

Hvordan kan kriterier for ph.d.-bedømmelser danne grundlag for ph.d.-uddannelse?

Pædagogiske perspektiver af en analyse af 41 ph.d.-bedømmelser fra et humanistisk fakultet

Lotte Rienecker, LearningLab, Danmarks Tekniske Universitet.

Peter Stray Jørgensen, forlagskonsulent og universitetspædagogisk freelancer.

Reviewet artikel

Pointen i denne artikel er at indholdet i ph.d.-bedømmelser kan og bør have pædagogiske implikationer for ph.d.-skrivning og ph.d.-vejledning. For at skabe grundlag for sådanne implikationer har vi analyseret 41 ph.d. bedømmelser fra 2009 på Det Humanistiske Fakultet med henblik på at identificere styrker og svagheder i afhandlingerne. Analysen er baseret på international empirisk forskning i ph.d.-bedømmelser. Analysen viste tydelige mønstre i bedømmelserne: Mens afhandlingernes analysedel vurderes meget positivt, samler kritikken sig om de afsnit hvor afhandlinger skal trække de store linjer op: forskningsoversigt hvor eget bidrag skal positioneres i forhold til feltets litteratur, og diskussionen hvor analysens resultater skal sættes i sammenhæng med det øvrige felts teorier og metoder. Vi afslutter artiklen med forslag til ph.d.-studerende, vejledere/undervisere, bedømmere og ph.d.-administration og -forskerskole.

Problemstillingen

Den ph.d.-bedømmelsesforskning som vi tager udgangspunkt i, er her repræsenteret af en tværfakultær empirisk bedømmelsesforskning af amerikaneren Lovitts, 2007 og en række store, tværfakultære undersøgelser af den australske forskergruppe SORTI samt den efterhånden omfattende lærebogslitteratur til ph.d.-skrivere og vejledere /bedømmere om afhandlingsskrivning set i et alignment perspektiv (fx Trafford & Leshem 2008, Murray 2008). Forskningen studerer bedømmelserne med henblik på at gøre ph.d.-studerende og deres vejledere opmærksomme på hvad "der kræves".

Forskning i ph.d.-bedømmelser

Den ph.d.-bedømmelsesforskning som vi tager udgangspunkt i, er her repræsenteret af en tværfakultær empirisk bedømmelsesforskning af amerikaneren Lovitts, 2007 og en række store, tværfakultære undersøgelser af den australske forskergruppe SORTI samt den efterhånden omfattende lærebogslitteratur til ph.d.-skrivere og vejledere/bedømmere om afhandlingsskrivning set i et alignment perspektiv (fx Trafford & Leshem 2008, Murray 2008). Forskningen studerer bedømmelserne med henblik på at gøre ph.d.-studerende og deres vejledere opmærksomme på hvad "der kræves".

En dansk undersøgelse

Den eneste foreliggende danske undersøgelse til nu, ud over denne artikel, er jura-professor Schaumburg-Müllers (2009) korte gennemgang af 79 ph.d.- og doktorbedømmelser fra jura, Aarhus Universitet 2000-2006. Han fandt at bedømmelserne vidner om mange problemer i afhandlingsskrivning, problemer som han bemærker at man skulle mene var basale. Han bemærker manglende fokus, problemformulering og rød tråd og også mangel på teoriudvikling og kritisk forholden sig som hyppige problemer i antagne afhandlinger, ifølge bedømmerne (2009, s.1). Schaumburg-Müllers optællingsresultater og metode fremgår ikke af artiklen, mens der bruges ord som "mange" og "en del" om ovenstående indtryk. Den internationale bedømmelsesforskning og lærebogslitteratur om ph.d.-afhandlinger handler om at eksplicite det som ellers kan være tavs viden for vejledere, bedømmere og ph.d.-skrivere. Som Schaumburg-Müller beskriver det:

"På denne måde kan kravene til videnskabelighed i [...] afhandlinger i nogen grad betegnes som 'tavs viden', en viden som er i miljøet, men som ingen ekspliciterer offentligt. Her lades især de ph.d.-studerende unødigt i stikken, fordi der ikke er klare udmeldinger om bedømmelseskriterier, og de overlades til at gætte sig frem til hvad et nærmere bestemt bedømmelsesudvalg vil mene. I værste fald lægges der forskellige kriterier til grund, og en sådan uoverensstemmende praksis fanges ikke, fordi man i begrænset omfang læser andre bedømmelser end dem man selv og eventuelt nogle nære kolleger er involverede i." (2010, s.3).

