forfattere i teams. Uenighed om inddragelse/udelukkelse af en undersøgelse blev løst ved konsensus og diskussion med samtlige reviewere.

#### *Dataekstraktion og kvalitetsvurdering*

Følgende data blev udtrukket fra studierne i forbindelse med dataekstraktion: forfattere, publiceringsår, tidsskriftets impact-factor, karakteristika for studiet, herunder: geografi, uddannelsesniveau, undersøgelsesdesign, antallet af deltagere i undersøgelsen, anvendt mål for frafald, anvendt mål for læringsteknologi, sammenligningsgrundlaget i studiet og resultater.

Vi opstillede en tjekliste til vurdering af studiernes kvalitet. Skemaet blev udviklet iterativt med afsæt i guidelines for afrapportering af empiriske studier (Mager & Nowak, 2012; Savin-Baden & Major, 2010; Tong, Sainsbury & Craig, 2007) og testet inden endelig brug. Følgende foci blev operationaliseret i tjeklisten: forskningsspørgsmål, forskningsdesign, dataindsamling og dataanalyse, resultater, perspektivering og etiske overvejelser. Skemaet rummede derudover to spørgsmål, der var specifikke for denne undersøgelse og tilknyttet vurdering af

studiernes metodebeskrivelse: 1. "Er beskrivelsen af teknologien klart beskrevet og relevant i forhold til formålet med denne?" og 2. "Er målingen af frafald klart beskrevet og berettiget?" De enkelte punkter i tjeklisten kunne tildeles 0 (ikke til stede), 1 (i nogen grad) eller 2 (til stede) point. Artiklerne kunne maksimalt opnå 40 point. Studier, der ikke opnåede (1) *en samlet pointscore  $\geq 30$  (ud af 40 mulige)* og (2) *mindst 8 point samlet* i de spørgsmål, der angik studiets metodebeskrivelse, blev ekskluderet.

#### Datasyntese

Målet med datasyntesen var at samle resultaterne fra forskellige typer af forskning for herigennem at kunne beskrive forhold for brug af læringsteknologi, der hhv. påvirker eller ikke påvirker studerendes frafald. For at syntetisere de forskellige studier udviklede vi temaer med afsæt i vores dataekstraktion. Temaerne opstod ved at identificere studierne operationalisering af læringsteknologi og de universitetspædagogiske eksperimenter i de studier, hvor læringsteknologi understøttede et sådant. Temaerne blev diskuteret blandt reviewerne og efterfølgende tilpasset.

#### Resultater

I det følgende sammenfattes resultaterne af de studier, der opnåede en score i kvalitetsvurderingen på minimum 30 point ud af 40 mulige og med mindst 8 point opnået i metodede- len. En typisk årsag til frasortering var en lav score i resultat- og/eller diskussionsdelen. Studier, som indgik i kvalitetsvurderingen, er beskrevet i [Appendix A](#), hvori de endeligt inkluderede studier er markeret med en asterisk (\*).

#### Inkluderede studier

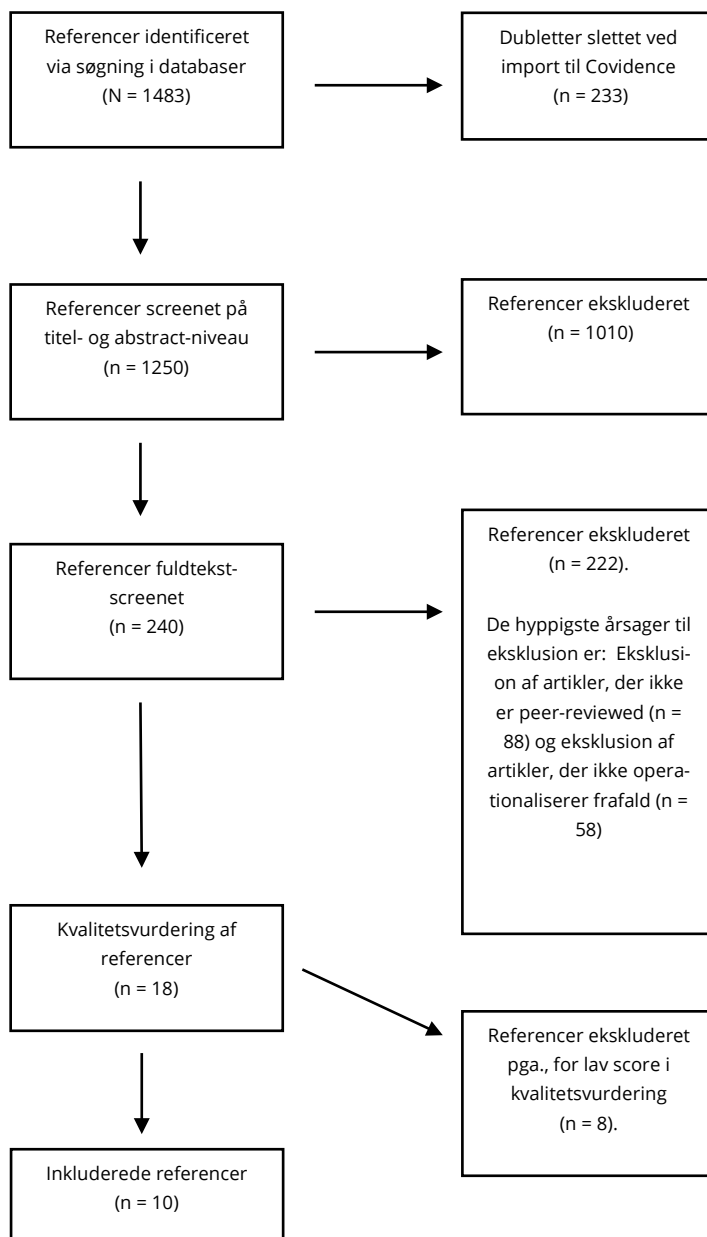
Fremsøgning af studier blev gennemført i september 2017. Den indledende litteratursøgning i databaserne resulterede i 1.483 referencer, der blev importeret til Covidence (Tabel 3).

Tabel 3: Resultat af databasesøgning

Dato	Søgestrategi	Database	Hits
25.9. 2017	Søgning på title, abstract og subject terms	EBSCOhost	774
do	Søgning på title, abstract og subject headings	Proquest	399
do	Søgning på title og topic	Web of science	78
do	Søgning på title, abstract, keywords	Scopus	366
Hits i alt			1617
Antal referencer importeret til Covidence	Baseret på databasernes egen reduktion af hits (fjernelse af dubletter, forkerte publikationstyper m.v.)		1483

240 referencer blev vurderet på fuldtekstniveau, og ud af disse blev 10 referencer vurderet som kvalificerede til at indgå i den endelige analyse (Figur 1).

Figur 1. Flowdiagram



De inkluderede studier dækker et bredt uddannelsesfagligt spektrum og stammer fortrinsvis fra amerikanske universiteter (n=8). De resterende studier (n=2) er gennemført i en hhv. engelsk og australsk kontekst. Størstedelen af studierne omhandler frafald blandt bachelorstuderende.

### **Operationalisering af læringsteknologi**

Vi fandt, at de inkluderede studier i særlig grad interesserer sig for campusbaseret undervisning (n=6) og onlineundervisning (n=6), dernæst blended learning-kurser (n=4) og teknologi-understøttet undervisning (n=3) (se Tabel 4). Studierne er heterogene i forhold til deres operationalisering af læringsteknologi og undersøgelsesdesign. Således undersøges effekten af anvendelse af læringsteknologi inden for (a) et kursus i et format, (b) et kursus i forskellige

formater og (c) forskellige kurser i forskellige formater. Tabel 4 beskriver reviewets definition af leveringsformater og fordelingen af studierne i forhold til de leveringsformater, der indgår i undersøgelserne. På grund af reviewets overvægt af amerikanske studier er definitionerne perspektiveret til en amerikansk praksis (Online Learning Consortium, 2015), hvor leveringsformater defineres ud fra omfanget af kursusaktiviteter, der foregår online (versus på campus).

*Tabel 4. Leveringsformater, der indgår i reviewet (bearbejdet efter Online Learning Consortium, 2015)*

Leveringsformat	Beskrivelse	Forkortelse i reviewet	Studier, der behandler formatet*
Campusbaseret undervisning	Undervisningen er organiseret som face-to-face-møder uden brug af læringsteknologi, og hvor læringsteknologi primært understøtter mere administrative opgaver.	F2F	Ashby, Sadera & McNary (2011); Hughes (2007); Liu & Stengel (2011); Powers, Brooks, Galazyn & Donnelly (2015); Wladis, Conway & Hachey (2017); Xu & Jaggars (2011)
Teknologi-støttet undervisning	Online kursusaktiviteter supplerer campusbaseret undervisning uden at reducere antallet af face-to-face-møder.	TEK	Garratt-Reed, Roberts & Heritage (2016); Griff & Matter (2013); Liu & Stengel (2011)
Blended (eller hybrid) learning	Online kursusaktiviteter er kombineret med campusbaseret undervisning, og de erstatter en væsentlig del af face-to-face-møderne. I amerikansk kontekst har man hidtil typisk defineret et kursus med op til 79 % online kursusaktiviteter som blended learning.	BL	Ashby et al. (2011); Hughes (2007); Powers et al. (2015); Wladis et al. (2017)
Onlineundervisning	Størstedelen eller al undervisning foregår online. I amerikansk kontekst defineres et kursus typisk som online, hvis 80 % eller derover af undervisningen foregår online.	O	Ashby et al. (2011); Garratt-Reed et al. (2016); Pittenger & Doering (2010); Robb & Sutton (2014); Wladis et al. (2017); Xu and Jaggars (2011)

\*Et studie kan behandle flere formater. Identifikation af de leveringsformater, der indgår i studierne, bygger dels på studierne egen beskrivelse heraf og dels på vores vurdering af det beskrevne leveringsformat jf. vores definition.

#### *Studiernes definition af frafald*

Definitionen af frafald varierer i de inkluderede studier. Størsteparten af studierne undersøger frafald på kursusniveau. Frafald på kursusniveau defineres bl.a. i forhold til (1) andelen af studerende der gennemfører hele kurser eller moduler i et kursus (f.eks. Liu & Stengel, 2011), (2) andelen af studerende der deltager i eksamen (f.eks. Inkelaar & Simpson, 2015), (3) de studerendes faglige præstationer, f.eks. andelen af studerende der opnår karakteren lavere end "C" (Wladis et al., 2017; Xu & Jaggars, 2011) eller (4) andelen af studerende der ikke opnår 70 % rigtige eksamenssvar (f.eks. Ashby et al., 2011).

#### *Syntese af inkluderede studier*

Vi identificerede fire temaer i de inkluderede studier: Hhv. studier, der undersøger: (1) de leveringsformater der er opstået i tilknytning til brug af læringsteknologi, (2) læringsteknologi, der faciliterer aktiverende F2F-undervisning, (3) læringsteknologi, der anvendes til adaptiv og automatiseret instruktion og (4) tilrettelæggelse af teknologiunderstøttede forløb med fokus på motiverende kommunikation og kursusdesign (se Tabel 5).

*Tabel 5: Temaer og studierne resultater*

Tema	Forfatter	Studiet	Resultat
<b>Leveringsformater</b>	Ashby et al. (2011)	Sammenligner tre leveringsformater F2F/BL/O inden for samme kursus i relation til antal studerende, der <u>består eksamen</u> .	Der ses ingen forskelle
	Garratt-Reed et al. (2016)	Sammenligner et O-forløb med et TEK-forløb i relation til antallet af studerende, der <u>består eksamen</u> .	TEK > O
	Xu & Jaggars (2011)	Sammenligner <u>frafald</u> mellem to formater F2F/O på to fag.	F2F < O
	Wladis et al. (2017)	Sammenligner A: Overordnet forskelle i <u>frafald</u> i F2F over for O/BL.  Underordnet B: Valgfrie/Obligatoriske C: Lower/upper level kurser	A: F2F < O og BL B: F2F < O og BL C: O < F2F
<b>Aktiverende F2F-undervisning</b>	Liu & Stengel (2011)	Sammenligner <u>frafald</u> på undervisningsforløb, der hhv. inddrager (TEK) eller ikke inddrager clickers (F2F).	TEK < F2F

<b>Adaptiv og automatiseret instruktion</b>	Griff & Matter (2013)	Sammenligner forskelle i <u>frafald</u> og <u>faglig præstation</u> mellem to formater (adaptivt, automatiseret instruktion vs. underviserudvalgte spørgsmål fra online spørgsmålsbank).	Ingen signifikante forskelle.
	Powers et al. (2015)	Sammenligner <u>frafald</u> mellem to formater: F2F vs. BL (med brug af adaptivt, automatiseret instruktion)	Ingen signifikante forskelle.
<b>Motiverende kommunikation og kursusdesign</b>	Pittenger & Doering (2010)	Undersøger, hvorvidt e-mails til stud. med opfordring til aktiv deltagelse, læringsaktiviteter med fokus på aktiv læring og fleksibilitet i forhold til gennemførelse af opgaver på et O-forløb kan bidrage til at <u>reducere frafald</u> .	≥ 95 % studerende gennemfører kurset.  Det er ikke muligt at gennemskue, om resultatet alene kan henføres til de kommunikative aspekter.
	Robb & Sutton (2014)	Sammenligner <u>frafald</u> mellem to grupper af studerende på et O-forløb:  Gruppe 1 modtager fem e-mails henover semestret (ikke personaliserede) med rådgivning om progression, påmindelser om at gennemse feedback, anerkendelse af arbejdsindsats og opfordring til at gennemføre kurset.  Gruppe 2: Modtager ikke e-mails.	Frafald gruppe 1 < gruppe 2.
	Hughes (2007)	Sammenligner <u>frafald</u> mellem to BL-kurser, hvor der på det ene kursus ydes en særlig proaktiv støtte til frafaldstruede studerende i form af ekstra vejledning og opmuntrende kommunikation.	Frafaldet er lavere på det kursus, der indeholder løbende support (6 % vs. 17 %).

*Tema: Leveringsformater*

Med introduktionen af nye formater for levering af kurser, såsom rene onlinekurser, er der opstået en interesse for at undersøge virkningen af de nye leveringsformater over for det traditionelle campusbaserede leveringsformat (Means, Toyama, Murphy, Bakia & Jones, 2010). I dette review indgår fire studier, der undersøger hvorvidt leveringsformatet indvirker på frafald.

Ashby et al. (2011) finder ingen signifikante forskelle på andelen af community college-studerende, der inden for hvert af studiets tre undersøgte leveringsformater består kurset (F2F=59 %; BL=48 %; O=65 %). Derimod finder man, at en større andel F2F-studerende gennemfører kursusforløbet, der leder op til den afsluttende eksamen (F2F=93 %; BL=70 %; O=76 %). Forfatterne påpeger, at den lavere kursusgennemførelse for hhv. BL- og O-studerende kan skyldes, at formaterne ikke appellerer til de studerende, men også at de faglige forudsætninger blandt studerende på community colleges generelt er lidt lavere end blandt studerende på øvrige colleges eller universiteter.

Garratt-Reed og kolleger (2016) finder, at forskellen på antallet af studerende, der bestod et TEK-forløb, var signifikant højere (96 %) end de, der bestod et O-forløb (91 %) ( $p = 0.018$ ,  $d = 0.47$ ). Ved en nærmere undersøgelse af eksamensresultatet finder de, at resultatet særligt skyldes, at de, der deltog i TEK-forløbet, opnåede signifikant bedre resultater i forbindelse med gruppeopgaver end de, der deltog i O-forløbet.

I Xu og Jaggars (2011) undersøgelse sammenlignes et F2F-format med et O-format i to introduktionsforløb (hhv. engelsk og matematik) ved brug af registerdata fra mere end 20 forskellige uddannelsesinstitutioner med lokale udgaver af de specifikke undervisningsforløb. Forskerne finder, at frafaldet er signifikant højere blandt studerende, der deltager i rene online forløb. På engelskkurset udgør frafaldet blandt O-studerende 19 % mod 10 % blandt F2F-studerende. På matematik er frafaldet på 25 % blandt O-studerende mod 12 % blandt F2F-studerende. Forfatterne argumenterer for, at denne forskel først og fremmest skyldes forskelle i studenterforudsætninger/karakteristika fremfor leveringsformatet.

Wladis et al. (2017) undersøger ligeledes betydningen af leveringsformatet (O/BL/F2F) for at nuancere, hvorvidt et O- eller BL-format er mere velegnet til bestemte kursustyper end et F2F-format. Omvendt, undersøger forskerne tillige betydningen af, om kurset er valgfrit, 'distributional' eller obligatorisk, kursets sværhedsgrad og den faglige disciplin (STEM vs. ikke-STEM). Samme kursus udbydes i forskellige formater (O/BL/F2F) i samme semester af samme underviser og gentages i en længere årrække (2004-2010). Resultaterne viser for det første, at der i perioden generelt er en bedre fastholdelse af de studerende i F2F-formatet (65,3 %) sammenlignet med O/BL-formatet (58,6 %), men at der er betydelige nuancer. På valgfrie kurser ses et signifikant højere frafald i O/BL-formatet sammenlignet med F2F-formatet over for de samme formater i obligatoriske kurser. Til gengæld gennemfører de studerende i højere grad et Lower Level-kursus online (49,9 %) end F2F (42,2 %), mens billedet er omvendt ved Upper Level-kurser. Generelt er den gennemsnitlige kursusgennemførelse signifikant bedre for Upper Level-kurser i forhold til Lower Level-kurser i både F2F-format og O/BL-format.



### *Tema: Aktiverende F2F-undervisning*

Gennem de seneste årtier har der været øget fokus på at udvikle universitetsundervisningen fra transmission af viden til mere aktiverende undervisning (Biggs & Tang, 2011). Ud af de 10 studier, der indgår i syntesen, undersøger ét, hvorvidt undervisning, der fokuserer på at aktivere de studerende i timerne, kan bidrage til at reducere frafald.

Liu & Stengel (2011) sammenligner frafald i kurser, der hhv. anvender (TEK)/ikke anvender clickers (F2F) til at løse multiple choice-opgaver præsenteret af underviseren på campus. I TEK-forløbet drøftes løsninger med underviseren. Den traditionelle F2F-udgave af kurserne består af forelæsninger ved samme underviser og tilsvarende øvelser i timerne - blot uden afrapportering og feedback. Liu og Stengel finder, at frafaldet er lavere i TEK-forløbet sammenlignet med F2F-formatet. Desuden opnår de studerende bedre resultater i de løbende tests fraset et enkelt eksempel, hvor studerende, der deltager i F2F-formatet, opnår bedre resultater. Forfatterne anfører, at årsagen til dette kan skyldes, at nogle af emnerne i den pågældende test er mere komplekse og muligvis ikke egner sig til det anvendte multiple choice-format.

### *Tema: Adaptiv og automatiseret instruktion*

Automatiseret instruktion eller adaptive læringsystemer vurderer og guider studerende igennem typisk mindre dele af et fagområde online (Willis et al., 2018, s. 50). Med afsæt i den studerendes besvarelser på opgaver tilpasses instruktionen den studerendes niveau, og herved kan den studerende opleve at få et individualiseret og 'just-in-time'-læringsforløb. Hensigten med denne type instruktion er typisk ikke at erstatte underviseren, men at tilbyde studerende et supplement til den eksisterende undervisning. Materialet er ofte udviklet af en ekstern part som f.eks. et forlag, der i forvejen udgiver centrale lærebøger inden for fagområdet. I dette review angår to af de inkluderede studier automatiseret instruktion.

I Griff & Matters (2013) studie med deltagelse af 6 institutioner undersøges effekten af et adaptivt læringsystem udviklet til bachelorstuderende inden for anatomi eller fysiologi. Indholdet, som de studerende skal arbejde med, udvælges af instruktorer. Kontrolgruppen får spørgsmål fra en spørgsmålsbank, ligeledes udvalgt af en instruktør. Resultaterne viser, at begge grupper af studerende forøger deres viden, og der ses ingen signifikant forskel på de to grupper i forhold til præstation eller frafald. Derimod finder forskerne, at der på to af de involverede institutioner er en forskel på eksperiment- og kontrolgruppen. Undersøgelsen kan ikke afdække årsagen til denne forskel. Forfatterne konkluderer, at nærværende studie ikke understøtter andre studiers fund af effekten af adaptive materialer, og de vurderer, at det i dette tilfælde muligvis kan skyldes instruktorens indflydelse.

I studiet af Powers et al. (2016) sammenlignes frafald for to grupper af studerende, der hhv. deltager i et F2F-forløb og i et BL-forløb, hvori der indgår et adaptivt læringselement (MyPsychLab). De to grupper gennemgår det samme faglige indhold, og undervisningsforløbene afvikles parallelt. Resultaterne af denne undersøgelse viser ingen forskelle i frafald mellem de to grupper. Frafallsprocenten udgjorde 16,8 % for BL-studerende og 14,8 % for F2F-studerende. Forfatterne oplyser endvidere, at ingen af de studerende, der deltog i BL-forløbet, og som faldt fra, havde besøgt MyPsychLab. Selvom undersøgelsen ikke kan udsige noget om de studerendes holdning til onlineundervisningselementerne, indikerer forskerne alligevel, at forskellen kan være et udtryk for, at denne gruppe ikke var motiverede af mulighederne og kravene til online selvstudie. De konkluderer derpå, at studerendes motivation i

forhold til at arbejde med online læringsmateriale kan udfordre antagelsen om, at adaptive undervisningsformater kan erstatte traditionel undervisning.

*Tema: Motiverende kommunikation og kursusdesign*

Selvom læring i høj grad afhænger af studerendes viden og evnen til at anvende denne, har forskning også vist, at motivation og interesse fremmer de studerendes læreprocesser (Bandura, 1993). Tre af de inkluderede studier undersøger effekten af motiverende elementer i undervisningen. To af disse studier omhandler O-forløb tilrettelagt som selvstudie, mens det tredje studie omhandler et BL-forløb.

Robb & Sutton (2014) sammenligner to grupper i et O-forløb, hvoraf en gruppe af studerende modtager fem motiverende og ikke personaliserede e-mails. Disse indeholder blandt andet: rådgivning om progression, påmindelser om at gennemse feedback, anerkendelse af arbejdsindsats og opfordring til at gennemføre kurset. Den anden gruppe modtager ingen mails, men har adgang til de samme informationer på en kursusside. Frafaldet blandt de studerende i gruppen, der modtager e-mails, er lavere sammenlignet med gruppen, der ikke modtager e-mails.

Pittenger & Doering (2010) undersøger studerendes oplevelse af et O-forløb, der er kendetegnet ved en høj gennemførelsesprocent ( $\geq 95\%$ ). Deres komparative analyse af fire O-kurser indikerer, at en velorganiseret kursusstruktur, ugentlige e-mails med forslag til deltagelse i kursets aktiviteter og opgaver, læringsaktiviteter med fokus på aktiv læring og fleksibilitet i forhold til gennemførelse af opgaver kan bidrage til at reducere frafald. Blandt andet rapporterer de studerende, at designet af de fire kurser indeholder en undervisningsmæssig stilladsering, der fanger deres opmærksomhed, opleves som relevant for dem og giver dem tillid til, at de kan gennemføre kurset. Overordnet set resulterer det i en høj grad af tilfredshed. Det kursus, som de studerende giver den laveste score mht. motiverende elementer, adskiller sig fra de øvrige ved ikke at have en tekstbog og ved udelukkende at bestå af online læringsmaterialer såsom audiopræsentationer og slide-handouts. Derudover kræver dette kursus deltagelse i fire eksaminer undervejs, i modsætning til de øvrige kurser, der kun kræver deltagelse i den afsluttende eksamen. Forfatterne overvejer, hvorvidt studerende demotiveres af de mange afprøvninger, der trækker på strukturer fra traditionelle F2F-kurser.

I Hughes' (2007) undersøgelse er der fokus på et BL-kursus, hvor 30 % af forelæsningserne er omlagt til online aktiviteter og online vejledning til frafaldstruede studerende. Underviseren følger de studerendes adfærd på en kursusside og giver særlig vejledning (af administrativ, teknisk, motiverende eller faglig karakter) til de studerende, der er mindre aktive. Hughes finder, at kombinationen af proaktiv hjælp og opmuntrende kommunikation øger antallet af studerende, der gennemfører kurset, sammenlignet med en tidligere version af BL-forløbet uden support. Frafaldet er hhv. på 6 % og 17 % i forløbet med support mod 25 % i forløbet uden særlig support til frafaldstruede studerende. Data fra det samme kursus i et F2F-format viser et frafald på 25-55 % på disse kurser. Hughes konkluderer deraf, at et veltilrettelagt BL-forløb med online aktiviteter og online vejledning til frafaldstruede studerende har en positiv betydning i forhold til at øge andelen af studerende, der gennemfører kursusforløbet.

### **Opsamling**

Vi stillede spørgsmålet: I hvilken grad kan læringsteknologi bidrage til at reducere frafald på de videregående uddannelser? Vi har i det foregående givet et overblik over et udsnit af nyere empirisk forskning, der undersøger dette. Som et væsentligt træk fandt vi, at frafald ikke er

entydigt operationaliseret i studierne. Frafald varierer eksempelvis fra manglende deltagelse eller beståelse af kursusmoduler til bestået/ikke bestået eksamen. Derudover fandt vi, at studierne undersøger inddragelsen af læringsteknologi forskelligt: (1) som leveringsformat og (2) med afsæt i undervisningseksperimenter eller kursusdesigns, der på forskellig vis inddrager læringsteknologi.

Den første gruppe, der undersøger leveringsformater, består af fire studier. Heraf finder et studie ingen forskelle mellem de undersøgte leveringsformater. De tre øvrige studier giver samlet set ikke et entydigt billede af betydningen af leveringsformatet på frafald. Resultaterne af studierne indikerer, at kurser med ingen eller nogle online kursusaktiviteter (F2F eller TEK) har lavere frafald blandt studerende i sammenligning med kurser med høj forekomst af online kursusaktiviteter (BL eller O). Det fremgår dog ikke klart af studierne, hvilke udfordringer som studerende eller undervisere oplever ved de forskellige leveringsformater.

I den anden gruppe indgår seks studier, der angår brugen af læringsteknologi forankret i en undervisningspraksis. Fokus i studierne er således ikke på betydningen af læringsteknologi alene, men også på koblingen af pædagogiske valg med læringsteknologi. Fire studier finder en positiv sammenhæng mellem lavere frafald og tiltag i undervisningen, der gennemføres vha. læringsteknologi. To studier finder ingen effekt af teknologiinddragelse.

Det næste spørgsmål er: Under hvilke betingelser kan brugen af læringsteknologi på videregående uddannelser bidrage til at mindske studerendes frafald? Det spørgsmål besvares i højere grad i de studier, der beskriver et pædagogisk rationale og de kontekstuelle forhold for brugen af læringsteknologi. Under afsnittet om aktiverende undervisning så vi, at undervisning med brug af et digitalt afstemningsværktøj til at aktivere studerende kan bidrage positivt til reducere af frafald (Liu & Stengel, 2011). I det konkrete tilfælde fandt man, at flere af de studerende, der anvendte afstemningsværktøjet i undervisningen, gennemførte kursusmodulet sammenlignet med de studerende, der ikke anvendte afstemningsværktøjet. Idéen bag digitale afstemningsværktøjer er at give studerende mulighed for at reflektere over fagets indhold, samtidig med at de får feedback på deres niveau og forståelse af emnet (Dyrberg, 2014). Feedback vurderes at være en betydningsfuld faktor – særligt i forbindelse med forelæsningsformatet, der ofte indebærer mindre interaktion, og hvor det kan være vanskeligt for forelæseren at fornemme, om de studerende forstår det forklarede (Dyrberg, 2014). I afsnittet om adaptiv og automatiseret instruktion gennemgik vi to studier, der undersøger effekten af automatiseret instruktion eller adaptive online læringssystemer, der tilbyder studerende individualiseret instruktion (Griff & Matter, 2013; Powers et al., 2015). Begge studier finder ingen effekt af brugen af adaptiv instruktion, men der er forskelle i de to studier, som er værd at fremhæve. I det ene studie indgik adaptivt materiale som et supplement, og i det andet studie erstattede det F2F-undervisningstid, hvilket i dette tilfælde også indebar reduceret mulighed for social og faglig interaktion mellem medstuderende og undervisere. I sidstnævnte studie (Powers et al., 2015) konkluderede man, at det er nødvendigt at understøtte de studerendes arbejde med materialet og i højere grad at skabe en bedre integration mellem dette materiale og resten af undervisningen for at højne læringsudbyttet og de studerendes motivation. Det peger på betydningen af tilrettelæggelse af undervisningsaktiviteter i undervisningen og betydningen af at sammentænke dette med studerendes nuværende niveau for viden og krav til eksamen. Sidstnævnte blev netop fremhævet som en medierende faktor i det første studie. I syntesen gennemgik vi endvidere tre studier, der peger på, at brugen af motiverende kommunikation og motiverende kursusdesign kan bidrage til at reducere frafald. Dette kan f.eks. være i form af personlige e-mails til de studerende (Robb & Sutton,

2014), kursusplanlægning med fokus på motiverende elementer (Pittenger & Doering, 2010) eller særlig monitorering af og støtte til frafaldstruede studerende (Hughes, 2007).

Samlet set indikerer disse studier, at underviseren spiller en central rolle i forhold til at tilrettelægge kurser og aktiviteter, der ved hjælp af læringsteknologi motiverer og aktiverer studerende.

### Diskussion og forslag til videre forskning

I reviewet har vi gennemgået studier, der undersøger, hvorvidt teknologi i undervisningen kan bidrage til at reducere frafald. Resultaterne viser, at det ikke er muligt at løse frafaldsproblematikker ved blot at tilføje mere læringsteknologi. Ydermere viser de, at der kan være en positiv sammenhæng mellem frafald og teknologiunderstøttet undervisning. Der, hvor vi ser positive fund, er særligt i de studier, hvori der indgår et pædagogisk rationale for brugen af læringsteknologi, f.eks. med henblik på at motivere og/eller aktivere studerende. Her finder vi det særligt interessant, at man selv med en forholdsvis lille indsats kan påvirke studerendes faglige integration og i sidste ende mindske studerendes frafald. Dette så vi blandt andet eksemplificeret i afsnittet om motiverende kommunikation og kursusdesign, hvor direkte og opmuntrende kommunikation kan bidrage til både den faglige og sociale integration. Vi så også, hvordan anvendelsen af et digitalt afstemningsværktøj i forelæsninger kan inddrage de studerende direkte i undervisningen, bringe deres viden i spil og øge deres opmærksomhed. Vi har endvidere set, hvordan læringsteknologier kan anvendes til at give feedback, der løbende kan støtte tilegnelsen af et fags teori, empiri og metode. Eksemplerne illustrer samtidig, at studierne særligt fokuserer på den faglige integration i form af interaktion mellem underviser og studerende eller mellem studerende og stoffet og i mindre grad fokuserer på interaktion de studerende imellem. Hvorvidt frafald kan mindskes via et fokus på det sidste aspekt, giver studierne os ikke svar på.

Til trods for at studierne, der indgår i denne undersøgelse, primært er baseret på en amerikansk universitetskontekst, finder vi, at temaerne, som de beskæftiger sig med, er sammenfaldende med nogle af de temaer, som vi er optaget af i den danske uddannelsessektor og i en dansk universitetspædagogisk kontekst. For eksempel er aktiverende undervisning og aktiv læring de bærende principper for uddannelse på flere af de danske universiteter, herunder blandt andet Syddansk Universitet (Syddansk Universitet, 2018), IT-Universitetet (IT-Universitetet, 2019) og Roskilde Universitet (Roskilde Universitet, 2019). Endvidere beskriver danske studier af undervisningsekspirer med læringsteknologi ofte undervisningsmetoder til aktivering af studerende – herunder de danske studier, som vi indledningsvist refererede til. Endelig anses adaptiv læring og automatiseret instruktion for at rumme så store potentialer for det danske uddannelsesmarked at Statens finansieringsfond, Vækstfonden, i 2018 skød 190 millioner kroner i udviklingen af adaptive læringsmoduler (Vækstfonden, 2018). Det kunne derfor være relevant at undersøge, om de nævnte temaer og konkrete strategier for undervisningsudvikling understøttet af læringsteknologi kan bidrage til at øge studerendes læring og motivation, og i forlængelse af dette bidrage til at reducere frafald i en dansk universitetskontekst.

Studierne savner generelt en beskrivelse af og kobling til kontekst- og studentfaktorer, som vi ved fra anden forskning påvirker de studerendes engagement og faglige tilknytning (se f.eks. Biggs & Tang, 2011; Ellis & Goodyear, 2010; Entwistle, 2007; van Dinther, Dochy & Segers, 2011). Endvidere savnes en kobling til den eksisterende forskning i online læring. Særligt amerikanske og canadiske forskere har beskæftiget sig med *community of inquiry* og

undersøgt betydningen af underviserens tilstedeværelse online og sociale interaktioner (Garrison, Anderson & Archer, 2010). Vi anbefaler derfor, at man i højere grad er opmærksom på den eksisterende forskning og kobler den til fremtidig forskning, der angår brugen af læringsteknologi til reduktion af frafald.

### Styrker og svagheder

I denne undersøgelse har vi gennemført et systematisk review for at gennemgå resultaterne af international, peer-reviewed forskning, der undersøger brugen af læringsteknologi og sammenholder det med tal for frafald. Et systematisk review bør, hvis det er veludført, spare andre for at gennemføre lignende litteratursøgninger, henvise til interessante studier af høj kvalitet og præsentere en syntese af resultaterne inden for området (Bearman et al., 2012). 10 ud af 1.483 studier angik undersøgelsens fokus. I [Appendix A](#) har vi beskrevet samtlige kvalitetsvurderede studier for at give læseren mulighed for at vurdere denne undersøgelses kvalitet såvel som kvaliteten af de studier, der indgår i analysen. Vi har syntetiseret resultaterne af de studier, der levede op til etablerede standarder for forskningskvalitet (10/18). Vi har syntetiseret studierne, for at andre aktører inden for området så vidt muligt kan anvende resultaterne i egen undervisning eller strategi- og udviklingsarbejde. Det lave antal inkluderede studier giver dog anledning til at være forsigtig med at generalisere resultaterne. Tilsvarende skal man være forsigtig med at overføre resultaterne direkte til en dansk kontekst, da hovedparten af studierne i dette review er foretaget i en amerikansk universitetskontekst.

Ved hjælp af systematiske review er det muligt at kondensere en stor del af den viden, der findes inden for et felt. Undersøgelsen her er dog af flere grunde begrænset i forhold til at blotlægge alle undersøgelser af frafald inden for de videregående uddannelser, hvor lærings-teknologi er blevet inddraget. Undersøgelsens søgestrategi er snæver. Den ekskluderer eksempelvis de studier, der ikke er identificeret ved hhv. frafald, læringsteknologi og 'higher education'. På samme vis er der ikke gennemført en supplerende screening af referencer, hvilket kunne have sikret identifikation af andre relevante studier. Selvom sprogrestrictioner ikke er ideelle, valgte vi i dette review at begrænse vores søgning til studier, der er publiceret i engelsksprogede peer-reviewede tidsskrifter. Derved kan viden fra undersøgelser publiceret på dansk eller andre sprog være gået tabt. Undersøgelsen repræsenterer således kun et udsnit af forskningen inden for området, og der vil i fremtiden derfor være behov for undersøgelser, der anvender andre søgestrategier.

### Referencer

- Aarhus Universitet. (2018). Aarhus Universitet investerer i teknologisk udvikling af undervisningen. Hentet fra <http://newsroom.au.dk/nyheder/vis/artikel/aarhus-universitet-investerer-i-teknologisk-udvikling-af-undervisningen/>
- Aarhus Universitet, & Uddannelses- og Forskningsministeriet. (2018). *Strategisk rammekontrakt 2018-2021* Hentet fra [https://www.ufm.dk/uddannelse/videregaende-uddannelse/universiteter/styring-og-ansvar/strategiske-rammekontrakter/au\\_strategisk-rammekontrakt-2018-21.pdf](https://www.ufm.dk/uddannelse/videregaende-uddannelse/universiteter/styring-og-ansvar/strategiske-rammekontrakter/au_strategisk-rammekontrakt-2018-21.pdf)
- Ashby, J., Sadera, W. A. & McNary, S. W. (2011). Comparing student success between developmental math courses offered online, blended, and face-to-face. *Journal of Interactive Online Learning*, 10(3), 128-140.

- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational psychologist*, 28(2), 117-148.
- Bearman, M., Smith, C. D., Carbone, A., Slade, S., Baik, C., Hughes-Warrington, M. & Neumann, D. L. (2012). Systematic review methodology in higher education. *Higher Education Research & Development*, 31(5), 625-640.
- Bertel, T. F. & Bentzen, T. Ø. (2018). Formativ evaluering med IT-systemet Peergrade. *Tidsskriftet Læring og Medier (LOM)*, 10(18).
- Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (4 ed.). Milton Keynes, UK: Open University Press.
- Caviglia, F., Dalsgaard, C., Davidsen, J. & Ryberg, T. (2018). *Students' digital learning environments*. Paper presented at the 11th International Networked Learning Conference.
- Copenhagen Business School. (2017). The RiBL project is a 6-year externally funded project. Hentet fra <http://blog.cbs.dk/ribl/the-ribl-project/>
- Danbjørg, D. B., Clemensen, J., Hansen, P. S., Thrysoe, L. & Rothmann, M. J. (2018). Blogs-Læringsfællesskab med rum til refleksion. *Tidsskriftet Læring og Medier (LOM)*, 10(18).
- Davidsen, J. & Ryberg, T. (2016). Samhørighed, interaktion og vidensdeling blandt studerende-erfaringer fra et IKT-pædagogisk udviklingsprojekt. *Dansk Universitetspædagogisk Tidsskrift*, 11(21), 57-71.
- Dørup, J., Gomme, J., Hansen, A. & Heiberg, B. (2005). Implementering af e-læring ved danske universiteter. *Tidsskrift for Universiteternes Efter-og Videreuddannelse (UNEV)*, 3(6).
- Dyrberg, N. R. (2014). Clickers-forbedring af traditionelle forelæsninger? *MONA-Matematik-og Naturfagsdidaktik*(2), 22-41.
- Ellis, R. & Goodyear, P. (2010). *Students' experiences of e-learning in higher education: the ecology of sustainable innovation*: Routledge.
- Entwistle, N. (2007). *Research into student learning and university teaching* (Vol. 1): British Psychological Society.
- Garratt-Reed, D., Roberts, L. D. & Heritage, B. (2016). Grades, Student Satisfaction and Retention in Online and Face-to-Face Introductory Psychology Units: A Test of Equivalency Theory. *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.00673
- Garrison, D. R., Anderson, T. & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The internet and higher education*, 13(1-2), 5-9.
- Griff, E. R. & Matter, S. F. (2013). Evaluation of an adaptive online learning system. *British Journal of Educational Technology*, 44(1), 170-176. doi:10.1111/j.1467-8535.2012.01300.x
- Henderson, M., Selwyn, N. & Aston, R. (2017). What works and why? Student perceptions of 'useful'digital technology in university teaching and learning. *Studies in higher education*, 42(8), 1567-1579.