Resultaterne af vores undersøgelse viste lidt andre udfordringer end dem som Schaumburg-Müller fandt. Nærværende undersøgelse blev igangsat på initiativ af en (dengang) forskerskoleleder som bemærkede mønstre i bedømmelsernes kritik, og at observationen igennem flere år var at mange bedømmelser var meget kritiske, også overfor antagne afhandlinger. Man ønskede at bruge bedømmelserne pædagogisk og fremadrettet. Pædagogisk interessant er det derfor hvad man kan lære af især de kritiske bemærkninger.

Formål

Med undersøgelsen her vil vi vise hvilke kriterier for bedømmelser af ph.d.-afhandlinger man kan analysere sig frem til i en årgang af bedømmelser på ét fakultet. Specifikt analyserer vi hvad der i ph.d.-afhandlingerne *ikke* har levet op til bedømmernes kriterier, og dermed peger vi på vejledningspædagogiske muligheder og vigtige temaer for vejledning, forskeruddannelse og information ud fra en *alignment* tankegang (Biggs & Tang, 2011). Metodisk er denne undersøgelses bidrag et analyse-skema for ph.d.-bedømmelser som kan danne afsæt for kommende analyser af ph.d.-bedømmelser til brug for udvikling af forskeruddannelse og vejledning.

Data

Data er 41 bedømmelser af humanistiske ph.d.-afhandlinger på ét fakultet, nemlig samtlige bedømmelser for ét år. Heraf var de 38 antaget til forsvar, dog 6 efter omskrivning på 6 måneder. Ud af disse er to afvist uden mulighed for at revidere, én efter omskrivning, men som afvises i anden omgang. Samplet er givet af at undersøgelsen blev foretaget på foranledning af ledelsen; der er derfor tale om en intern undersøgelse. Bedømmelser er normalt fortrolige, og det er ekstraordinært at vi har fået lov til at læse bedømmelserne i papirversion som efter dataindtastningen blev maku-leret.

Teori: Kriterier for bedømmelse af ph.d.-afhandlinger

Læringsmål for ph.d.-uddannelse kan udlæses af Qualification Framework of the European Higher Education Area som bl.a. nævner systematisk forståelse af et felt, forskningsmetode, forskningsbidrag som kan publiceres, kritisk analyse, vurdering og syntese af nye ideer (<http://www.grossroads.eu/qualifications-framework-of-the-european-higher-education-area/third-cycle-phd>). Udover internationale qualification frameworks har lokale forskerskoler retningslinjer for kriterier og bedømmelsers udformning. Alligevel er det et fund i bedømmelsesforskningen at der er usikkerhed om kriterier for ph.d.-bedømmelser, at de er forskellige og at eksisterende retningslinjer ikke bliver brugt konsistent (Holbrook et al. 2004). Nogle bedømmelser kommenterer fx på anvendte metoder, andre ikke. Forskellige bedømmelseskriterier kan gøre ph.d.-bedømmelser svært sammenlignelige og underminere tillid til bedømmelserne.

Kritiske vurderinger sætter kriterier

Kritiske, negative vurderinger er et omdrejningspunkt for bedømmelsesforskningen, og specielt hvilke kriterier der bliver afgørende og har *make-or-break*-karakter for afhandlingens antagelse. Store tværfakultære studier af flere tusinde bedømmelser, foretaget af Holbrook et al. 2007, og Holbrook et al. 2004, har vist at *make-or-break*-karakter har i særdeleshed kandidaternes litteraturkendskab – og kvalificering (*coverage* → *critical appraisal*). Hvis bedømmerne finder det "i orden", er sandsynlighe-

den for godkendelse meget høj. Det drejer sig ikke kun om bredde, men om dybde, om sætten i relation og forståelse og anerkendelse af kontekster, om kompetencen at kunne reviewe feltets litteratur. Også tilstedeværelsen af et klart bidrag i en ph.d.-afhandling er vigtigt for at opnå graden – men mange afhandlinger antages dog uden at der i bedømmelsen eksplicit nævnes et bidrag (*contribution*) – med dette ord eller med synonymer. Holbrook et al. skriver om kandidatens *scholarship* (belæsthed) i forhold til kandidatens *research* (egen undersøgelse) som to piller i hvad der bedømmes, og de har fundet at *scholarship* i mange bedømmelser beskrives som mere udslagsgivende end om kandidatens egne undersøgelser har ført til resultater, eller er helt metodisk velgennemførte.