- Henrie, C. R., Halverson, L. R. & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36-53.
- Hughes, G. (2007). Using Blended Learning to Increase Learner Support and Improve Retention. *Teaching in Higher Education*, 12(3), 349-363.
- Inkelaar, T. & Simpson, O. (2015). Challenging the 'distance education deficit' through 'motivational emails'. *Open Learning*, 30(2), 152-163.  
doi:10.1080/02680513.2015.1055718
- IT-Universitetet. (2019). Pædagogiske Principper på IT-Universitetet i København. Hentet fra <https://www.itu.dk/om-itu/paedagogiske-principper>
- Jensen, M. B., Lindorf, M., Norn, M. T. & Aller, S. H. (2015). *Styrk uddannelseskvaliteten gennem digitale læringsteknologier*. Hentet fra [https://dea.nu/sites/dea.nu/files/dea\\_-\\_styrk\\_uddannelseskvaliteten\\_gennem\\_digitale\\_laeringsteknologier.pdf](https://dea.nu/sites/dea.nu/files/dea_-_styrk_uddannelseskvaliteten_gennem_digitale_laeringsteknologier.pdf)
- Littell, J. H., Corcoran, J. & Pillai, V. (2008). *Systematic reviews and meta-analysis*: Oxford University Press.
- Liu, W. C. & Stengel, D. N. (2011). Improving Student Retention and Performance in Quantitative Courses Using Clickers. *International Journal for Technology in Mathematics Education*, 18(1), 51-58.
- Mager, U. & Nowak, P. (2012). Effects of student participation in decision making at school. A systematic review and synthesis of empirical research. *Educational research review*, 7(1), 38-61.
- Means, B., Toyama, Y., Murphy, R., Bakia, M. & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning. A Meta-Analysis and Review of Online Learning Studies* Hentet fra <https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>
- Møller, R. B. (2016). Introduktion til diskussionsforum som læringsplatform. *Dansk Universitetspædagogisk Tidsskrift*, 11(21), 158-165.
- Online Learning Consortium. (2015). Updated E-Learning Definitions. Hentet fra <https://onlinelearningconsortium.org/updated-e-learning-definitions-2>
- Pittenger, A. & Doering, A. (2010). Influence of motivational design on completion rates in online self-study pharmacy-content courses. *Distance Education*, 31(3), 275-293.  
doi:10.1080/01587919.2010.513953
- Powers, K. L., Brooks, P. J., Galazyn, M. & Donnelly, S. (2015). Testing the Efficacy of MyPsychLab to Replace Traditional Instruction in a Hybrid Course. *Psychology Learning and Teaching*, 15(1), 6-30. doi:10.1177/1475725716636514
- Price, L. & Kirkwood, A. (2011). *Enhancing learning and teaching through technology*. York, UK: Higher Education Academy.
- Robb, C. A. & Sutton, J. (2014). The importance of social presence and motivation in distance learning. *Journal of Technology, Management, and Applied Engineering*, 31(2), 1-10.

- Roskilde Universitet. (2017). RUC får 20 mio. til virtuelle læringsteknologier. Hentet fra <https://ruc.dk/nyheder/ruc-far-20-mio-til-virtuelle-laeringsteknologier>
- Roskilde Universitet. (2019). Pædagogisk profil på Roskilde Universitet. Hentet fra [https://intra.ruc.dk/fileadmin/assets/adm/dokumenter\\_og\\_noegletal/vaerdier/Paedagogisk\\_profil\\_DK.pdf](https://intra.ruc.dk/fileadmin/assets/adm/dokumenter_og_noegletal/vaerdier/Paedagogisk_profil_DK.pdf)
- Savin-Baden, M. & Major, C. H. (2010). *New approaches to qualitative research: Wisdom and uncertainty*. Routledge.
- Styrelsen for Forskning og Uddannelse. (2018). *Frafald og studieskift. Spørgeskemaundersøgelse blandt frafaldne studerende ved videregående uddannelser*. Hentet fra København: <https://ufm.dk/publikationer/2018/filer/frafald-og-studieskift.pdf>
- Syddansk Universitet. (2018). Bærende principper. Retrieved from [https://www.sdu.dk/da/om\\_sdu/institutter\\_centre/c\\_unipaedagogik/baerende\\_principper](https://www.sdu.dk/da/om_sdu/institutter_centre/c_unipaedagogik/baerende_principper)
- Syddansk Universitet, & Uddannelses- og Forskningsministeriet. (2018). *Strategisk rammekontrakt 2018-2021*. Hentet fra [https://ufm.dk/uddannelse/videregaende-uddannelse/universiteter/styring-og-ansvar/strategiske-rammekontrakter/sdu\\_strategisk-rammekontrakt-2018-21.pdf](https://ufm.dk/uddannelse/videregaende-uddannelse/universiteter/styring-og-ansvar/strategiske-rammekontrakter/sdu_strategisk-rammekontrakt-2018-21.pdf)
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research*, 45(1), 89-125.
- Tong, A., Sainsbury, P. & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International journal for quality in health care*, 19(6), 349-357.
- Troelsen, R. (2011). Frøfald på de videregående uddannelser–hvad ved vi om årsagerne? *Dansk Universitetspædagogisk Tidsskrift*, 6(10), 37-44.
- Uddannelses- og Forskningsministeriet. (2018). Frøfald. Tal, analyser og undersøgelser om frøfald på videregående uddannelser, som løbende bliver udarbejdet af ministeriet. Hentet fra <https://ufm.dk/uddannelse/statistik-og-analyser/frafald>
- Ulriksen, L., Madsen, L. M. & Holmegaard, H. T. (2010). What do we know about explanations for drop out/opt out among young people from STM higher education programmes? *Studies in Science Education*, 46(2), 209.
- Vækstfonden. (2018). Ny investering: Area9 kan blive fyrtårn inden for uddannelsesteknologi. Hentet fra <https://www.vf.dk/nyheder-og-analyser/nyheder/2018/20180110-area9-faar-investering.aspx>
- van Dinther, M., Dochy, F. & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational research review*, 6(2), 95-108.
- Willis, R., Yeigh, T., Lynch, D., Smith, R., Provost, S., Sell, K. & Turner, D. (2017). *Towards a Strategic Blend in Education: A review of the blended learning literature*. Lulu. com.
- Wladis, C., Conway, K. & Hachey, A. C. (2017). Using course-level factors as predictors of online course outcomes: a multi-level analysis at a US urban community college. *Studies in higher education*, 42(1), 184-200. doi:10.1080/03075079.2015.1045478



Xu, D. & Jaggars, S. S. (2011). The Effectiveness of Distance Education across Virginia's Community Colleges: Evidence from Introductory College-Level Math and English Courses. *Educational Evaluation and Policy Analysis*, 33(3), 360-377.

# Takeaway Teaching – A design for redesign

Hanne Balsby Thingholm<sup>1,a</sup>, Anders Hjortskov Larsen<sup>b</sup>

<sup>a,b</sup>Center for undervisningsudvikling og digitale medier, Aarhus Universitet

Practice paper, peer-reviewed

This paper introduces a new teaching design called Takeaway Teaching. The design consists of pre-designed generic themes and activities that support the development of study strategies in higher education. These themes and activities are easily redesigned by the teacher and integrated into the academic curriculum. The purpose of the design is twofold. It supports the didactic reflection and development of the teacher and the development of the students' study competence. This paper presents the theoretical framework of the design, based on a theoretical approach of metacognition and second-order teaching and learning, and an example of a redesign of one of the generic themes; Literature Search.

## Takeaway teaching – A design for redesign

This paper is a practical contribution to university teaching and presents a new teaching design called *Takeaway Teaching*. The purpose of this design is, by means of pre-designed generic themes, to make it easy for university teachers to integrate activities into the academic curriculum and explicitly teach study strategies. This design is an alternative to generic courses in study skills. We think that an alternative to generic study courses is important for two reasons. Students who need help to develop study competence are often not motivated to sign up and attend the courses. Furthermore, generic courses often lead to a transfer problem for the students and they find it hard to integrate the skills they have learned into their academic discipline.

The Takeaway Teaching design and its theoretical framework have a strong student-oriented perspective, and support the idea that teachers may teach students how to actively learn course content, rather than the idea that the primary objective is for teachers to teach content. This perspective includes a focus on *how* students study, and suggests that study strategies be addressed as an integral part of the curriculum, rather than as something that the students must figure out on their own. The design consists of the following themes: Academic Reading, Academic Writing, Academic Oral Presentation, Peer-feedback, Group Work, MA Preparation, BA Preparation, Literature Search and a Project-oriented theme. Currently the design consists of 13 pre-designed mini-plans or scripts, which are available from the digital resource AU Educate: <http://educate.au.dk/forloeb/>. The themes of the Takeaway Teaching design inspire university teachers to implement various study strategies into their curriculum and teaching. The themes and their activities may be implemented directly or redesigned by the teacher in order to match the students and the subject area. When redesigning the theme, the teacher may choose to remove and add elements and activities to it, in order to qualify the students' achievements with respect to a given academic curriculum. In this way,

---

<sup>1</sup> Contact: hbt@tdm.au.dk

the design of Takeaway Teaching is flexible, and aims to develop teaching competence as well as study competence.

### Theoretical framework of Takeaway Teaching

The theoretical framework of the Takeaway Teaching design is inspired by theories of system theory and metacognition. This combination emphasizes communication and suggests directing didactic attention towards teaching students how to study. In the following sections, we present the theoretical concepts underlying the Takeaway Teaching design: metacognition and second-order teaching. When combining these concepts, it becomes meaningful to talk about metacommunication.

#### *Metacognition in teaching and learning*

Theories about metacognition have been related to teaching and learning since the late 1970s, when John Flavell defined metacognition as: “knowledge and cognition about cognitive phenomena” (Flavell, 1979 p. 906), or put more straightforwardly, “thinking about thinking.” According to Flavell, metacognitive knowledge has three aspects: knowledge about yourself as a *person*, knowledge about the *task*, and knowledge about the *strategies* appropriate for solving the task (Flavell, 1979 p. 907). Since Flavell’s definition of metacognition, the concept has been defined and explored in various ways. However, the literature for the field of metacognition reflects a consensus that distinguishes between two elements: *metacognitive knowledge* and *metacognitive control*. In order to develop metacognitive knowledge and control, three dimensions of knowledge are essential:

1. *Declarative* knowledge about *what* you know about the task in which you are engaged. This relates to the content dimension.
2. *Procedural* knowledge about *how* you can solve the task, and about which strategies to use. This relates to procedure and method.
3. *Conditional knowledge* about *why* and *when* the task is being done. This relates to purpose, state and context.

(Schraw, Crippen, & Hartley; 2006 Kuhn & Dean, 2004).

According to Harrison and Vallin (2017), metacognition is valuable in various ways related to students’ learning processes. From research on collaborative learning, we know that metacognition is important for collective content learning, from motivation research we know that metacognition is important for successful learning strategies, and from literature on critical thinking, we know that metacognition is often regarded as the basis for critical thinking. According to Coutinho (2008), students improve their learning skills when they are trained in metacognitive skills, and metacognition is an important predictor for academic achievement. Having metacognitive skills makes it easier to apply one’s knowledge in different contexts (Pintrich, 2002; Veenman, 2015), and to engage in creative problem-solving (Hargrove & Nietfeld, 2015). According to Frenkel (2014), metacognitive skills are not genetically determined. Some students may have stronger metacognitive skills than others, but everybody can learn metacognitive skills to a certain extent. We agree with the description of metacognition as a “sleeping potential” (Frenkel, 2014).

### *Second-order teaching*

From this perspective, teaching students how to study not only includes metacognitive orientation, it also identifies the importance of second-order teaching. Second-order teaching is a theoretical concept that draws on Luhmann's understanding of second-order observation and Bateson's understanding of context learning. The theoretical analysis of the concept of second-order teaching is articulated by Thingholm and Keiding (2018). In the following sections, we extract the main points of this work by explaining the concepts of second-order observation and context learning. We also describe how, from this perspective, reflection on the implications for teaching emerges as two-sided, considering both the teaching and learning processes.

### *Second-order observation*

According to the German sociologist Niklas Luhmann, an observation is always an interpretation. Consequently, different observers or systems will make different observations or interpretations of *what* they see – a first-order observation. A second-order observation is a certain kind of observation or interpretation that observes the first observation, or in other words, observes *how* something has been observed (Luhmann, 2007, p. 143). Second-order observations observe the immediate, i.e. first, observation. An example is the teacher who finds the students skilful and hardworking. This would be the teacher's first observation, an observation of *what*, of skilful students. When the teachers observe their first observation, it becomes possible to understand *how* they have made distinctions in the first observation. Does the teacher compare the students to last year's students? Does the teacher refer to a test or an exam? Has the teacher been observing the students for a long time, or just one week? (Thingholm and Keiding, 2018).

Second-order observations relate to how an observation or interpretation has been observed or interpreted. Second-order observations aim to understand the distinctions behind the first observation. According to Luhmann, any observation or interpretation is based on a distinction and this distinction is essential to understanding any given question (Luhmann, 2007, p. 144). Therefore, second-order observations are crucial to the teacher's didactic reflections, in order to plan, conduct and evaluate teaching and learning.

In line with second-order observation, which is the observation that observes an observation, second-order teaching is teaching that observes teaching. In order to understand how this is possible, we turn to Luhmann's understanding of teaching and learning. Teaching and learning are closely connected, but at the same time, they are two very different concepts or systems. Teaching is a *social* system maintained by *communication*. Teaching is the interaction that takes place when teachers and students engage in a dialogue. On the other hand, learning is an individual endeavour and refers to the *psychological* system that is maintained by thoughts and *consciousness*. Learning occurs in the mind of each student (and teacher), whereas teaching is a joint effort. When learning takes place in the students' minds, it is not possible for the teacher to *directly* observe the students' learning processes; the teacher cannot access the students' minds. It is only possible to observe the students' learning, including their reflections on their own learning, *indirectly*, in the form of communicative constructions shared through classroom communication. An example would be when the students correctly/incorrectly translate a text, fail/succeed to solve a mathematical equation, or share their concerns about a theory.

The learning process needs to be communicated in one way or another, in order for the teacher to observe it. Or, more correctly, the teacher can only observe a communicative contribution as an expression of the learning process (Luhmann, 2000, p. 152).

Second-order teaching is about teaching and learning that is explicitly communicated in a learning context. This is possible only *if* the learning process and the students' study strategies are addressed in the classroom, or, as Luhmann says, if they are thematized in the communication. From this perspective, second-order teaching may be seen as classroom communication *about* classroom communication. The purpose of this is to create study transparency through metacommunication (Thingholm, 2016), knowing that teaching and learning will never be fully transparent (Bengtzen & Barnett, 2017), but rather an attempt to create transparency despite opaque complexity, in order to achieve relative transparency (Luhmann, 2000, p. 152).

### *Context learning*

The concept of second-order teaching draws not only on Luhmann's concept of second-order observation, but also on Gregory Bateson's theory of logical categories of learning and communication (Bateson, 2000). When learners learn any given content, they simultaneously *also* learn about the context in which that content appears. This is based on learning studies, which show that test participants who have to repeatedly memorize senseless rhymes, learn not only the rhymes themselves, but also learn *how* to learn rhymes (Bateson, 2000; Keiding & Laursen, 2005). The learners may or may not be conscious of the context of their learning, or the strategy that they use to learn the rhymes. This is comparable to second-order observation that observes *how* something has been observed. With respect to education, several themes are embodied in context learning: students learning about the type of given task or assignment, about academic norms and values, about the school and class as a specific learning environment in which to collaborate, and generally, learning about themselves as learners (Laursen, Keiding, & Johansen, 2003).

As mentioned, context learning may take place without the students being conscious of how they learn and which strategies they make use of. In that case, context learning is only available for systematic reflection to a certain extent. To develop systematic reflection on study strategies or context learning has certain metacommunicative implications for teaching and learning. These implications are revealed below.

### *Implications for teaching practice: metacommunication*

In an educational setting, and inspired by the concepts of metacognition and second-order teaching, the focus moves from "thinking about thinking" to "classroom communication about classroom communication." These concepts share a meta-perspective, but the first concept, thinking about thinking, refers to cognition and the psychological system, which is invisible for everybody else except the individual student. The second concept, communication about communication refers to the social system and the teaching taking place in the classroom as a joint effort. It is communication that the students, as opposed to cognition, can connect to and engage in with the teacher and each other. The purpose of metacommunication is to understand understanding by explicitly addressing not only the academic curriculum (the what dimension), the strategies for achieving the curriculum (the how dimension) and the purpose of introducing the curriculum (the why dimension), but also asking the students to reflect on and explain *how they understand* the curriculum, the strategies and the

overall purpose. This kind of communication helps the students to observe their own learning process. They learn to do second-order observation in relation to their own learning and understanding of any given topic. They learn to observe how they have observed by identifying the distinction(s) behind their first observation and understanding. The teacher can facilitate this kind of meta-learning by asking the students to explicitly explain how they constructed meaning and how they relate to the curriculum. When facilitating this second-order learning, the teacher references the didactic dimension described above: What did you understand from the curriculum? (content dimension). Which strategies did you use? Could you have done anything differently? (method dimension). How is this meaningful to you? (purpose dimension).

Thereby, two parallel perspectives appear for the teacher to reflect on: 1) the teaching perspective and 2) the learning perspective. The first of these two perspectives involves reflection in relation to the classical didactic questions about “what-how-why”: content, methods and purpose. This perspective is mainly concerned with the teaching dimension and how the teacher presents content. The second perspective concerns the learning perspective and involves reflection on how the students learn content. In other words, the learning dimension concerns the context and how it is possible to explore it. Moreover, the focus is not on the teacher’s instruction, but on the students’ learning processes (thinking and doing). From this perspective, the “what-how-why” questions are related to the learning process: What am I, as a learner, about to learn? How will I best learn this content? Why do I have to learn this content? Why is it important to me? (Thingholm and Keiding, 2018).

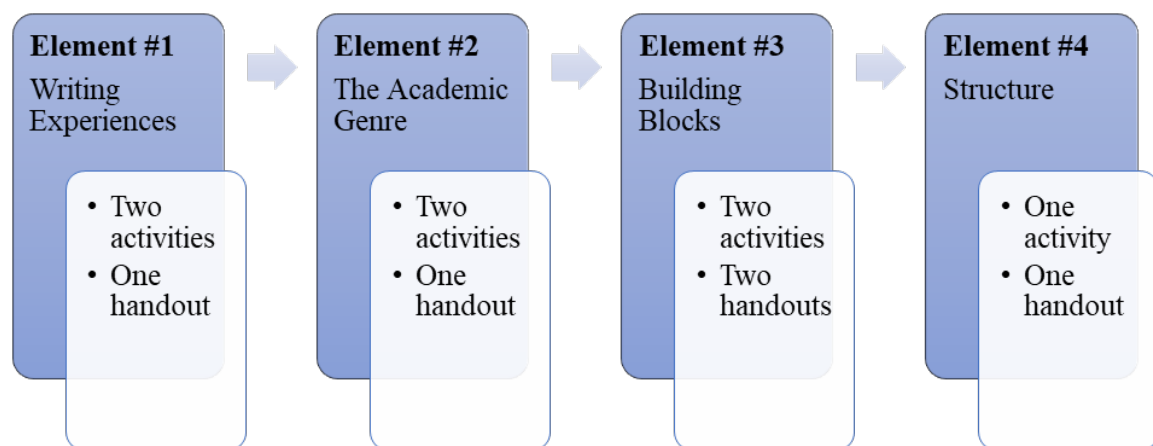
When teaching study strategies explicitly, the acquisition of relevant study strategies becomes more than just unconscious context learning, and is no longer something that students either do or do not manage. The responsibility for systematically acquiring relevant study strategies shifts from belonging to the individual student, to being something that pertains to classroom communication. Second-order teaching takes the form of a joint dialogue concerning *what* knowledge and skills students need to learn, and *how* and *why* they might learn these skills, but it also includes a ‘looking back’ perspective – an observation of how the students did understand the curriculum and an understanding of this understanding – which distinctions were underlying their understanding. Second-order teaching is metacommunication, and helps the students not only to understand the academic content, but also to understand the various study strategies used to understand the academic content. Put in theoretical terms, second-order teaching explicitly helps the students to observe how they have observed the academic content. In this way, the students become self-regulated learners who are able to control and regulate their own learning processes. Similar ideas are suggested in theories of self-regulated learning (Weinstein, Bråten, & Andreassen, 2013).

This kind of metacommunication calls for a kind of dialogic pedagogy, in the sense that the teacher needs to engage in a dialogue with the students about how they study, about their strategies for acquiring new knowledge and skills, about their norms and values and about when they enjoy learning. It is equally important that the teacher introduces new study strategies, and that the students share their strategies, in order for them to acquire a repertoire of useful strategies. The Takeaway Teaching design is structured to support this kind of metacommunication in the classroom, and explicitly deals with which study strategies the students employ and how these strategies may be developed over time.

## The design of Takeaway Teaching

Above, we have unfolded the theoretical framework for the Takeaway Teaching design. Below, we describe the structure of the design and its themes. It consists of nine generic themes and 13 pre-designed mini-plans or scripts. The reason for choosing this pre-designed scripting design was to make it easy for university teachers to choose one (or more) relevant theme(s), read the content, adjust or redesign it and integrate it into the lesson plan. Each theme is developed by one or two subject developers who have a professional interest in the specific content. For example, the Literature Search theme is developed by a librarian who has professional experience of providing workshops for students on how to search and manage literature. Each theme is structured around a different number of elements and each element includes a number of activities. This structure recurs in each theme, in order to make the design consistent and easily accessible to the teachers. Figure 1 is an example of the structure of one of the themes; Academic Writing – Genre and writing process.

Figure 1. The structure of the theme Academic Writing - Genre and writing process



In the scripts, the activities are described in detail and the developers have included relevant didactic comments aimed at the teachers. If necessary, handouts for the students are included. Furthermore, for each element, the developer gives an estimate of how long it will take the students to complete the activities. The themes are all generic and may be directly implemented in the lesson plan, or they may be adjusted and redesigned by the teacher to better match the curriculum and the students. The design helps the teacher to facilitate meta-learning not only by addressing strategies, for example academic reading and writing, but also by asking the students to think about how they usually reflect on strategies for academic reading and writing. According to the theoretical frame, students are asked to observe how they observe. The students have the chance to discuss and understand their current way of understanding academic reading and writing, an understanding that develops during their years of studying. Using the design makes it easier for the teacher to facilitate discussion about the students' different understandings of how and why to read and write.

When students are explicitly taught different generic strategies for studying they have the chance to relate to their own knowledge and use of strategies. An example of understanding academic writing could be the progression from understanding writing as a skill, *learning to write*, to understanding writing as a process of knowledge construction, *writing to learn*. The themes and the activities help to foster the understanding of this learning perspective. Exer-

cises and handouts contribute with questions about the students' construction of knowledge in order to help them reflect on their own learning and develop academically. For example, they reflect on their own writing experiences and how academic writing is far more than a product; it is an ongoing cognitive process that creates new knowledge.

The themes are not constructed as online learning. It is not sufficient for the teacher to send the website's URL to the students or digitally conduct the activities without integrating them into the curriculum. In this sense, the design is not a "plug-and-play" activity. The design is available on an online learning resource, but the content is not an independent online course. Just as with any other teaching activity, the teacher needs to reflect on how to implement a Takeaway Teaching theme and its activities when planning the curriculum and the lessons. The activities must be facilitated by the teacher, but may be conducted by the students in class, or between classes as homework or group work.

The content of each theme is visible on the website and may be downloaded for use, but the content of each theme is also uploaded to the online study platform, Blackboard, in order to make it easy to implement the activities in the rest of the lesson plan. Whether the teacher chooses a printed or a digital version of the theme, they must read it thoroughly, possibly making changes here and there, removing unwanted activities, or adding supplementary activities and materials of their own. This redesign may be done by each individual teacher, but it may also be conducted by groups of teachers in the same discipline who want to try out the same theme. In this way, the teachers may develop more discipline-specific elements and activities if they wish to do so.

### **Example of a redesign process**

In this section, we present an example of a redesign of a Takeaway Teaching theme in practice. In this example, the Literature Search theme was integrated into the Communication on the Internet course. This course is a part of the Information Studies master's programme at Aarhus University. The course is a 10-ECTS point course with 12 physical, face-to-face classroom sessions, various online exercises and one webinar. In the autumn of 2017, the class consisted of 18 students who, at the end of the course, handed in a short assignment using theories from the curriculum, combined with a case study of their own choosing and with additional literature. The course teacher and co-author of this paper, Anders Hjortskov Larsen, found the Literature Search theme appropriate, because he wanted to promote and support his students' skills in searching for relevant literature for their case study and assignments. Typically, the students would do the literature search on their own at home or maybe go to the library for help. With the pre-designed activities, Larsen found that it was possible to integrate the students' literature search into the lesson plan and facilitate the processes of literature search within the discipline at hand.

After requesting the Literature Search theme by e-mail, it was uploaded to the Blackboard course module. The Literature Search theme is based on activities, texts, tasks and videos showing the actual use of a database with truncation and other technical search strategies. This theme consists of three independent scripts: Inspirational Literature Search, Systematic Literature Search, and Reference Management. The first script focuses on a literature search as inspiration for identifying a problem statement or research question. The second script focuses on a systematic literature search in order to find relevant literature to enable discussion of the problem statement or research question. The third script is an introduction to the



reference programme EndNote. In this case, the teacher chose to implement the first and second scripts into the curriculum, and omit the third script about EndNote.

This was the first broad decision of how to use the theme. Another general choice that led to a redesign of the given material involved the choice of digital media. In this case, the LMS Blackboard was not adequate for the way the course was being taught. Therefore, in order to fit the digital media setup of the course, the search results and discussions surrounding them were moved from Blackboard to a Facebook page and a WordPress blog, which were already integrated parts of the course.

The two scripts were timed to fit into the overall progression of the course, in order to motivate the students in the most effective way. The students handed in assignments on a weekly basis and the two scripts were timed to fit with the progression of these course assignments. Therefore, the first script was timed to fit with the third course assignment. Here, the students came up with new literature, and discussed it online in blog posts with their co-students and the teacher. Five weeks later, the weekly assignment was to further develop the students' research questions, and find additional literature that could support their analyses. Therefore, the second script was integrated. The requirement for this weekly assignment was to apply additional literature based on the new search strategies that they had learned. This requirement was aligned with the exam.

After the students had worked with the Takeaway Teaching theme, an evaluation was conducted as an in-class workshop, and later on in the official course evaluation, multiple questions were asked concerning the literature theme from the Takeaway Teaching design. The students discussed the focus on their study skills as something they already had worked with during their bachelor's programme, but they also mentioned that they found the various activities useful for the fulfilment of the weekly tasks. Some students mentioned the search strategies as something they knew about from earlier introductions in their studies, but had never used or reflected on.

#### *Looking back at the redesign*

Integrating the Takeaway Teaching theme opened up an opportunity to have a critical discussion about relevant literature and alternatives to the titles in the syllabus handed out by the teacher. In relation to the different kinds of knowledge; declarative, procedural and conditional, the class had a chance to reflect on and explain the three dimensions: what, how and why. The what dimension relates to what is known about the assignment at hand. Here, the students are asked to refer to relevant literature in order to support their academic argument in a text. The how dimension relates to knowledge about how to solve the task, in this case the ways in which relevant literature within the field are found, in other words which search strategies to use. The why dimension relates to knowledge about the purpose and the context, in this case the purpose of strengthening academic argumentation through references and to the learning taking place while searching for relevant literature, e.g. understanding that the syllabus is only a small part of a much greater academic field.

How did the re-design of the Takeaway Teaching theme support the teaching and learning within this example?

With regard to the what dimension, using the literature search theme did not change the demands of the assignment. The students were still asked to refer to relevant literature to support their argumentation just as before, but implementing the search activities increased

the focus on literature search and made the students reflect on the search skills, which they had gradually developed during their studies over the years.

With regard to the how dimension, the integration of the theme led to two new elements: Firstly, the students' literature search was explicitly integrated and related to the actual discipline, thereby reducing the problem of transfer from generic search courses at the library. Secondly, by presenting and trying out various search strategies in class the students gathered a greater repertoire of strategies as opposed to students doing the literature search on their own.

With regard to the why dimension, the integration of the theme made it possible for the teacher to be explicit about the purpose and the context of the literature search. This kind of discussion opened up an opportunity for the students to understand their own understanding of searching for literature – how to do it and why to do it. These kinds of meta-reflections do not happen just because a theme from the Takeaway Teaching design is implemented in the lesson plan, but the design does create an opportunity or space for the teacher and the students to allow these meta-reflections to take place. Larsen experienced that time was an important factor for creating these meta-reflections. Developing an awareness and use of search strategies as knowledge construction takes time and does not happen in one semester. Furthermore, developing metacommunication is not easy and it is an ongoing process for the teacher as well as for the students. Larsen found that implementing the Takeaway Teaching theme provided an opportunity for developing metacommunication, but it was still difficult and unfamiliar to him to have this kind of communication with the students. Using the same theme over time, this might change and metacommunication may gradually become more natural for the teacher.

Looking more closely at the example above, the teacher experienced an improved focus on discussions about searching for literature and a critical perspective on the literature of the syllabus. The process of searching for applicable literature in relation to the assignments within the research field supported the students' reflections and understanding of the academic field. Furthermore, it supported the students' general understanding about how to create new knowledge. In other words, the students learned to search for literature but to a certain extent, they also searched to learn and understand the literature of the field.

In this example, using the literature became a more active and reflective part of the students' learning process instead of being a demand from the teacher. The literature search thereby opened up learning processes within a concrete and real context instead of study skills and activities being disconnected from the actual curriculum. In this way, the literature search activities and strategies qualified the students' learning outcomes and their reflections about their knowledge acquisition throughout the course. This was visible in the way in which the students in their assignments referred to literature outside the syllabus. The students were more reflective and creative in their use of literature than previous students. Thereby, the students achieved a broader understanding of the subject knowledge of the course. By using the Takeaway Teaching theme and explicitly connecting search strategies to concrete assignments, the teacher found a greater alignment throughout the course.

## Challenges and potentials of Takeaway Teaching

To sum up, Takeaway Teaching inspires university teachers to explicitly address and develop study strategies as a part of their curricula. The design has potential with regard to two aspects. Firstly, the teachers reflect didactically on how to teach students to study a given discipline, and by implementing the Takeaway Teaching themes they make teaching and learning more transparent by addressing how students understand their understanding, and how developing study strategies is essential for knowledge construction. Secondly, from this explicitness the students gain by developing new study strategies and becoming more meta-oriented with respect to their own learning processes.

Therefore, we argue that strengthening the teachers' didactic ability while developing the students' study skills improves the quality of teaching and learning. The design differs from ordinary textbooks by *not* presenting a certain academic discipline or content but by presenting useful strategies that support *the learning* of any given discipline. Takeaway Teaching is designed for university teaching and learning. However, the basic idea and design may also be useful at other educational levels.

Various aspects of Takeaway Teaching as a teaching design are challenging, and in the following section, we will highlight five challenges. Firstly, convincing teachers to try the new design has been challenging. Their hesitation is not related to the actual Takeaway Teaching design, or to the activities included in the themes, but mainly concerns the time needed to become familiar with the design, to spend time redesigning the activities and also to present it during lessons. Secondly, some teachers are unfamiliar with Blackboard, the university's LMS, and therefore find that integrating activities into their classrooms via Blackboard is an unmanageable task. We aim to resolve this by providing the described activities and ideas in various formats, such as print, new online formats and downloadable PDFs. Thirdly, it may be difficult for the teachers to adopt someone else's understanding and thinking about study strategies. The developer of a Takeaway Teaching theme, for example, Academic Reading, may have a different approach to reading scientific texts than the teacher who is going to apply the content and activities of the Takeaway Teaching theme. Fourthly, as we emphasized in the title of this paper, Takeaway Teaching is a *design for redesign*. This puts the teacher in charge of the content, the ways in which the study strategies are taught, and using meta-communication in the own classroom. The many redesigns of the actual content of Takeaway Teaching make it difficult to know precisely *how* the study strategies are being taught in different contexts and whether or not the second-order observations are fostered. Lastly, the Takeaway Teaching themes might limit the teachers' creativity if they believe that the themes and activities in the Takeaway Teaching design are conceived as the *only* way to teach study strategies, which is quite contrary to the intention of the design. Takeaway Teaching presents only *some* ways to teach study strategies, and is not an exhaustive collection of how study strategies might be taught in higher education.

Despite the possible challenges with using Takeaway Teaching, we do find that the Takeaway Teaching design is valuable and innovative because it manages to address the teacher and the students at the same time. Often, both generic teaching courses for teachers and generic study courses for students are detached from the curriculum, which may cause transfer problems related to their teaching and study practices. Takeaway Teaching integrates the development of practical teaching and study competence in the same lesson.

Furthermore, teachers and students have difficulty finding time to attend generic courses.

Takeaway Teaching aims to address this problem, because it relates directly to teaching and study practices. As an alternative to generic courses that are detached from the classroom, the Takeaway Teaching design may reach and benefit a large number of teachers and students. During the first semester, in the autumn of 2017, Takeaway Teaching was used in 42 different courses at Aarhus University and reached 337 students. These numbers have been tracked from Blackboard, and more teachers may have used printed themes downloaded directly from the website. This number is far more than the teaching development centre of the faculty could provide generic courses for during one semester.

At the moment, further research related to the Takeaway Teaching design is being carried out, and the detailed results of a qualitative study that closely follows five university teachers and their students using the Takeaway Teaching design will be published later this year. Practical knowledge about using the design will be obtained, in order to qualify the ongoing design of Takeaway Teaching and its related themes.

## References

- Bateson, G. (2000), *Steps to an ecology of mind. With a new foreword by Catherine Bateson*. Chicago: Chicago University Press.
- Bengtsen, S., & Barnett, R. (2017), "Confronting the dark side of higher education", *Journal of Philosophy of Education*, Vol. 51No. 1, pp. 114-131. Doi:10.1111/1467-9752.12190.
- Coutinho, S. (2008), "Self-efficacy, metacognition, and performance", *North American Journal of Psychology*, Vol. 10 No. 1, pp. 165-172.
- Flavell, J. H. (1979), "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry", in: *American Psychologist*, Vol. 34, No. 10, pp. 5 Doi: <http://psycnet.apa.org/doi/10.1037/0003-066X.34.10.906>.
- Frenkel, S. (2014), "Metacognitive components in learning to learn approaches" in: *International Journal of Psychology: A Biopsychosocial Approach*, Vol. 14 Doi: <http://dx.doi.org/10.7220/2345-024X.14.5>.
- Hargrove, R. A., & Nietfeld, J. L. (2015), "The impact of metacognitive instruction on creative problem solving", *The Journal of Experimental Education*, Vol. 83, No. 3, pp. 291-318. Doi:10.1080/00220973.2013.876604.
- Harrison, G. M., & Vallin, L. M. (2017), "Evaluating the metacognitive awareness inventory using empirical factor-structure evidence", *Metacognition and Learning*. Doi: 10.1007/s11409-017-9176-z.
- Keiding, T. B., & Laursen, E. (2005), *Interaktion og Læring. Gregory Batesons bidrag: Unge Pædagoger*.
- Kuhn, D., & Dean, J. D. (2004), "Metacognition: A bridge between cognitive psychology and educational practice", *Theory Into Practice*, Vol. 43, No. 4, pp. 268-273. Doi: 10.1207/s15430421tip4304\_4.
- Laursen, E., Keiding, T. B., & Johansen, L. Ø. (2003), IT og Bateson: Et læringsteoretisk perspektiv. H. Mathiasen (Ed.), *IT og læringsperspektiver*: Alinea.

- Luhmann, N. (2000), *Sociale systemer*. København: Hans Reitzels.
- Luhmann, N. (2007), *Indføring i systemteorien*. København: Unge Pædagoger.
- Pintrich, P. R. (2002), "The Role of metacognitive knowledge in learning, teaching, and assessing", *Theory Into Practice*, Vol. 41, No. 4, pp. 219-225. Doi: 10.1207/s15430421tip4104\_3.
- Schraw, G., Crippen, K. J., & Hartley, K. (2006), "Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning", *Research in Science Education*, Vol. 36, No. 1, 111-139. Doi: 10.1007/s11165-005-3917-8.
- Thingholm, H. B. (2016), "Study Transparency through Meta-communication", paper presented at the SRHE International Conference on Research into Higher Education, Wales.
- Thingholm, H. B., & Keiding, T. B. (2018), "Second Order Teaching", *Journal of the International Society for Teacher Education*, Vol. 22, No. 1.
- Veenman, M. V. J. (2015), "Teaching for metacognition", *International encyclopedia of the social & behavioral sciences (2)* (pp. 89-95). Doi: <https://doi.org/10.1016/B978-0-08-097086-8.92136-6>.
- Weinstein, C. E., Bråten, I., & Andreassen, R. (2013), *Læringsstrategier og selvreguleret læring: Teoretisk beskrivelse, kortlægning og undervisning*. R. Andreassen, H. Bjerresgaard, I. Bråten, J. Hattie, M. Hermansen, T. N. Hopfenbeck, P. O. Kirkegaard, C. Madsen, H. Timperley, C. E. Weinstein, & T. S. Wille, *Feedback og vurdering for læring* (pp. 145-174). Frederikshavn: Dafolo.