De kriterier som undersøgelserne foretaget af Holbrook et al, 2007, respektive Lovitts, 2007, har angivet som de vigtigste (bl.a. Holbrook et al.s overkategorier *scholarship* og *research*, og kategorien *contribution*) har dannet baggrund for vores analyse-skema, og de er blevet holdt op imod vores danske empiri.

Vi har endvidere brugt fakultetets interne skrift Retningslinjer for bedømmelse (2009), idet der heri angives et sæt kriterier, nemlig at bedømmelserne skal afspejle afhandlingens stærke og svage sider, og der ønskes vurdering af afhandlingens formål, bidrag med ny viden og originalitet, empirisk grundlag, samt de(n) valgte metodiske og teoretiske tilgang(e), resultater og konklusion, struktur, præsentationsform, akribi og akkuratesse Disse kriterier genfindes i det analyseskema vi har konstrueret (bilag 1).

Metode

Data var givne, vi har derfor ikke foretaget nogen dataudvælgelse.

Materialet; de 41 bedømmelser udgjordes af 7-14 sider lange dokumenter. Første del af disse er en nøje gennemgang af afhandlingens argumentation og dokumentation. Anden del består af vurderende kommentarer. Ofte er sidste del den længste. Det er kun de vurderende kommentarer vi analyserer her.

Skemadesign

Hovedredskabet i analysen er et skema til registrering af de vurderende kommentarer i bedømmelserne.

Kategorierne i analyseskemaet har vi konstrueret gennem en kombination af input fra henholdsvis empirien, dvs. bedømmelserne, og teorien. Kategori 1 er bidraget som al litteratur og vejledninger på området, bl.a. Qualification Framework fremhæver som et afgørende kriterium, kategori 2-10 udgøres af tekstelementerne i afhandlingens standardstrukturer (se fx Feak & Swales, 2009, s.61), kategori 11-14 er kategorier som tilsammen udgør afhandlingens fremstilling, (eng. *presentation*). Fremstilling går på tværs af de enkelte tekstelementer.

Afhandlingens indholds- og tekstelementer der vurderes i bedømmelserne

	Positivt	Negativt
1. Nybrud, bidrag		
2. Indledning med problemformulering og begrebsdefinitioner, afgrænsning m.m.		
3. Forskningsoversigt, kontekstualisering, faghistorie		
4. Metode (og metodekritik)		
5. Empiri		
6. Teori		
7. Analyse(r)		
8. Diskussion		
9. Resultater, konklusion		
10. Perspektivering, med samfundsrelevans, kontekstualisering		
FREMSTILLING		
11. Argumentation		
12. Struktur		
13. Positionering af skriveren, med kritisk distance, personligt engagement		
14. Sprog og formalia, tabeller, figurer mv.		

Figur 1: Skema til registrering af kommentarer i bedømmelserne

Bedømmelsesforskningen byder på forskellige kategorier at analysere bedømmelsens vurderende udsagn i. Vi har valgt at kategorisere i standardstruktur-modellen, fordi en model der kategoriserer i velkendte dele af afhandlingen letter en pædagogisk overførbare på skrive- og vejledningssituationen: Klager mange bedømmelser fx over manglende metodeafsnit, så er det brugbare informationer for skrivere, vejledere og andre.

Vi har fundet kategorierne velegnede. Kun ganske få kommentarer i vores materiale kan vi ikke rubricere som en kommentar til et tekstelement eller til fremstillingen. Kategoriernes grænser er ikke præcise, nogle kommentarer kan placeres flere steder, fx kan kommentarer om bidrag også, med anden formulering, placeres under konklusion og perspektivering.

Kategorierne opdeler vi i to grupper. I tråd med Lovitts (2007) og Holbrook et al. (2007) opdeler vi alle vurderende kommentarer i "positive" og "negative". Kun i et fåtal af tilfælde har det været svært at afgøre om en kommentar skulle forstås som negativt eller positivt vurderende, det fremgår næsten altid af konteksten.

Analysemetode

Hver bedømmelse er blevet registreret i sit eget individuelle skema ved at indtaste samtlige vurderende sætninger i analyseskemaet. Da vi ikke havde elektroniske versioner, har vi skrevet kommentarerne ind fra papirversion, i praksis en indskrivning af næsten al tekst i den vurderende del. Derefter blev alle kommentarer i de 41 bedømmelser til hvert af punkterne i analyseskemaet samlet i dokumenter med afsnitenes titler, dvs. et skema for samtlige kommentarer til Bidrag, Empiri, Metode osv.

Der er fremstillet 41 skemaer til registrering af hver bedømmelse, og 14 skemaer til hver af kategorierne, opdelt i grupperne Positivt og Negativt. Således kan ros og kritik til hvert tekstelement ses samlet. Vi har begge læst alle bedømmelser og alle bedømmelsesskemaer for gensidigt tjek af placering af de enkelte udsagn i de enkelte tekstelementer. Tvivl om placeringen af et bedømmelsesudsagn kan opstå hvis bedømmernes sproghandling ikke er klar. Fx kan en sekvens hvor der spørges til afhandlingens argumentation, men ikke vurderes/kritiseres, opfattes som en kritisk vurderende sekvens, men også som en neutral, spørgende sekvens der foregriber spørgsmål til det mundtlige forsvar. Vi har i sådanne tilfælde undladt at medtage enkelte passager fra bedømmelserne i skemaerne som vi ikke begge kunnet afkode som vurderende, positivt eller negativt. Således har enighed om placering i skemaet været en forudsætning for inklusion i registreringskemaerne.

Hvert af de 14 tekstelementer i de 14 skemaer har vi analyseret ud fra hvilke kriterier for god afhandlingskrivning der kommenteres på af bedømmerne, dvs. hvilke kriterier der er genkommende. Resultaterne af denne detaljerede sproglige analyse af bl.a. anvendte adjektiver kan ses i bilag 1 og i undersøgelsen (Rienecker & Stray Jørgensen 2011, bilag 2).

De 14 tekstelementskeemaer giver overblik over omfang og indhold af de vurderende kommentarer til hvert felt, samt forholdet mellem ros og kritik, talt op på antallet af kandidater der modtager positive/negative/både positive og negative kommentarer på det pågældende tekstelement. Et eksempel er kommentarerne til tekstelementet Forskningsoversigt, med 25 positive kommentarer til 34 kandidater og negative kommentarer til 24 kandidater. Det har vist sig ikke at være fremgangsrigt at kvantificere mængden af henholdsvis positive og negative kommentarer i linjer eller sidetal da et gennemgående fund er at negative kommentarer generelt er længere, de kalder i højere grad på argumentation og dokumentation, mens positive ofte indskrænkes til et kort, positivt udsagn. Kun 4 passager med positive kommentarer er > 16 ord lange, mens 22 passager med negative kommentarer er > 16 ord lange. (16 ord er arbitrært valgt som en kort tekstpassage).

Det er de overordnede resultater af optællingerne og indholdsanalysen af bedømmelserne, vi beskriver nedenfor.

Analyse og resultater

Vi analyserer ud fra skemaets tekstelementer hvilke dele af afhandlingerne der vurderes/ikke kommenteres/vurderes, hvilke dele der roses og kritiseres, og hvad kriterierne er for vurderingerne.

Som en forudsætning for resultaterne skal det nævnes at 38 af 41 afhandlinger var empiriske (to fra Filosofi var teoretiske, og en enkelt fra Dansk blev betegnet som et essay). Samtlige andre afhandlinger analyserer empiri og indeholder analyse som et hovedelement. Der er altså i de fleste tilfælde tale om samme genre og struktur i afhandlingerne på tværs af fakultetets mange discipliner. Strukturerne i afhandlingerne bruger således analyseskemaets tekstelementer, men ikke nødvendigvis dem alle. Ikke alle afhandlinger indeholder fx metodeafsnit, selvstændigt perspektiveringsafsnit osv. – Men: De der ikke gør det, får hyppigt kritik for hvert af de udeladte standardelementer. Det er åbenbart at bedømmerne holder afhandlingerne op imod deres kriterier for videnskab og afhandling, og holder den foreliggende afhandling op imod standardelementer for en empirisk afhandling, med variationer. Bedømmerne kommenterer således meget på fraværende og altså savnede elementer.

Som nævnt er der overensstemmelse blandt bedømmerne om hvilke ca. 14 kategorier der overhovedet kommenteres på, men der kommenteres ikke hver gang på de samme 14.