# DUT Guide om akademisk skrivning

Tine Wirenfeldt Jensen<sup>a,1</sup>

<sup>a</sup>METoDo og SDU Universitetspædagogik, Syddansk Universitet

At kunne skrive er ikke bare en forudsætning for at kunne bestå eksamener på videregående uddannelser, det er også en metode til at lære sit fag og kunne gå i dialog med det. Alligevel er systematiske indsatser på området relativt sjældne på danske videregående uddannelser. Det betyder, at den enkelte underviser kan opleve at stå alene med opgaven med at støtte de studerende i at opbygge de nødvendige skrivekompetencer.

På baggrund af forskningslitteratur præsenterer artiklen ni tips til at understøtte udviklingen af akademiske skrivekompetencer og læring henvendt til undervisere. Rammen er et undervisningsforløb, og tipsene kan også bruges i vejledningen.

## Baggrund

Lea og Street (1998) tilbyder en nuanceret forståelsesramme for studerendes skrivning på videregående uddannelser, og rammen tjener samtidig til at indfange udviklingen på feltet på et overordnet niveau. Lea og Street (1998) identificerer tre modeller for akademisk skrivning: som studieteknik, som akademisk socialisering eller som akademisk *literacy*.

I en studieteknisk optik forstås skrivning som en teknisk færdighed, der uproblematisk kan overføres på tværs af kontekster, og som studerende burde mestre ved studiestart. I modellen akademisk socialisering opfattes studerendes skrivning som en del af en større socialiseringsproces ind i den akademiske kultur. I et akademisk *literacy*-perspektiv forstås akademisk kultur hverken som en homogen eller statisk størrelse. Skrivning opfattes som en social praksis med fokus på at skabe og udfordre mening, en praksis som er tæt forbundet med fagenes vidensforståelser og de studerendes identitetsudvikling.

Selvom forfatterne giver modellen akademisk *literacy* forrang, så afviser de ikke de to andre forståelser, men opfatter dem som indlejrede i hinanden, således at akademisk *literacy* medtænker de to første modeller i en større kontekst. De tre modeller vægter i stigende grad fagdisciplinens betydning for skriveudvikling og læring. Akademisk *literacy*-forståelsen synliggør, at akademisk skrivning ikke meningsfuldt kan adskilles fra faglig læring (Dysthe et al., 2001). Samtidig åbner forståelsen et refleksionsrum for at diskutere fagets vidensformer og videnskabsteoretiske positioner.

Lea og Streets modeller kan bruges som refleksionsramme for underviseres tiltag i forhold til studerendes skrive- og læreprocesser. Den omfangsrige håndbogslitteratur på området (Booth et al., 2016; Rienecker og Jørgensen, 2017, m.fl.) kan fx kritiseres for at være udtryk for studieteknikmodellen (Hagström, 2005), men med Lea og Streets forståelse af de tre modeller som indeholdt i hinanden er det muligt som underviser at forankre udvalgte dele af denne type materiale i faget. Det forudsætter et fokus på at koble materialet til faget og dets

---

<sup>1</sup> Kontakt: [tine@metodo.dk](mailto:tine@metodo.dk)

vidensformer og give plads til fælles diskussion og kritik. På samme måde kan peer feedback forstås som en del af en akademisk socialiseringsproces, men ved at give plads til kritisk dialog om fagets videnskabsteoretiske positioner, traditioner og konventioner kan aktiviteten også forstås i et akademisk *literacy*-perspektiv.

De forskellige fag stiller i kraft af deres vidensforståelser også forskellige medierende redskaber til rådighed for de studerendes skrivning (Blåsjö, 2004), og i arbejdet med at understøtte akademisk skrivning er det afgørende, at der eksisterer en tæt kobling til den konkrete fagdisciplin (Carter et al., 2007). Faglige undervisningsforløb indeholder dermed et stort potentiale til at understøtte studerende i at oparbejde akademiske skrivekompetencer. Særlig hvis undervisningen ikke reduceres til kun at have et normativt sigte, men at skrivning også opfattes som en vigtig arena for, at studerende kan lære deres fag og udvikle en selvstændig faglig identitet og stemme (Ivanič, 1998).

### Praksispunkter

- Sørg for overensstemmelse mellem skriftlige eksamener, opgavens læringsmål og undervisningen
- Brug skrivning til at understøtte læreprocessen
- Ekspliciter vurderingskriterier
- Synliggør argumentationens rolle
- Diskuter forventningerne til selvstændighed i opgaven
- Diskuter krav til akademisk sprog
- Understøt problemformulerings- og litteratursøgningsprocessen
- Giv redskaber til at kvalificere skriveprocessen
- Integrer struktureret peer feedback og selvevaluering

### Tip 1

#### **Sørg for overensstemmelse mellem skriftlige eksamener, opgavens læringsmål og undervisningen**

At der bør være en intern sammenhæng (*alignment*, Biggs & Tang, 2007) i uddannelsesaktiviteter er ikke nyt. Alligevel er det ikke altid, at der er den nødvendige overensstemmelse mellem en skriftlig opgave som eksamensform, de opstillede læringsmål for opgaven og så tilrettelæggelsen af den undervisning, de studerende deltager i undervejs i forløbet. Hvis eksamensformen er skriftlig, bør underviseren sikre, at forløbet indeholder rige muligheder for, at de studerende skriver og får feedback på deres tekster. Dette gælder uanset eksamensformen (bundet eller selvstændig problemformulering, hjemme eller stedbundet, få timer eller måneder til rådighed). Brugen af en arbejdsportfolio i undervisningen kan være en metode til

at give de studerende mulighed for løbende at producere, samle og reflektere over tekst produceret i undervisningen, og portfolio kan også på forskellige måder danne udgangspunkt for den endelige bedømmelse (Dysthe & Engelsen, 2003).

Som underviser har man et begrænset råderum i forhold til de rammer, der er fastlagt for undervisningens mål og prøveformer. Det er dog ofte muligt at tilrettelægge egne undervisningsaktiviteter, formulere stillede problemformuleringer og at udarbejde egne vejledninger til opgaven. I dette arbejde er det centralt at sikre, at valg og materialer er kompatible med de overordnede rammer og mål (Skov, 2015). Den manglende interne sammenhæng kan ellers medføre, at underviserens pædagogiske praksis kommer i konflikt med censurs forpligtelse til at bedømme efter studieordningens kriterier, hvilket kan påvirke eksamens validitet.

## Tip 2

### Brug skrivning til at understøtte læreprocessen

I undervisningens planlægning er det læringsfremmende at tilrettelægge aktiviteter, hvor de studerende kan skrive og får lejlighed til at forholde sig til det skrevne. Det giver både de studerende mulighed for at tilegne sig stoffet ved at arbejde aktivt med det og træner dem i at skrive.

Man kan fx tilrettelægge skriveøvelser, de studerende skal arbejde med i løbet af undervisningen, og disse kan gøres til genstand for feedbackprocesser eller selvevaluering (se tip 9). Det øger motivationen og læringsudbyttet, hvis studerende får mulighed for at anvende feedback fremadrettet, fx i forbindelse med omskrivning eller diskussion (Dysthe, 2011; Nicol & Macfarlane-Dick).

Udover deciderede skriveøvelser er der mange andre måder, hvorpå underviseren kan integrere skrivning som læringsaktivitet i undervisningen (Bean, 2011). Ved introduktion af nyt stof og begreber i undervisningen kan man give de studerende 3-5 min. til at formulere deres forståelse ved fx at beskrive et konkret eksempel eller sætte det nye i relation til allerede kendt stof (Kristiansen, 2017). Det skrevne kan efterfølgende danne afsæt for diskussioner i mindre grupper eller i plenum. Man kan også arbejde med reflekterende skrivning, hvor studerende fx i slutningen af undervisningen nedskriver, hvad de vigtigste pointer var, eller hvad de fandt særligt svært, og derefter diskuterer deres svar med hinanden (Kristiansen, 2017). Endelig kan man også bede de studerende formulere sig på skrift som forberedelse, fx ved at besvare spørgsmål stillet af underviser eller formulere egne spørgsmål til en tekst og dele materialet på e-læringsplatformen inden undervisningen (Skov, 2015).

## Tip 3

### Eksplíciter vurderingskriterier

Studerende og undervisere deler ofte ikke forståelse af, hvad der "tæller" i en opgave, og studerende kan grundlæggende være i tvivl om, hvad formålet overhovedet er med deres opgave (Norton, 1990; Kristiansen, 2010; Hegelund & Kock, 2003). Hvis de ikke forstår kravene til opgaven, så kan de ikke udvælge relevante skrivestrategier og opnår ikke det intendede læringsudbytte.

Studieordningen indeholder en række kriterier, men man kan ikke forudsætte, at de studerende på egen hånd kan afkode disse (Skov, 2015). Det er derfor nødvendigt at bruge tid på



at diskutere dem, give mulighed for spørgsmål og selv uddybe særligt vigtige pointer (Nicol & Macfarlane-Dick, 2006). En del af de spørgsmål, der opstår i opgaveskrivningen fx i forhold til prioritering af stoffet, vil ideelt set kunne afklares ved at konsultere studieordningens beskrivelse af læringsmål. De studerende skal altså både lære at afkode en studieordningsbeskrivelse og bruge den aktivt undervejs i skriveprocessen.

De læringsmål, som er opstillet i studieordningen, vil sjældent kunne stå alene. For yderligere at ekspliciter og synliggøre kravene til opgaven, kan man udarbejde en *rubric*, dvs. et skema, der udfolder kriterier og giver mulighed for at vurdere, i hvilken grad de er opfyldt (Reddy & Andrade, 2010). Udover at give en bedre forståelse af kravene til opgaven vil brugen af rubrics også kvalificere peer feedback (se tip 9). Man kan inddrage de studerende i arbejdet med at afkode og opstille kriterier, fx i forbindelse med en skriveøvelse.

Det hjælper også de studerende til at forstå kriterier, hvis de får udleveret eksempler på konkrete, gode præstationer, som de kan bruge til at sammenligne deres eget og andres arbejde med (Orsmond et al., 2002).

#### Tip 4

##### Synliggør argumentationens rolle

For at kunne skrive gode opgaver er det en forudsætning at kunne argumentere på fagets præmisser og at kunne analysere og diskutere andres argumentation. Det er også en forudsætning at være bevidst om, at selvstændige opgaver er karakteriseret ved at være vidensproducerende og ikke blot vidensrefererende.

Som en metode til at give de studerende en forståelse for argumentationens rolle forslår Kock & Hegelund (2003) at anvende Toulmins argumentmodel som en forståelsesramme for akademiske tekster som ét, overordnet argument. Toulmin anvender kategorierne påstand, belæg, hjemmel, gendrivelse, rygdækning og styrkemarkør til at synliggøre argumentationens enkelte delelementer og indbyrdes sammenhæng (Toulmin, 1958). Kategorierne er velegnede til at vise argumentstrukturer i akademiske tekster, herunder sammenhængen mellem data/materiale, teori, metode og metodediskussioner. Modellen kan både bruges som en analysemodel i forhold til andres tekster og som et arbejdsredskab i forbindelse med selvstændige opgaver.

I undervisningen kan man støtte de studerendes forståelse af argumentationens rolle ved fx at fremhæve udvalgte passager i forløbets tekster og sammen diskutere deres funktion i forhold til de overordnede pointer/fund. Man kan også analysere uddrag fra en opgave med henblik på at identificere og kategorisere de forskellige argumentelementer (Kock & Hegelund, 2003).

#### Tip 5

##### Diskuter forventningerne til selvstændighed i opgaven

Selvstændighedsbegrebet kan forstås på mange forskellige måder (Jensen, 2018). Man kan derfor ikke forvente, at de studerende af sig selv kan afkode, hvad det indebærer at vise selvstændighed i det specifikke fag, eller at de har oplevet en progression i opgaveskrivningsarbejdet, der har rustet dem til at vise selvstændighed. Man kan heller ikke forvente, at de i udgangspunktet kan navigere mellem kravet til selvstændighed og kravet til at angive kilder. Det stigende fokus på plagiatkontrol kan i nogle sammenhænge føre til misforståelser og

uheldige skrivestrategier, der direkte modvirker selvstændighed (Pittam, et al., 2009; Jensen et al., 2016). Så hvis der eksisterer en forventning om selvstændighed i opgaven (eksplicit eller implicit), er det nødvendigt at tale med de studerende om, hvordan de kan forstå kravet om selvstændighed inden for den specifikke faglighed, og hvordan de kan arbejde sig henimod at indfri dette krav på en konstruktiv måde. Hvad der i særlig grad fokuseres på er forskelligt på tværs af fag, men undervisere/vejledere vil ofte koble selvstændighed med fx emnevalg, problemformulering, teori- og metodevalg samt inddragelse af empiri og brug af kilder (Jensen, 2018; Rienecker et al., 2019).

## Tip 6

### Diskuter krav til akademisk sprog

Der hersker mange misforståelser om akademisk sprog. Det kan resultere i, at studerende skriver i et unødigt komplekst sprog, de ikke helt behersker, hvilket kan føre til næsten ulæselige opgaver. Andre fokuserer for ensidigt på finpudsningen af enkelte formuleringer og for lidt på fx at skabe sammenhængende argumentation. "Dårligt sprog" i opgaver kan også være et udtryk for en læreproces, et tegn på at studerende kæmper med at tilegne sig nye fagspecifikke skrivemåder og nyt stof. Udviklingen af skrivekompetencer er ikke en lineær proces, og det er normalt, at processen involverer regression (Gere, 2019; Kristensen, 2017).

Gennem diskussion af krav til akademisk sprog i undervisningen kan de studerende få lejlighed til at forholde sig til egne antagelser. Som et redskab til at diskutere emnet kan man inddrage Signe Skovs skema over kriterier for opgavesprog, der er inddelt i kategorierne videnskabeligt, klart og korrekt sprog (Skov, 2006; Skov, 2008). I diskussionen bør kriterierne kobles til faget gennem eksempler og gerne suppleres med fagspecifikke krav. Fremfor at have et rent anvisende sigte bør diskussionen inddrage baggrund og ræsonnementer for kriterierne og synliggøre koblinger til videnskabsteoretiske forståelser. Det kan åbne for et fælles refleksionsrum, der kan muliggøre en erkendelse af, at kravene ikke er statiske og entydige, men et resultat af kontinuerlige forhandlinger (Holm & Clemensen, 2017). Denne bevidsthed kan hjælpe de studerende til selv at kunne analysere fremtidige kommunikationssituationer og træffe kvalificerede valg på egen hånd.

## Tip 7

### Understøt problemformulerings- og litteratursøgningsprocessen

Det er udfordrende for mange studerende at problemformulere og søge litteratur i forbindelse med skrivning af selvstændige opgaver. At give redskaber til at understøtte skriveprocessen (se tip 8) er en hjælp, men kan ikke stå alene.

Processen kan fx understøttes, ved at den studerende arbejder sig igennem en form for spørgeguide udarbejdet og evt. tilpasset til formålet. Booth et al. (2016) tilbyder en sådan spørgestruktur under overskrifterne *From Topics to Questions* og *From Questions to a Problem* (ibid. s. 33 ff.). Den interaktive problemformuleringsguide *Scribo* (Rienecker, 2014) opstiller også en række spørgsmål og kan derudover opsamle svarene og give adgang til eksempler på opgaver på tværs af fag. De tekster, de studerende producerer ved hjælp af sådanne spørgeguides, er velegnede til at danne udgangspunkt for peer-aktiviteter og underviservejledning.

Begge de nævnte spørgeguides kobler problemformuleringsprocessen med litteratursøgning. At kunne formulere sit vidensbehov, orientere sig i og fremsøge kilder og kritisk evaluere disse er dybt forbundet med faglige læreprocesser og kan ikke forstås som en teknisk kompetence (Andreasen et al., 2008). De nævnte kompetencer er en forudsætning for at kunne situere en problemformulering fagligt i en selvstændig opgave. De studerende har derfor brug for at kunne udarbejde relevante søgestrategier samt at opnå en bevidsthed om, at arbejdet med at lave oversigter over søgeord og søgeskemaer kan være en frugtbar metode til at skærpe og afgrænse en problemformulering. Dermed bliver litteratursøgeprocessen i sig selv en vidensproducerende aktivitet (Bruce & Hughes, 2010).

Man kan som underviser med fordel samarbejde med bibliotekernes informationsspecialister med kendskab til fagområdet, særligt hvis fokus er på at skrive selvstændige opgaver, hvor der er en forventning om, at den studerende selv finder og evaluerer kilder.

## Tip 8

### Giv redskaber til at kvalificere skriveprocessen

For at støtte de studerendes skriveproces kan man introducere opdelingen mellem tænkeskrivning og præsentationsskrivning (Dysthe et al., 2001). Tænkeskrivning (også kaldet processkrivning) dækker over tekst, der har til formål at frembringe/afprøve ideer og klargøre tanker, og som ikke vægter formelle krav. Præsentationsskrivning dækker over den færdige opgavetekst, der lever op til alle genrens krav. Tænkeskrivning er en metode til at arbejde sig hen mod færdig opgavetekst. Opdelingen kan hjælpe de studerende med at komme tidligere i gang med skrivearbejdet, at skrive bedre opgaver og kan afhjælpe udbredte problemer som fx udsættelse og perfektionisme (Dysthe et al., 2001, Elbow, 1973).

For både at understøtte de studerendes brug af tænkeskrivning og deres læring kan man inddrage aktiviteten i undervisningen. De studerende kan fx hurtigskrive (Elbow, 1973) i 5 min. om udvalgte faglige temaer eller spørgsmål og derefter få tid til selv at læse og revidere teksten. Øvelsen kan tjene som optakt til fx produktion af tekst til en wiki, skriveøvelser eller opgaveskrivning (Bean, 2011). Metoden kan understøttes gennem inddragelse af digitale redskaber (se fx nonstopskrivningsmodulet på Studypedia.au.dk; Jensen og Koefoed, 2003).

Akademisk skrivning indebærer at arbejde med revision af tekster. Revision kan kvalificeres ved at de studerende forholder sig til teksten på et globalt niveau før et lokalt niveau (Dysthe et al. 2001). Dyste et. al (2001) forslår, at man i prioriteret rækkefølge fokuserer på følgende fem niveauer: skrivesituation (formål, modtager); indhold og struktur; sætninger; ordvalg og retskrivning og endelig tegnsætning og formalia. Alle niveauer er vigtige, men modellen kan orientere studerende mod at forholde sig til tekstens overordnede formål, argumentation og struktur, før de sætter fokus på specifikke formuleringer og tegnsætning. Det giver fx ikke meget mening at arbejde længe på at finpudse sætninger, der så viser sig indholdsmæssigt ikke at have en plads i den overordnede argumentation. Princippet kan også med fordel anvendes i forbindelse med feedback.