Tekstdele der ikke kommenteres på

Kategori:	Antal af 41 som ikke kommenteres mht.:
Bidrag	12
Indledning	8
Forskningsoversigt	7
Metode	11
Empiri	14
Teori	8
Analyse	4
Diskussion	8
Konklusion	9
Perspektivering	20
Positionering	20
Argumentation	13
Struktur	12
Sprog og formalia	9

Det er altså ikke de samme kriterier, man kan forvente bedømt i hvert tilfælde, men dog altid en række af de opstillede. Selv ikke bidrag, formål og fokus, metode og fremgangsmåder, dvs. hele afhandlingens forskningsdesign, kan man forvente altid vil blive vurderet eller kommenteret eksplicit. Fx indeholder 12 bedømmelser ikke eksplicitering af om der er bidrag som foreskrevet i Qualifications Framework og Retningslinjer for bedømmelse, 2009.

Positive kommentarer

De positive kommentarer samler sig i dette sample fra humaniora mest tydeligt om elementerne Empiri og Analyse. Kategorien analyseafsnit får flest positive kommentarer af alle kategorier.

Kun 4 afhandlinger får ikke kommentarer til analysen. 33 afhandlingsskrivere får positive kommentarer for deres analyser, 28 får (også) negative kommentarer til deres analyser.

Typiske kommentarer er

"Analysen forekommer nuanceret, skarp og overbevisende. NNs kritiske refleksioner over ... dokumenterer kompleksiteten ..."

"De [analyser] er grundige rent empirisk set med omhyggelig og detaljeret redegørelse for ligheder og forskelle."

Negative kommentarer

Den negative kritik går især på (i rækkefølge efter prominens i bedømmelserne):

- Diskussion (for lidt af) får negative kommentarer i 24 bedømmelser.
- Forskningsoversigt (især huller) får negative kommentarer i 21 bedømmelser.
- Metode (fravær af metode eller fravær af eksplicitering, fravær af methodediskussion) får negative kommentarer i 14 bedømmelser.
- Begrebsdefinitioner og problemformuleringer (fravær af, uklare, glidende) får negative kommentarer i 11 bedømmelser.

Negativ kritik og kritiske diskussionspunkter fylder meget i de samlede vurderinger. I mange bedømmelsers vurderende afsnit får man mere at vide om kritikpunkter end om kandidatens demonstrerede forskningskompetencer, også når afhandlingen overordnet vurderes positivt.

Den største udfordring er "Det store billede"

Vi ser altså at analyser og detailarbejde får megen ros og kun lidt kritik, mens de afsnit og tekstelementer hvor "det store billede" skal tegnes op inviterer til mest kritik: Det gælder især indledning, forskningsoversigt, teori, diskussion, konklusion, per-

spektivering, samt argumentationen i sin helhed. Kritikken går på at disse afsnit helt mangler eller er ufuldstændigt, utilfredsstillende ført igennem. Det er således her at samplets ph.d.-skrivere har de største udfordringer, og udfordringen ligger i etableringen af et større perspektiv, dvs. at se hvor eget bidrag hører hjemme i et fagligt landskab, kunne kvalificere feltets øvrige bidrag og se implikationer af egen forskning.

Her er et eksempel på hvordan en bedømmelse efterlyser "Det store billede" i en mere samfundsvidenskabeligt vinklet afhandling:

"... only limited effort is made to situate debates, forms and developments in social and political organisation in x in relation to larger y and z contexts ... one of the surprising aspects of the thesis is the near-absence of consideration of the significance of the x context for understanding ... it is difficult to make sense of ...without taking into account the influence of x political and civil society. Similarly, larger world-historical-geopolitical shifts might have been given greater attention in seeking to understand the shifts in the x discourse ... But what, in the eyes of the author, do these viewpoints add or change regarding our prior picture and understanding of x and y in z context?"

Se også eksempel citatet nedenfor under bidrag.

De største kritikpunkter– 3 nedslag

Bidrag, forskningsoversigt og diskussion er alle elementer som er relaterede til at se "Det store billede", nemlig hvordan eget arbejde er relateret til feltets øvrige bidrag. Derfor er det interessant at en stor del af de alvorlige kritikker sætter ind her.