Skriveprocesser er ofte usynlige i undervisningen. Det kan derfor være lærerigt at diskutere, på hvilke betingelser faglige tekster er blevet til, fx hvordan de ofte gennemgår gentagne omskrivninger på baggrund af peer reviews. Dette vil være med til at synliggøre, at revision og omskrivning af tekst er en naturlig del af skrivearbejdet. Underviseren kan også vælge at inddrage eksempler på egne tekster på forskellige stadier for at give indblik i, hvordan en faglig tekst kan udvikle sig.

## Tip 9

### Integrer struktureret peer feedback og selvevaluering

Feedbackforskningen viser, at peer feedback på skriftlige produkter kan være en effektiv metode til at lære. Når de studerende skal forberede og give kriteriebaseret feedback, arbejder de aktivt med at forstå og tilegne sig kriterier for faglige tekster. Denne træning overfører de efterfølgende i arbejdet med deres egne opgaver (Nicol, Thomson & Breslin, 2014; Gibbs, 1999). Dermed indeholder peer feedback et stort læringspotentiale, der ikke nødvendigvis forløses ved underviserfeedback, der gives i forbindelse med en karakter (Butler, 1987). Peer feedback anvendes som formativ feedback, så de studerende får mulighed for at anvende den feedback, de modtager, fremadrettet (Nicol & Macfarlane-Dick, 2006; Sadler, 1989). Man kan fx introducere skriveøvelser, hvor studerende efter peer feedback får mulighed for at arbejde videre med teksten, inden evt. aflevering til underviser.

For at peer feedback får den ønskede effekt, er det vigtigt, at studerende forstår rationalet bag metoden og får retningslinjer til at arbejde med den i praksis. Det vil fx være nødvendigt at skabe en fælles forståelse af, at arbejdet med at give andre feedback udgør et vigtigt træningsrum for opbygning af egne skrivekompetencer. Det er også nødvendigt med fælles retningslinjer for, hvordan man giver og modtager feedback, for at kvalificere processen og for at undgå, at negative sociale dynamikker opstår (Dysthe, Hertzberg & Hoel, 2001; Hvass & Heger, 2018). Der findes både digitale og analoge redskaber, der kan strukturere og understøtte peer feedback i undervisningen (*Peergrade*, se fx Bertel & Bentzen, 2018; *Tekstfeedbackspillet*, se fx Jensen, Jensen & Jørgensen, 2013).

Peer feedback sigter mod at sætte den studerendes egne handlinger i undervisningen i fokus og give dem mulighed for at oparbejde en evne til selv at foretage valg og vurderinger, baseret på ekspliciterede kriterier i deres skriftlige arbejde. En anden metode til at opnå denne evne er selvevaluering, hvor de studerende får mulighed for løbende at vurdere egne præstationer og øve sig i at regulere deres egen læring på en struktureret måde (Nicol & Macfarlane-Dick, 2006). Man kan fx bede studerende om selv at identificere styrker og svagheder ved deres tekst i forhold til ekspliciterede kriterier og aflevere refleksionerne på skrift inden evt. underviserfeedback.

### Konklusion

Det er især undervisere på forløb med skriftlige eksamensformer, som har et ansvar for at understøtte de studerende i at oparbejde de nødvendige faglige skrivekompetencer, men i alle sammenhænge kan skrivning integreret i den specifikke fagdisciplin anvendes som en effektiv metode til at lære.

Da den faglige disciplins egenart spiller en så afgørende rolle, kan man ikke give en entydig opskrift på, hvordan man bedst arbejder med akademisk skrivning - det afhænger af fagets genstandsområde, tænkemåder og vidensforståelser. Men de ni tips præsenteret i denne artikel tilbyder undervisere på tværs af fag et fleksibelt udgangspunkt til at tilrettelægge deres undervisning, så den kan understøtte opbygningen af akademiske skrivekompetencer og læring.

## Referencer

- Andreasen, L. B., Jørgensen, L., & Nielsen, J. L. (2008). Udvikling af informationskompetence i problemorienterede læreprocesser. *Antologier*, 75-90.
- Bean, J. C. (2011). *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom*. John Wiley & Sons.
- Bertel, T. F., & Bentzen, T. Ø. (2018). Formativ evaluering med IT-systemet Peergrade. *Tidskriftet Læring og Medier (LOM)*, 10 (18).
- Biggs, J.B. & Tang, C. (2007). *Teaching for Quality Learning at University* (3rd ed.). Buckingham: SRHE and Open University Press.
- Blåsjö, M. (2004). *Studenters skrivande i två kunskapsbyggande miljöer*. Acta Universitatis Stockholmiensis. Stockholm Studies in Scandinavian Philology. New Series. 37. Almqvist & Wiksell International.
- Booth, C, G. G. Colomb, J.M. Williams, J. Bizup & W. T. Fitzgerald (2016) *The Craft of Research*, The University of Chicago Press
- Bruce, C., & Hughes, H. (2010). Informed learning: a pedagogical construct attending simultaneously to information use and learning. *Library & Information Science Research*, 32(4), A2-A8.
- Butler, R. (1987) Task-Involving and Ego-Involving Properties of Evaluation: Effects of Different Feedback Conditions on Motivational Perceptions, Interest and Performance, *Journal of Educational Psychology*, 78(4), 210-216.
- Carter, M., Ferzli, M., & Wiebe, E. N. (2007). Writing to Learn by Learning to Write in the Disciplines. *Journal of Business and Technical Communication*, 21(3), 278-302.
- Dysthe, O., Hertzberg, F., & Hoel, T. L. (2001). *Skrive for at lære*. Århus: Forlaget Klim.
- Dysthe, O., & Engelsen, K. S. (Eds.). (2003). *Mapper som pedagogisk redskap: Perspektiver og erfaringer*. Abstrakt.
- Dysthe, O. (2011). 'What is the Purpose of Feedback when Revision is not Expected?' A Case Study of Feedback Quality and Study Design in a First Year Master's Programme. *Journal of Academic Writing*, 1(1), 135-142.
- Elbow, P. (1998). *Writing Without Teachers*. Oxford University Press, USA.
- Gibbs, G. (1999) Using assessment strategically to change the way students learn, in: S. Brown & A. Glasner (Eds) *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches* (Buckingham, SRHE/Open University Press).
- Hagström, E. 2005. Meningar om uppsatsskrivande i högskolan. Ph.d.-afhandling. Örebro University, Department of Education.
- Hegelund, S., & Kock, C. (2003). A Good Paper Makes a Case: Teaching Academic Writing the Macro Toulmin Way. In *Teaching Academic Writing in European Higher Education* (pp. 75-85). Springer, Dordrecht.
- Hvass, H., & Heger, S. 2018 Brugbar peer feedback: Instruktion og træning, før de studerende

selv skal give og modtage. *Dansk Universitetspædagogisk Tidsskrift*, 13(25), 59-70.

Holm, L., & Clemensen, N. (2017). At lære sig" de kloge damers" sprog: Studerendes perspektiver på akademisk skrivning. *Dansk Universitetspædagogisk Tidsskrift*, 12(23), 37-51.

Ivanič, R. (1998). *Writing and Identity*. John Benjamins Publishing Company.

Jensen, T.W. (2018). *Det danske universitetsspeciale - topografi, tekster og tendenser*. Aarhus Universitetsforlag.

Jensen, T. W. & Hansen, L. K. (2003). Next Stop in the Study Metro: Supporting Learning of Freewriting by Providing Interactive Feedback. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (Vol. 2003, No. 1, pp. 1291-1296).

Jensen, T. W., Jensen, G. S., & Jørgensen, A. H. (2013). Tekstfeedbackspillet: en metode til at strukturere peer feedbackprocessen. I R. Troelsen, & L. Rienecker (red.), 17 ting, du kan gøre ved din undervisning, DUNK12 Konferencepublikation. (s. 61-66), Dansk Universitetspædagogisk Netværk (DUN).

Jensen, T. W., Bay, G., & Andersen, P. (2017). Udvikling af studerendes akademiske skrivekompetence - en model for en indsats på fakultetsniveau. *Dansk Universitetspædagogisk Tidsskrift*, 12(22).

Kristiansen, B. (2017). *Om at skrive på universitetet*. Syddansk Universitetsforlag.

Kristiansen, B. (2010). Tekstproduktion og vidensproduktion. *Dansk Universitetspædagogisk Tidsskrift*, 5(9), 50-54.

Lea, M. R. & B. V. Street. 1998. "Student Writing in Higher Education: An Academic Literacies Approach." *Studies in Higher Education* 23.2.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative Assessment and Self-Regulated Learning: A Model and Seven Principles of Good Feedback Practice. *Studies in Higher Education*, 31(2), 199-218.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking Feedback Practices in Higher Education: A Peer Review Perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122.

Norton, L. S. (1990). Essay-Writing: What Really Counts?. *Higher Education*, 20(4), 411-442.

Olsen P.B. & K. Pedersen. (2018). Problemorienteret projektarbejde. Samfundslitteratur

Orsmond, P., Merry, S., & Reiling, K. (2002). The Use of Exemplars and Formative Feedback when Using Student Derived Marking Criteria in Peer and Self-Assessment. *Assessment & Evaluation in Higher Education*, 27(4), 309-323.

Pittam, G., Elander, J., Lusher, J., Fox, P., & Payne, N. (2009). Student beliefs and attitudes about authorial identity in academic writing. *Studies in Higher Education*, 34(2), 153-170.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448.

Rienecker, L. (2014). *Scribo*. Samfundslitteratur

- Rienecker, L. & P.S. Jørgensen. (2017). *Den gode opgave - håndbog i opgaveskrivning på videregående uddannelser*. Samfundslitteratur.
- Rienecker, L., G. Wichmann-Hansen & P.S. Jørgensen (2019). God vejledning - af specialer, bacheloropgaver og projekter. Samfundslitteratur.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Skov, S. (2015). *Læringsorienterende kursusdesign*. Samfundslitteratur.
- Skov, S. (2006). Hvordan bedømmer man det sproglige i universitetsopgaver? *Dansk Universitetspædagogisk Tidsskrift*, 1(1), 18-25.
- Skov, S. (2008.) *Bundne opgaver - hjemmeopgaver og eksamensopgaver på videregående uddannelser*. Samfundslitteratur.
- Toulmin, S. E. (2003). *The Uses of Argument*. Cambridge University Press.
- Wingate, U. 2006. "Doing away with 'study skills'." *Teaching in Higher Education*. 11.4:457-469.97

# DUT Guide on teaching and learning in the international classroom

Karen M. Lauridsen<sup>a,1</sup>

<sup>a</sup>Centre for Teaching and Learning, Aarhus University

This DUT Guide offers seven pieces of good advice for those teaching international programmes to diverse student audiences. In essence, the advice is all about being a good lecturer in a higher education context, so the points addressed here are primarily the additional factors that are important to keep in mind when the programme content is internationalised, and the students have different educational, cultural, and linguistic backgrounds.

1. Assess your higher education environment.
2. Assess and adjust the international and intercultural components of your course.
3. Discover and respect the disciplinary, cultural, and linguistic diversity among your students.
4. Consider your students' learning needs.
5. Scaffold your students' learning: Focus on strengths and alleviate weaknesses.
6. Be explicit in your communication and behaviour.
7. Consider the group dynamics: Turn challenges into opportunities.

## Background

The *international classroom* and *international programmes* are cover terms that may reflect a range of different realities across higher education institutions (HEIs). They are often used interchangeably to cover the students' diverse educational, cultural, and linguistic backgrounds, the content of the course or programme, the language of instruction (in Denmark often English), or a combination of all of these factors. They are rarely used to indicate specific ways of teaching in the classroom, even though many lecturers encounter educational, cultural, and linguistic challenges that have implications for how successful they are when teaching such programmes.

In the following, it is assumed that the international classroom reflects the diversity of the student cohort, but also that it is a learning space in which the HEIs seek to establish what is known as *Internationalisation at Home*:

---

<sup>1</sup> Contact: kml@au.dk



*Internationalization at Home is the purposeful integration of international and intercultural dimensions into the formal and informal curriculum for all students within domestic learning environments.* (Beelen & Jones, 2015:69)

Internationalisation at home is part of the *internationalisation of higher education* (de Wit et al., 2015) and of the *internationalisation of the curriculum* (Leask, 2015).

Within this conceptual framework, the terms *international classroom* and *internationalised programmes* ideally cover the following:

- A curriculum content that, depending on the discipline, includes examples from different cultures or geographical areas, or comparisons across such cultures and areas. Examples of this would be not only teaching the political system in a given country, but comparing it with other systems in other countries (political science); understanding and comparing issues related to caring for the terminally ill in different cultures (medicine or nursing); or an interdisciplinary programme or course on clean water supply in different parts of the world.
- A set of intended learning outcomes that reflect such international and intercultural components, including the development of students' intercultural competences and leading to a set of graduate attributes for the programme as a whole. An example would be that the successful graduate has demonstrated the ability to lead and manage multicultural teams within the disciplinary context (e.g. engineering).
- A programme taught in a language that may or may not be the local or national language of the higher education institution. The obvious example in a Western European context is English Medium Instruction (EMI), which does not in itself make a programme international, but may be one characteristic of an international programme; such a programme may also be taught in the local (national) language, however.
- A programme taught to a diverse group of students. In addition to general demographic differences, race, gender, beliefs, etc., the students may come from different educational backgrounds (disciplines or national systems) and have different first languages and cultures. They may be mobile (international) students, or they may be local. Whether mobile or local, all students come together and learn from and with each other.

These characteristics of international programmes have implications for how they are taught. The advice offered below may be considered good academic practice in any classroom and is therefore always important to keep in mind; but the points made become even more crucial in an international classroom with diverse student cohorts.

Research on how to turn the principles of the conceptual underpinnings into academic practice is still an emerging field. The advice offered below is inspired by and developed from Leask and Carroll's Good Practice Principles (Carroll, 2015; Leask & Carroll, 2013); the IntlUni Project (Lauridsen & Cozart, 2015; Lauridsen & Lillemose, 2015), and the EQUiIP Project (forthcoming).

## Tip 1

### Assess your higher education environment

In Europe, most HEIs will have policies in place for the internationalisation of research and education. The question is how these policies and strategies are transformed into the academic practice of the international classroom. How conducive is the institutional environment to international programmes as outlined above? Does the HEI leadership and management team consider the number of programmes taught through English and the number of mobile students as the main indicators of the level of internationalisation, or are there in fact measures in place that seek to develop the international and intercultural components of the educational programme and of the study experience for all students? The answers to these questions determine the agency that you as a lecturer will have for teaching international programmes and navigating international classrooms.

## Tip 2

### Assess and adjust the international and intercultural components of your course

Let us take Danish higher education as an example. In order to attract a substantial number of non-Danish students, a programme or course would typically be taught through English. So when you develop your EMI course, or when you take over responsibility for a course developed by someone else, you would be well advised to start by ascertaining whether internationalised learning outcomes and graduate attributes have been defined. If not, the course description may just be a translation into English of an existing course that has been taught through Danish. If that is the case, you will need to decide whether or not you want to work with the powers that be and change the course description.

Whatever you choose to do will have implications for the content of the course and how you teach it. Experience tells us that, more often than not, students' motivation for learning in a multicultural environment and engaging in interaction with peers who have a different background than they do themselves, is determined by the extent to which this interaction is a requirement and an integrated part of the course and programme.

A course description with internationalised learning outcomes would be most conducive to intercultural learning and group dynamics. Examples of international and intercultural disciplinary content have already been given in the introduction above. Just as important is the way such content is taught. Let us consider an example.

It is important to recognise the resources in the classroom – whether in a large lecture hall or a small class setting. All students have prior knowledge and experience which may be leveraged in the classroom to the benefit of all students. Imagine a course in business management. Leadership hierarchies differ considerably between cultures, some of them high and complex, others flat and with more agency and responsibility for the individual employee. From the outset, students would tend to think that leadership hierarchies are similar everywhere; they would therefore not be aware of the vast differences that may be observed even between cultures that are otherwise relatively close to each other. When discussing a business case, the lecturer may invite students to suggest solutions that would work in their respective home cultures, thereby demonstrating the cultural differences to all students in the classroom. The lecturer must obviously encourage students from diverse backgrounds to come forward and explain their culture-specific solutions to the business case in order for all

the students to have a richer learning experience as a result of working with it.

When a lecturer decides to make all students' tacit knowledge on a given topic more explicit to everyone in the classroom, it has implications for how s/he plans the teaching sessions. In a large lecture hall, it may be something as simple as using questions in an electronic student response system to spark short pair or peer discussions as an activity during the lecture. In a small class setting, it will have implications for how student engagement in peer groups is organised and managed.

### Tip 3

#### **Discover and respect the disciplinary, cultural, and linguistic diversity among your students**

Part of your course planning is also knowing for whom you are planning the course. You will need to have an idea of the diversity among your students and reflect on how you can take that into consideration when you plan and teach your course.

Your students may have different disciplinary backgrounds: If you are teaching a course that attracts students from different disciplines, you will want to know which disciplines they come from and how you can leverage their diverse knowledge and skills for the benefit of all students in the classroom. You may, for instance, design a group task that can only be solved if students draw on their different content knowledge, different research methodology skills (statistics, production or analysis of quantitative or qualitative data), or whatever is appropriate within the context of your course or discipline. With a clear purpose, the task becomes meaningful in context for the students, and they learn to appreciate their differences as strengths in the task-solving process.

Your students may have different educational backgrounds: Some of your students may be used to an educational culture where most or all teaching takes the form of lectures where students are expected to listen and take notes in order to regurgitate what they have heard at the exams. Others, like most Danish students, will be used to an educational culture where students are expected to be actively engaged in peer collaboration in and between classes. If you expect your students to be actively engaged, you will need to consider how you will introduce this to and manage this with students from other educational backgrounds. One possibility here would be to design student collaboration so that all students in a group have clear roles and so that – in the course of the academic term – they all try the different roles at least once. Such roles could be as experts of a specific aspect of the disciplinary content or in specific methods (see the examples in the previous paragraph); it could be as note-taker and the one responsible for uploading a summary of outcomes in a blog or other medium on the learning management platform, as moderator of discussions in the group, or as presenter in the large group on behalf of the peer group, etc. If you explain from the beginning what this entails and why you do it, all students will benefit from working with each other and – again – learn to appreciate the differences among them.

Your students may have different cultural backgrounds: In addition to what has already been mentioned, your students may have different religions and beliefs, or come from cultures or countries that are or have been at odds or at war with each other. In most cases you would not need to take this into consideration in your teaching, but there may be instances when it would be most respectful to do so, or instances when you can grasp the opportunity to make it a learning moment for all. For instance, in a history class, you may have students present

the same historical events from different cultural (national, regional, etc.) perspectives before you discuss the similarities and differences among these perspectives with them; this will help them appreciate that there is not one historical truth. One possibility would be to have students present from the perspective of their home culture, another is to make sure that all students present from a perspective that is not their own. Similar activities could be framed within different theories or schools of thought, different political parties, etc.