Bidrag

12 bedømmelser vurderer afhandlinger uden at kvalificere at arbejdet repræsenterer et bidrag til videnskaben. I stedet anerkendes fx en stor dataindsamling. Vi konstaterer at de samme afhandlinger modtager megen kritik i øvrigt, så der er i det store og hele tale om samplets svageste fjerdedel. Positivt udtrykt er det muligt at erhverve en ph.d.-grad på baggrund af et solidt empirisk arbejde.

Det kritiseres også når skrivere ikke selv ekspliciterer eget bidrag og således ikke synes at være bevidste om feltet og egen positionering i relation til det, fx:

"The author appears to be almost too modest to show that she has contributed significantly to current theoretical understandings of [begreb x]. Yet, the thesis could have been enriched with a clearer spelling out what it is, exactly, that this study adds to the existing body of knowledge on x ... But what, in the eyes of the author, do these viewpoints add or change regarding our prior picture and understanding of x and y in z context?"

Forskningsoversigt

Forskningsoversigt kritiseres i 21 bedømmelser for mindre og større huller, for manglende kvalificering af og positionering i forhold til den omtalte litteratur, og for bias.

Huller er den hyppigste anke mod litteraturinddragelse. Bedømmerne fremfører i mange tilfælde dog ikke argumenter for hvad de manglende resurser så potentielt kunne have betydet for og tilført afhandlingen. Forskningsoversigt (*state-of-the-art*) er i den internationale, empiriske forskning om ph.d.-bedømmelser et centralt kritikpunkt fra bedømmere.

Et eksempel på hvordan dette formuleres i vores danske sample:

"... desuden savner man en diskussion af NNs begreb y, som i den internationale diskussion sammenfatter de tendenser, som afhandlingen berører ved at citere netop NN ... i betragtning af den relative betydning som forfatteren tilskriver x, forekommer det problematisk at han ikke redegør for state of the art i x teori og praksis, eller i det mindste diskuterer de positioner han refererer til i relation til disciplinens samlede udvikling. - Bibliografien tager ikke hensyn til arbejder som ville være naturlige [5 nævnes]. Han påpeger ganske rigtigt at en forskningsoversigt ikke kan indeholde alle tekster i en kontrovers ... påpeger forfatteren, at han fokuserer på fremtrædende skikkelser som skal opfattes som repræsentative. Hvilke disse repræsentanter er, og hvorfor de skal opfattes som repræsentative, er dog uklart ... Gennemgående er det billede der gives af feltet i det hele taget noget forældet ... Samtidig er det problematisk at diskutere x uden at inddrage fag y."

De mange negative kommentarer til forskningsoversigt-elementet tyder på at det ikke lykkes 20 ud af 41 afhandlingsskriverne at overbevise bedømmelsesudvalgene om kandidatens belæsthed (*scholarship*), informationskompetence og overblik over og fornuftige valg mellem tilgængelige resurser.

Diskussion

I alt 24 afhandlinger kritiseres for for lidt diskussion, og diskussion bliver hermed det mest negativt vurderede element i disse 41 afhandlinger. Det er diskussioner af oplagte aktører på feltet, af centrale begreber og ideer, altså i mange tilfælde af temaer der anslås af skriveren som centrale, men som siden ikke diskuteres, men "tages for givet" som nogle bedømmere skriver. Som en bedømmelse formulerer det:

"Indvoendingerne mod denne anden del er nærmere de debatter og problematikker, der med fordel kunne have været inddraget, ikke mindst med x-konteksten i betragtning, men nu lades uberørt. Det havde været oplagt at inddrage debatten om ... [X + to andre debatter som bedømmerne også synes burde have været taget med]."

Ikke i et eneste tilfælde vurderes det at der er for meget diskussion i en afhandling. Samtlige kommentarer enten anerkender den diskussion der måtte være, eller – og dobbelt så ofte – vurderer graden, dybden, nuanceringen, positioneringen som for lidt i bredde og dybde, for lidt ført igennem, for lidt relateret til eget udgangspunkt i forskningsspørgsmål og forskningsoversigt, for lidt positionerende og udviklende i forhold til feltets aktører og positioner, for lidt gåen i (konstruktiv, udviklende) dialog med sig selv som undersøger og med feltets bidragsydere i øvrigt. Her er altså en række skuffede bedømmerforventninger afspejlet. Bedømmerne ønsker sig at skrive-
ren træder et trin op, ser sin egen undersøgelse og sit felts bidrag i fugleperspektiv, gerne med et kritisk blik for bidragenes (teoretiske, især) huller og dermed feltets udviklingsmuligheder.