Your students may have different linguistic backgrounds: Many of the students will not have the language of instruction as their first language. While this may create challenges for them (see also under tip 4 below), you may exploit the fact that many of them will be able to read disciplinary literature in another language. This may be helpful in activities such as those mentioned in the previous paragraph above, but it may in fact also be an asset in a number of other types of individual or group assignments. The only thing is that the lecturer should be brave enough to let students use electronic or printed sources that s/he cannot check because s/he does not master all the languages present in the international classroom.

So the question remains how you obtain the information needed to take all these differences into account in the planning and teaching of your course. The most obvious solution is to consider which differences you would like to leverage in a particular course and then have the students fill in a short, targeted survey (e.g. on the learning management platform) just before or at the beginning of the course.

#### **Tip 4**

##### **Consider your students' learning needs**

Tip 3 above focused on how you may leverage the differences in your student cohort and create learning moments for them all. There is obviously a flipside to this coin. The differences among them may also mean that some students have learning needs that must be addressed if they are to successfully complete your course. Again, this may pertain to disciplinary content, methodologies, skills or language, and you will need to reflect on how you can assist your students in the process in a meaningful way (and without going way beyond the time you have been allotted for the course). What follows here assumes that the students have been screened appropriately (during the application and registration process) so that, generally speaking, they master the prerequisites for the programme or course.

If there is some disciplinary content that they need to cover or recapitulate, you may refer them to some supplementary literature if this need is not alleviated by working together in peer groups – or even if it is. When doing so, you should be mindful of the fact that the students in question will need to read this supplementary literature in a foreign language, and that this often takes extra time. So, if possible, it would be a good idea to (also) refer the students to literature in their own language.

As for methodologies, much can be learned by working with peers in the course of the academic term, but here supplementary reading may also be necessary. Whether or not students are already familiar with a given methodology, you should always make sure that they are given the opportunity to work with it in the course of the academic term so that they are prepared to do so when their work is assessed during or at the end of the academic term. The same applies to any skills, the mastery of which students have to demonstrate when their work is assessed. Practice makes perfect!

Finally, the language component. All students may struggle with academic reading and writing, even in their first languages, and academic literacy and academic writing are skills that must be learned – and therefore also taught – in any programme. In addition to that, some students may have language issues even though they have passed the language tests and demonstrated the minimum level of language proficiency required for a specific programme or course. If that is the case, it is important that you as the lecturer know where students can improve their language skills. In Denmark, most HEIs do not have a language centre for such purposes, which makes it even more important that all students master the language of instruction to a sufficient level before the beginning of the course.

### **Tip 5**

#### **Scaffold your students' learning: Focus on strengths and alleviate weaknesses**

Following directly from tips 3 and 4 above, it is important that you scaffold your students' learning. Help them appreciate their strengths and leverage them as indicated under tip 3, and help them alleviate their weaknesses by addressing their learning needs as indicated under tip 4. However, even if you do that, your students would still benefit from a robust structure (scaffold) to guide them through the course and support their learning. Again, such scaffolding may take different forms depending on the discipline or course, but an obvious example would be to provide support materials or develop them in collaboration with the students.

### **Tip 6**

#### **Be explicit in your communication and behaviour**

As already indicated in the above sections, students come with diverse prior knowledge and experience, and you can in fact not take anything for granted. While students may learn from each other in peer groups, it is still very important that you as the lecturer are very explicit in your communication and behaviour. You must tell the students exactly what is expected of them in the course of the semester and at the exams; you must also give them the opportunity to try it out while they are still learning – write in the specific genre that is required, use specific technology as required, take on roles in peer interaction, etc. Be aware that some students may be uncomfortable with your direct communication as they will be used to a much more indirect form of communication in their home cultures. If/when you observe that that is the case, take a moment to explain and discuss this with the students and help them overcome their uneasiness.

### **Tip 7**

#### **Consider the group dynamics: Turn challenges into opportunities**

There will be times when students – mobile and local – are challenged in the international classroom. Individual students or peer groups may encounter situations where their differences seem to become a barrier for them to move forward in their learning process. When you observe such a situation, you cannot ignore it. It does not just go away. You will have to stop your lecture or small class procedures and seize the opportunity to resolve the issue. If you do that and help the students resolve a (potential) conflict, you also help them learn disciplinary content at the same time as they develop their intercultural competences. The issue turns into a learning moment for them.

## Conclusion

It is especially important to keep group diversity and the group dynamics in mind when teaching in the international classroom. Your students are all learning – not only disciplinary content, but also peer group interaction. If you pay attention to it, plan your teaching accordingly, and address issues as they arise, you also help your students develop their intercultural competences, which will be valuable graduate attributes when they start their professional careers.

Two points will have become apparent in the sections above: The seven tips reflect what we would consider good teaching in any context, and they are interlinked and overlap to a certain extent. This is very much on purpose and also unavoidable. It is all higher education teaching and learning, irrespective of the characteristics of the student target group. The aim of this short DUT Guide has been to make the reader aware of some of the factors that are specific characteristics of the international classroom, and which the lecturer would be well advised to address when teaching international programmes to diverse student audiences.

## References

- Beelen, J. & E. Jones. 2015, "Redefining Internationalization at Home", in Curaj, A., L. Matei, R. Pricopie, J. Salmi & P. Scott (Eds), *The European Higher Education Area. Between critical reflections and future policies*. Springer.
- Carroll, J. 2015, *Tools for Teaching in an Educationally Mobile World*, Internationalization in Higher Education Series, Routledge.
- De Wit, H.; F. Hunter; L. Howard; E. Egron-Polak. 2015, "Internationalisation of Higher Education", Study requested by the European Parliament's Committee on Culture and Education, available at:  
[http://www.europarl.europa.eu/RegData/etudes/STUD/2015/540370/IPOL\\_STU\(2015\)540370\\_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2015/540370/IPOL_STU(2015)540370_EN.pdf)
- EQUiiP. Forthcoming/2019, "Educational Quality at Universities for inclusive international Programmes", available at: <http://www.EQUiiP.eu>.
- IntlUni. 2015, "The Challenges of the Multilingual and Multicultural Learning Space", available at: <http://www.IntlUni.eu>.
- Jones, E. 2017, "Problematizing and reimagining the notion of the 'international student experience'", *Studies in Higher Education*, 42.5.
- Lauridsen, K.M. & S.M. Cozart. 2015, "Teaching and Learning in the international classroom: quality principles and lessons learned from the IntlUni project", in *Internationalisation of Higher Education. An EAIE Handbook*, 2015,3. Raabe.
- Lauridsen, K.M. & Lillemose, M. K. (Eds). 2015, *Opportunities and challenges in the multilingual and multicultural learning space*, Final document of the IntlUni Erasmus Academic Network 2012-15, available at:  
[http://intluni.eu/uploads/media/The\\_opportunities\\_and\\_challenges\\_of\\_the\\_MMLS\\_Final\\_report\\_sept\\_2015.pdf](http://intluni.eu/uploads/media/The_opportunities_and_challenges_of_the_MMLS_Final_report_sept_2015.pdf)

Leask, B. 2015, *Internationalizing the Curriculum*, Internationalization in Higher Education Series. Routledge.

Leask, B. & Carroll, J. 2013. *Learning and Teaching Across Cultures. Good Practice Principles and Quick Guides*, available at: <http://www.ieaa.org.au/documents/item/397>

# Videnskabelse på universitetet. Veje til integration af forskning og undervisning.

Af Tine Damsholt, Hanne Nexø Jensen, Camilla Østerberg Rump (red.)

Udgivet 2018 på forlaget Samfundslitteratur, 241 sider, ISBN: 9788759330470

Anmeldt af Jakob Egholm Feldt<sup>a,1</sup>

<sup>a</sup>Forskningscenter for problemorienteret projektlæring, Roskilde Universitet

I bogens indledning citerer redaktørerne Københavns Universitets prorektor Thomas Bjørnholm for, at "forskningsbaseret undervisning er universitetets DNA" (s. 13). Det er sandt, i hvert fald i den Humboldt-inspirerede europæiske tradition, og vi er mange universitetsforskere, som glade vil nikke genkendende til det. Men hvad er det så? Hvordan gøres det? Hvordan beskriver man formelt, normativt og teoretisk forskningsbaseret undervisning på en tilstrækkeligt pluralistisk måde til, at beskrivelsen faktisk baserer sig på virkelighedens praksisser og samtidigt begrebsliggør for os noget mere alment, der kan give os bedre dømmekraft i forhold til praksisserne? Det er udfordringen for denne interessante og læseværdige bog, som giver os gode svar på ovennævnte spørgsmål: Det kan gøres gennem den forskningsbaserede undervisnings etnografi. Bogens redaktører og forfattere har ved hjælp af et omfattende etnografisk materiale, interviews med studerende og undervisere og undervisningsprogrammer på tværs af hovedområder på Københavns Universitet skabt interessante typologier over forskellige registre, som forskningsbaseret undervisning modtages i af de studerende og gives fra af underviserne.

Ofte er bøger om undervisning enten for praktiske; how-to modellen, hvor de hæmmer den selvstændige refleksion og modarbejder underviseren som "lærer", eller også er de så abstrakte, at de er et identitetsspejl for underviseren uden at have en praktisk virkning i forhold til undervisningen. Det er en svær genre. "Videnskabelse på universitetet" lykkedes dog godt i forhold til at opløse modsætningen, fordi den fremlægger en stor mængde data fra forskningsprojekterne bag, så forskellige sammenpasninger og uoverensstemmelser mellem registre blandt deltagere i forskningsbaseret undervisning kan studeres og medbedømmes af læseren. Bogen er således ikke en anvisning til bestemte handlinger, men en invitation til refleksion over praksis.

Bogen tager sit udgangspunkt i en dobbeltforpligtelse, som bliver sat til diskussion: forpligtelsen på den forskningsbaserede uddannelse på højeste internationale niveau og forpligtelsen på masseuniversitetet. Man kan mene, at højeste internationale niveau bliver sat af steder, hvor forpligtelsen på at få alle med er mindre end i Danmark. Bogens nok mest kontroversielle kapitel (kap. 7 af Tine Ravnsted-Larsen Reeh) beskriver, hvor megen modstand et projekt, der ville lave særlige talentforløb for de dygtigste studerende, mødte. Projektet ville sætte talentfulde studerende i kontakt med eksternt finansierede forskningsprojekter på Teologisk Fakultet i København. Tre forskellige forløb mellem talentfulde studerende (udvalgt på motiveret ansøgning og karakterer) og forskningsprojekter blev opsat: ét efter en semi-

---

<sup>1</sup> Kontakt: feldt@ruc.dk



narmodel, hvor studerende præsenterede deres arbejde og fik feedback; ét efter en mere traditionel undervisningsmodel, hvor forskere underviste i deres egen forskning, og ét efter en forskningslignende model, hvor de studerende deltog i forsknings(lignende) arbejde.

Blandt de væsentligste udfordringer viste sig at være tid og kønsforskel. De studerende (og forskerne) har ikke tid til at lave eksperimenter, arbejde med interesserede emner og til at tage fejl og lave om. Det er i virkeligheden en politisk og ledelsesmæssig falliterklæring, hvis man har bare den mindste ide om, hvordan kvalitet i videnskabelse opstår. I forhold til udfordringen med kønsforskellen viste det sig, at mandlige studerende havde markant lettere ved at se sig selv på et talentprogram, og de var mindre bange for at fejle, hvorimod kvindelige studerende havde svært ved at tale om sig selv som talentfulde trods gode papirer. Det er desværre en velkendt udfordring i forskningsverdenen, som, hvis den blev bedre løst, ville kunne frisætte en masse talent. De positive resultater af projektet er generelt interessante, men jeg vil fremhæve ét aspekt, nemlig det, at evalueringerne viste, at de studerende på talentkurserne oplevede en af-instrumentalisering, dvs. de oplevede at være drevet af nysgerighed og lyst. Modstanden mod projektet var rettet mod undervisningsdifferentieringen til fordel for de dygtigste og kom fra "universitetsbefolkningen", som forfatteren vagt kalder "dem", men som vel var ment som andre forskere og undervisere på teologi. Det er tankevækkende, at en logik, som er helt indforstået i forskningen, er kontroversiel i forhold til universitetsundervisningen. I undervisningen accepterer vi uden problemer at bruge resurser på differentiering til fordel for de svageste studerende. Projektet på teologi var baseret på Dreyfus og Dreyfus (1986), hvor antagelsen er, at læring *om* maksimalt kan føre til kompetence og kyndighed, men at "excellence/virtuositet" kræver kontekstuelle erfaringer, dvs. ikke-regelgjorte erfaringer baseret på en stor sum af gentagelser, som det ses i elitesport og musik. Det kan også kaldes intuition eller abduktive kompetencer.

Bogens tre mest interessante kapitler sammen med ovennævnte er kapitel 1-3, hvor typologierne og begreber sammenholdes med materialet og forskningslitteraturen. I kapitel 1 præsenteres forskellige perspektiver på forskningsbaseret undervisning kaldet FUI, forskning og undervisningsintegration. Der lægges især vægt på Changs (2005) beskrivelse af forskellige former for FUI, som også anvendes fast gennem bogen. Den første form er "det vejledte forskningsfællesskab", som er en idealtipe karakteriseret ved, at de studerende definerer problemstillinger, gennemfører en undersøgelse og skriver en afrapportering. De studerende deler viden og materialer med hinanden og lærer således via praktisering af forskningsprocesser, og de er aktivt videnskabende mere end viden-reproducerende. Det kan kaldes en forskningsproces, fordi forskning bliver betragtet som en praksis, en adfærdsform, og ikke en produktion af originale resultater.

De andre former er "naturtalentet", hvor talenter gives frihed og inspiration til at forløse deres talent – "underviser præsenterer sin egen forskning" og "slavearbejdet" – hvor studerende sættes til at udføre rutinearbejde for underviseren. Alle former kan noget og bliver brugt i undervisningen, men det er klart, at det vejledte forskningsfællesskab kræver mere pædagogisk planlægning og institutionel understøttelse. Samtidigt kan denne form også rumme de andre former i sig, hvis den bliver tænkt som noget mere end et aspekt af den enkelte undervisers kursusplanlægning. Det er værd at knytte en bemærkning til den for undervisere måske mest eftertragtede form for FUI, nemlig "underviser præsenter sin egen forskning". Her er metaforerne oftest rettet mod underviseren, som får god feedback fra de studerende, lærer noget ved at udtrykke sig klart og kan spille bold op ad de studerende.

Måske er denne form for FUI i virkeligheden den svageste; hvor meget får man ud af at være muren, når der spilles bold?

I kapitel 2 præsenteres resultaterne af en omfattende undersøgelse af de studerendes møde med FUI. Det er meget interessant at få præsenteret en typologi af studenterrationaler baseret på originalt interviewmateriale. Forfatterne (Tine Damsholt og Marie Sandberg) opstiller fire rationaler: det udviklingsorienterede, det joborienterede, det omsættelsesorienterede og det professionsorienterede. For flere af disse studenterrationaler opleves FUI faktisk som irrelevant. Især når undervisningen bliver fokuseret på underviserens egen forskning. Uden motiverede studerende virker FUI ikke, hvilket gør pluraliteten af rationaler i undervisningslokalet til en stor udfordring. Som forfatterne påpeger, skal FUI give mening på flere måder samtidigt (s. 55).

I kapitel 3 vendes blikket mod underviserne. Her opstilles også fire rationaler: adskillelsesrationalet (undervisning og forskning er helt forskellige ting), helhedsrationalet (undervisning og forskning flyder sammen), investorrationalen (undervisning skal give afkast i forskning) og læringsrationalet (de studerendes læring er i centrum). Her er det en vigtig pointe, at FUI i princippet ikke er vigtigt for rationale 1 og 4, fordi ved rationale 1 er det en utopi, og ved rationale 4 er undervisningen kun til for de studerendes lærings skyld. Igen er der mange tankevækkende udsagn fra informanterne, som flittigt gengives til læserens vurdering.

FUI er vigtigt, pointerer forfatterne, fordi "vi skal træne dem [de studerende] i at producere den viden, der skal til for at løse fremtidens problemstillinger" (s. 16). Således skal de studerende på universitetet kvalificeres til at skabe viden snarere end at konsumere den. Bogen er et grundigt og gennemtænkt bidrag til at kvalificere udviklingen af universitetets undervisningsrække, nemlig forskningsbaseret undervisning. I for mange sammenhænge på universitetet gælder det, bogen kalder læringsrationalet, som et sammenfald af ekspertviden om læring med økonomiske rationaler, men denne bog løfter blikket og sætter universitetets mål og kulturer i samspil med etnografisk materiale og et væld af konkrete eksempler.

## Referencer

- Chang, H. (2005). Turning an undergraduate class into a professional research community. *Teaching in Higher Education*, 10, 3
- Hubert L. Dreyfus & Stuart E. Dreyfus (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press

# Kig op! Undervisning uden et blåt skær

Skrevet af Peter Holdt Christensen

Udgivet 2018 på forlaget Samfundslitteratur, 184 sider, ISBN: 978-87-593-3207-8

Anmeldt af Kim J. Herrmann<sup>a,1</sup>

<sup>a</sup>Center for Undervisning og Læring, Aarhus Universitet

Der findes næppe den underviser på universitetet eller andre videregående uddannelser, som ikke på et eller andet tidspunkt står med det problem, at en del af de studerende – det være sig få eller mange – er fraværende i undervisningen, fordi de læser mails, skriver på Facebook eller tjekker nyheder på nettet. I hvert fald har jeg endnu til gode at møde en underviser, som ikke har en holdning til studerendes brug af bærbare og mobiltelefoner i undervisningslokalet.

Og netop holdninger er der mange af, når det gælder studerendes brug af digitale medier. Til gengæld skorter det ifølge Peter Holdt Christensen på fakta og på aktiv stillingtagen. Derfor har han for nylig udgivet bogen *Kig op! Undervisning uden et blåt skær* med det formål at skrive en samlet fremstilling af, hvorfor og hvordan skærmtid udgør et problem for undervisningen. Dels for at muliggøre en saglig diskussion af konsekvenserne af skærmtid i undervisningen, dels for at gøre det muligt for undervisere og uddannelsessteder at tage stilling og handle.

## Skærmtid

Men først tilbage til hvad bogen handler om og ikke handler om. Christensen er interesseret i fænomenet 'skærmtid', hvilket han definerer som tiden brugt på digitale aktiviteter, hvor disse aktiviteter ikke er relevante for den interaktion som foregår i undervisningslokalet. Bogen handler således ikke om digitale aktiviteter uden for det fysiske undervisningsrum, ej heller handler bogen om digitale aktiviteter, som er relevante for undervisningen, som fx at tage noter på en bærbar.

Og det er en del af pointen. At studerendes brug af digitale medier i undervisningslokalet ikke er et problem *per se*, men et problem *de facto*. Dette har dels at gøre med, at mobile apparater og internetadgang er en helt selvfølgelig del af det moderne undervisningslokale, dels at mange mennesker (heriblandt studerende) har udviklet noget, der nærmer sig en afhængighed af nyheder og sociale medier, og denne afhængighed er rykket med ind i undervisningslokalet. Symptomerne er et stigende antal studerende, der har svært ved at koncentrere og fordybe sig, som multitasker uhensigtsmæssigt, og som har svært ved at byde ind i den faglige samtale. Konsekvenserne for den enkelte studerende er et ringere udbytte og lavere karakterer. For den enkelte underviser er konsekvensen, at det bliver sværere at afkode de studerendes signaler og dermed justere tempo og retning i undervisningen, ligesom mange undervisere oplever at blive demotiverede.

---

<sup>1</sup> Kontakt: kh@au.dk

## Energifyldte interaktioner

Men – og det er pointen – problemet er i virkeligheden ikke individuelt. Det er først og fremmest kollektivt. Og for at forstå det, gør Christensen sig umage med at forklare, hvordan vi kan forstå værdien af det *fysiske* møde mellem underviser og studerende og studerende imellem.

Christensen forklarer med henvisning til sociologen Randall Collins, at når vi som mennesker mødes i en fysisk forsamling, har vi mulighed for at indgå i såkaldte energifyldte interaktioner. 'Energien' opstår, fordi vi er fysisk tilstedeværende, fordi det fysiske rum afgrænser os fra andre, fordi vi er gensidigt opmærksomme, og fordi vi i fællesskab oplever følelser. Er disse fire betingelser opfyldt, har vi mulighed for at opleve energifyldte interaktioner, som er afgørende for vores motivation og engagement. Det samme gælder for mødet i det fysiske undervisningslokale.

Det fysiske møde kan dog lige så vel dræne os for energi i de tilfælde, hvor disse betingelser ikke er opfyldt, og det er netop her, skærmtid udgør en stigende trussel. Den lette adgang til underholdning på nettet gør nemlig, at det i dag er meget nemmere at undslippe det fysiske samvær og flygte ind i det digitale rum, som tilmed er konstrueret med det formål at give os hurtige belønninger, mens fordybelse og engagement i en faglig samtale er meget mere krævende og belønningen meget mere langsom.

Pointen er følgende: Det fysiske møde mellem underviser og studerende rummer – og har altid rummet – et afgørende potentiale i form af energi og motivation for de studerendes læring. Men hvor det før i tiden var meget svært at undslippe den fysiske forsamling, når først vi valgte at tage del i den, så er det i dag meget nemt og uendelig fristende at trække sig tilbage. Og dermed risikerer vi at undergrave forudsætningerne for, at undervisning fylder os med energi og motiverer os til at lære. Konsekvensen er, at det er blevet væsentligt sværere at forløse undervisningens vigtige potentiale for de studerendes læring.

## Religionskrig og reaktans

Når problemet med skærmtid er så stort og berører så mange, er det oplagt at spørge, hvorfor der ikke er flere, der handler og tager stilling. Svaret involverer psykologi, ideologi og individuelle interesser. Psykologi, fordi vi som mennesker sjældent er opmærksomme på, at vi er uopmærksomme, og fordi vi som mennesker per reflex yder modstand, når vi oplever, at andre vil begrænse vores frihed (reaktans). Ideologi, fordi der ifølge Christensen på mange uddannelsessteder hersker en teknologioptimisme, hvorved man med kritik af teknologien risikerer at udløse en religionskrig, hvor man stemples som teknologifjendsk. Individuelle interesser, fordi det for den enkelte underviser er forbundet med hurtig og sikker konflikt og langsigtet og usikker gevinst at tage kampen op. Intet er dermed mere oplagt end ikke at gøre noget.

Sidstnævnte – ikke at gøre noget – er dog ikke en pædagogisk acceptabel mulighed, for Christensens hovedpointe er følgende: Den digitale adfærd påvirker de aktiviteter og den læring, der foregår på uddannelsesstedet, og derfor er skærmtid en pædagogisk udfordring, der skal tages stilling til. Bogens sidste kapitel er således en diskussion af, hvad der kan gøres, og under hvilke omstændigheder vi kan gøre os håb om at begrænse eller endda helt forbyde skærmtid.

## En diskussion på baggrund af fakta

Christensens bog udmærker sig efter min vurdering på en række punkter, hvoraf jeg vil fremhæve tre.

For det første er det et helt afgørende bidrag, at Christensen på et teoretisk og empirisk grundlag beskriver, hvorfor og hvordan skærmtid påvirker undervisning i det fysiske rum. Debatten om Facebook i undervisningen har raset lige siden de første bærbare fandt vej til auditorierne i starten af årtusindskiftet, men karakteristisk for denne debat har været, at den har været baseret på holdninger, og at det hurtigt er blevet en skyttegravskrig mellem undervisere, der beskylder studerende for at være dovne, og studerende, der beskylder undervisere for at være kedelige. Det er derfor mit håb, at vi med Christensens bog kan gå ind i en ny fase, hvor vi kan diskutere problemerne på baggrund af viden og en fælles forståelse af, hvorfor det er så svært at lade mobilen ligge i tasken – uanset om vi som studerende deltager i undervisning, eller om vi som undervisere deltager i et møde.

Et andet væsentligt bidrag er, at Christensen formår at formulere, hvad der er det unikke og særlige ved det fysiske møde mellem undervisere og studerende. Selv i en tid med store fremskridt inden for digital undervisning har de fleste undervisere nok stadig en fornemmelse af, at mødet med de studerende, ansigt til ansigt, er vigtigt og værdifuldt. De færreste vil dog nok kunne begrunde hvorfor. Og det er netop begrundelser der skal til, hvis man vil gøre sig håb om at begrænse skærmtid.

I forlængelse heraf er det, for det tredje, et stort bidrag, at Christensen med brug af teori og empiri formår at forklare, at undervisning i sin essens er et fælles anliggende. Et samarbejde. En samtale mellem flere parter. Dermed punkterer han effektivt den myte, at man kan deltage i undervisning og samtidig være på nettet med den begrundelse, at "det går jo ikke ud over andre end mig selv". Jo, det gør det! For som Christensen dokumenterer, er det umuligt at være til stede sammen uden at påvirke hinanden. Det er jo det, der netop er pointen ved at mødes. Ingen studerende er isolerede øer, og hvis det store potentiale i det fysiske møde skal forløses, så kræver det, at alle (eller i hvert fald langt de fleste) er fysisk til stede og mentalt nærværende.

### Et fælles ansvar

Man kan diskutere, om de fortrinsvis amerikanske økonomiske, sociologiske og psykologiske teorier, som Christensen referer til, nødvendigvis og i alle henseender er gyldige i konteksten af en dansk videregående uddannelse. Man kan også diskutere, om Christensen, trods ihærdige forsøg, undgår at falde i den fælde at tegne et lettere stereotyp billede af de studerende.

Dette ændrer dog ikke på, at bogen *Kig op!* er anbefalelsesværdig til et meget bredt publikum, og at Christensens hovedpointe er svær at afvise:

*Fordi skærmtid er et fælles anliggende, mener jeg [...] at det er uansvarligt, hvis ikke uddannelsessteder, fagansvarlige, undervisere og studerende sørger for at diskutere rammerne for brugen af skærme. Et frugtbart undervisningsmiljø kræver nemlig, at vi i perioder også er sammen uden et blåt skær. (s. 49)*

Den anbefaling er hermed givet videre.

# 1968: Studenteroprør og undervisningsrevolution. En fortælling om opgøret med traditionel universitets- undervisning

Skrevet af Finn Hansson

Udgivet 2018 på Roskilde Universitetsforlag, 117 sider, ISBN: 978-87-7867-512-5

Anmeldt af Birger Steen Nielsen<sup>a,1</sup>

<sup>a</sup>Institut for Mennesker og Teknologi, Roskilde Universitet

Finn Hanssons bidrag til historieskrivningen om 68-oprøret (68) og dets virkninger skiller sig ud fra de ikke så få andre udgivelser i anledning af 50-året for 68. Den er som en række andre bidrag nok skrevet indefra – Finn Hansson var selv en central deltager i oprøret og har en lang universitetskarriere bag sig, der førte ham fra Sociologisk Institut på Københavns Universitet (KU) til Copenhagen Business School (CBS), hvor han bl.a. arbejdede som forskningsbibliotekar – og den trækker også på hans egne erindringer, men den kombinerer dette med et stædigt gravearbejde i bedste sociologisk tradition: Han har været vidt omkring og dybt nede i tilgængeligt kildemateriale, og det har gjort det muligt for ham at fortælle en *anden* historie om 68 end den gængse.

Det er historien om, hvordan opgøret med de stivnede og perspektivløse undervisningsformer og de måder at studere på, der prægede universitetet dengang, ikke bare var en vigtig side af oprøret, men også må ses som et omdrejningspunkt for det – en drivkraft – der dengang havde afgørende betydning for oprørets indre organisering og sammenhængskraft. En udvikling som derudover blev grundlaget for en pædagogisk og studiemæssig model, der har rakt langt ud over det kortvarige oprør i sig selv – med Hanssons indledende ord:

*Alle steder i det danske uddannelsessystem arbejder man i dag med gruppearbejde og projekter. Denne pædagogiske model har eksisteret i mange år og har [...] vist sig at være ganske sejlivet. Men hvordan opstod den egentlig? Hvor kom den fra? Og hvordan blev den så udbredt? (s. 9)*

## Gruppe- og projektarbejde: Hvor kom det fra?

At 68 gav anledning til, at undervisningsformerne – mere eller mindre og med forskellig rækkevidde på de forskellige fag – blev grundlæggende fornyet, er ikke en ny iagttagelse. Men i Hanssons rekonstruktion af, hvordan det fandt sted, vil jeg fremhæve to dimensioner, der i særlig grad falder i øjnene. *Den første dimension* angår det andet af hans tre spørgsmål: Hvor kom denne pædagogik fra?

Bogens rekonstruktion er her rigt facetteret. Jeg hæfter mig ved, at det lykkes Hansson at vise, hvordan de pædagogiske forsøg på nyorientering kun kan forstås, hvis man ser, hvordan studenteroprøret var en del af både det langt mere omfattende ungdomsoprør med dets eksperimenteren med nye måder at mødes og leve sammen på – kollektive modkulturer

---

<sup>1</sup> Kontakt: bsn@ruc.dk

– samtidig med, at der var stærke reformpædagogiske traditioner at trække på – traditioner, der på deres side var inspirationskilder for ungdomsoprøret og blev aktiveret og nyfortolket gennem det. Der henvises til denne sidste side, men det er først og fremmest den første, Hansson forbilledligt beskæftiger sig med. Herigennem bidrager hans fortælling – ud over dens hovedærinde – også til at tegne et langt mere nuanceret billede af "det politiske 68" end det, der fortrinsvis verserer. Ikke bare i betydningen, at "det politiske 68" var *mere* end politik, men mere grundlæggende i den forstand, at "det politiske" var uløseligt forbundet med forsøget på at skabe nye måder at leve på. I det perspektiv er det et scoop, at hans hoved-"case" er det Sociologiske Institut, der vel nok står som eksemplet på den radikale, marxistisk orienterede studenterbevægelse.

"Det alternative", orienteringen mod nye livsformer og "det politiske" – stærkest markeret i protesterne mod Vietnamkrigen og den teoretiske orientering mod en ny, kritisk marxisme – løb sammen i det, der var bevægelsens centrale kendetegn: det antiautoritære, positivt formuleret: ønsket om autonomi. Man ønskede frie, selvbestemte undervisnings- og studieformer, fordi de gamle ikke var til at ånde i, men også for at kunne gøre studierne "samfundsrelevante": kritiske. Denne bestræbelse udløste uanede energier og opfindsomhed. Der blev arbejdet hårdt, både med omformningen af undervisnings- og studieformerne og med de kritiske *faglige* studier og undersøgelser.

### **Hvordan kunne den nye pædagogik slå igennem?**

Den anden dimension, jeg vil fremhæve, besvarer Hanssons tredje – ovenstående – spørgsmål. Der blev lagt masser af arbejde i selvorganiseringen af studium og undervisning, men samtidig også med udformningen af nye studieordninger. Det var år med en intensiv mødeaktivitet, stormøder, gruppemøder, men også studienævns- og institutmøder. Det giver bogen et levende billede af – et billede, der samtidig viser, hvordan det forstokkede og autoritære institutmiljø, der generelt kendetegnede universitetet, men i *særlig* grad Sociologisk Institut, også udgjorde en selvstændig drivkraft til oprøret – en drivkraft, der havde form af oprør, modstand og nok også med elementer af militant uforsonlighed, som ikke uden videre harmonerede med frihedsbestræbelsen. At karakteristikken af datidens Sociologisk Institut som autoritært og forstokket ikke blot er et partsindlæg, viser bogen – gennem sit kildemateriale, herunder Rektor Mogens Fogs udtalelser og indgreb – overbevisende.

Uden studenternes ihærdighed og vedholdende engagement kunne den nye pædagogik ikke have slået igennem. Det er imidlertid bogens store fortjeneste – som historieskrivning og fortolkningsramme til at forstå denne vidtgående *samfundsmæssige* reformproces – at den viser, hvordan studenternes bestræbelser "nedefra", uden at de vidste af det, imødekom – delvist strandede – bestræbelser fra centraladministrationen, der ønskede et opgør med universiteternes ældgamle autonomi, som blev anset for blokerende for moderniseringen af det danske samfund, hvor netop uddannelsessektoren blev betragtet som et væsentligt moment i opbygningen af vækst- og velfærdssamfundet. 68-oprøret blev grebet som en anledning til at få gang i denne reformproces. Det er en spændende og tankevækkende historie, som jeg kun kan anbefale, man selv læser efter i bogen.

Resultatet blev jo ikke bare en uskadeliggørelse af oprørets bestræbelser, men en vidtgående institutionalisering af dem i former, der ganske vist var modsigelsesfulde, men altså også indebar en virkelig fornyelse og åbning af universitetet, som af de studerende (os) dengang nok blev betragtet som mangelfulde, men også som en sejr. Universiteternes nærmest feudale autonomi – professorvældet – var en stopklods for en fri og kritisk videnskab, og uden

kompromiset med "statsmagten" var en reformering ikke lykkedes. Men set i et aktuelt perspektiv var åbningen for den statslige intervention i universiteternes styre jo også et tveægget sværd. I dag står etableringen af en stærkere autonomi på universiteterne igen på dagsordenen, men i en helt anden politisk-økonomisk konstellation, hvor en kulturkonservativ tilbagevenden til det traditionelle autonome universitet, selv hvis den kunne lade sig gøre, ville være perspektivløs.

I studenternes perspektiv beskriver Hansson arbejdet med at udforme og bidrage til indførelsen af den nye pædagogik som en omfattende læreproces, bestemt af lige dele opfindsomhed og "administrativt håndværk". Især fremhæver han, hvordan det var nødvendigt, at man lærte sig at bruge et administrativt sprog, der kunne ligge langt fra de aktivistiske former for retorik, der jo også prægede oprøret. Det ligner det, som Knud Illeris senere kaldte en "dobbelt-kvalificering", og betydningen heraf vurderes i bogen højt og beskrives bl.a. i forhold til indførelsen af gruppe- og projektarbejde ved de nyoprettede universitetscentre. Det peger frem mod den indflydelse, som studenterrådene og DSF i mange årtier har kultiveret og udøvet. Men måske gås der lidt for hurtigt uden om tvetydighederne i denne form for dobbeltkvalificering?

### **Oprørets universitetspædagogiske frugter – i dag**

Bogens afsæt er, som forordet beskriver, at "alle steder i det danske uddannelsessystem arbejder man i dag med gruppearbejde og projekter". 68 lever videre. Men hvis man skulle ønske sig mere, end det bogen giver, kunne det være en mere indgående diskussion af denne vurdering. I hvert fald inviterer den til det. Hansson åbner selv denne diskussion, men udfolder den ikke. Hans ærinde har også væsentligt været et stykke korrigerende, sociologisk historieskrivning. En nærmere beskæftigelse med selve gruppe- og projektpædagogikken med sammenhængen mellem læringsformer og kritisk erkendelse er der ikke. Det kunne måske have givet anledning til udfoldelse af diskussionen om denne pædagogiks status i dag.

For der er gode grunde til at spørge, hvor meget der egentlig er tilbage af det kritiske og skabende projekt- og gruppearbejde fra 68. Måske ikke så meget. Hansson nævner selv, at universitetsdemokratiet er blevet afskaffet, men det rækker jo langt ind i undervisnings- og studieformer. Rekonstruktionen af den nye pædagogiks opkomst viser bl.a. også, at dens centrum og indre drivkraft var knyttet til studenternes stræben efter autonomi og ønsket om, at deres studier skulle have kritisk samfundsrelevans. Demokratisering af universiteterne var uadskilleligt forbundet hermed. I dag er det institutionelle perspektiv snarere "samfundsrelevans uden kritik". Og hvor "selvvalgt" er tilgangen i "grupperarbejde og projekter" i dag i "det danske uddannelsessystem"? Dette kan man selvfølgelig – på universiteterne, i hvert fald – søge at træde op imod. Og det sker jo også. Men her virker den stærke beskæring af studietid, kravet om hurtigere gennemførelse og ønsket om kortere studier kontraproduktivt.

Studerende og lærere kan stadig i et vist omfang finde omveje, tage sig tid og gå imod hovedtendensen. Og som bogens historie viser, er modstand ikke den ringeste drivkraft. Man skal heller ikke glemme, at projekt- og gruppearbejde stadig *i sig selv* har en iboende tendens mod at åbne for autonomi i læreprocesserne. Det er en pædagogik, der ikke fuldstændig kan teknokratiseres, men som også kan opfordre til "usædvanlige læreprocesser", som det engang hed. Også i dag findes der hos studenter og unge mennesker i det hele taget – på ny og måske i voksende omfang – stærke ønsker om et rigtigt og meningsfuldt liv og om en uddannelse, hvor man kan beskæftige sig med noget vigtigt; noget der *betyder* noget. Og sådanne øn-



sker kan netop 68-pædagogikken komme i møde. Det kan Finn Hanssons storartede bog også minde os om.

# DANSK UNIVERSITETSPÆDAGOGISK TIDSSKRIFT

Årgang 14 Nummer 26/2019

Titel: Fra data til beslutninger

Sponsoreres af Dansk Universitetspædagogisk Netværk

## REDAKTION

Kim Jesper Herrmann, Center for Undervisning og Læring (AU) (ansvarshavende)

Julie Borup Jensen, Institut for Læring og Filosofi (AAU)

Sofie Kobayashi, Institut for Naturfagenes Didaktik (KU)

Katrine Lindvig, Institut for Naturfagenes Didaktik (KU)

Bolette Rye Mønsted, Institut for Læring og Filosofi (AAU)

Lotte Dyhrberg O'Neill, SDU Universitetspædagogik (SDU)

Kamilla Pedersen, Center for Sundhedsvidenskabelige Uddannelser (AU)

Bo Skøtt, Institut for Design og Kommunikation (SDU)

Rie Troelsen, SDU Universitetspædagogik (SDU)