Diskussion af undersøgelsens metode og transfer

Kategorier

De kategorier i vores registreringsskema som bedømmerne har kommenteret mindst på er: Positionering og Perspektivering som således synes mindre relevant som selvstændige kategorier.

Sample

Også samplet: Bedømmelser fra bare ét fakultet, og bare ét år kan diskuteres. Vi har i dette sample medtaget i alt 9 afhandlinger som blev enten afvist eller skulle revideres. Mange negative kommentarer går på disse 9 afhandlinger idet afvisningen eller revisionskravet skal retfærdiggøres af forbehold, og i et sample på 41 fylder 9 så kritiske bedømmelser meget. På den anden side ligger en stor del af det uddannelsesmæssige potentiale i studiet af indholdet i de mest kritiske bedømmelser, og derfor er netop de lærerige.

Kontekst og transfer

Vi har grund til at tro at resultater af enhver given analyse af et sæt bedømmelser har begrænset transfer – andre fakulteter/bedømmere kan have andre vægtninger, om end ikke radikalt andre kriterier i deres bedømmelser. Resultaterne ser vi som af mere lokal og aktuel værdi, og netop i det lokale og aktuelle ligger også en del af bedømmelsesundersøgelsers pædagogiske potentiale. Dette er den første egentlige undersøgelse af ph.d.-bedømmelser i Danmark, og metoden og registreringsskemaet betragter vi som under udvikling, og som en metode der vil skulle kontekstualiseres til det fakultet eller forskningssammenhæng som den anvendes på. Vi har kun analyseret ét fakultets bedømmelser fra ét år. Kategorierne som udsagn i en bedømmelse grupperes i, bør defineres ud fra de kriterier for bedømmelser som er givet i de officielle tekster som de pågældende bedømmere vurderer ud fra, men man må samtidig have øjnene åbne for temaer som empirien lægger op til (her fx begrebsdefinition som bliver synligt når man analyserer indholdet af kommentarer til tekstelementet

Indledning). Et eksempel på at kriteriernes vægtning måske er fakultetsafhængig, ser vi i betydningen af diskussion i samplet. Den megen kritik af afhandlingernes diskussionsafsnit er et fund som er overraskende i forhold til de forventninger vi kunne have ud fra bedømmelseslitteraturen som i langt højere grad giver forventninger om at *contribution* og *scholarship* vil blive genstand for mest kritik. Muligvis er det specielt for samplet her fra et humanistisk fakultet hvor på den ene side diskussion mellem og i forhold til faglige positioner vægtes højt, og på den anden side forfattes ph.d.-afhandlinger individuelt, og forskergrupper har traditionelt været en sjældenhed. Dvs. at diskussionsmuligheder ikke altid ligger lige for, men skal etableres.

Implikationer

Vi finder at de vigtigste implikationer man kan uddrage for henholdsvis ph.d.-studerende, ph.d.-uddannelse og -vejledning samt bedømmelse, er de følgende:

Implikationer for ph.d.-studerende

Undersøgelsen viser at hvis man som skriver ønsker at skrive sig frem mod en overvejende positiv vurdering af undersøgelsen, så er det især vigtigt at være opmærksom på

- hvad eget bidrag er og at eksplicitere dette/disse bidrag selv
- at have eksplicit problemformulering og formål med undersøgelsen
- at have og demonstrere både scholarship og research
- at forholde sig til kilder, kvalificere og indplacere og afgrænse eksplicit
- at demonstrere næranalyser og gåen i dybden
- at både tage afsæt i, forholde sig til og gerne udvikle "Det store billede"
- at foretage en grundig, begrundet litteratursøgning og udvælgelse/inddragelse
- at være begrebsligt og teoretisk/metodisk eksplicit, præcis, konsistent og begrundende
- at opbygge en argumentationslinje og selv trække argumentationens implikationer og perspektiver
- at give plads for modargumentation, andre og modstridende positioner hvor det er relevant for undersøgelsen
- at diskutere, kritisere og udvikle andres og egne positioner, undersøgelser, metoder og teorier
- være opmærksom på de genkommende kriterier: relevans, præcision, systematik, begrundelse for valg, og at læse korrektur for sammenhæng og korrekthed.

Skemaet på Bilag 1 med temaer i bedømmelserne kan også give ph.d.-skrivere et indtryk af hvad bedømmere og derfor også de selv bør have i baghovedet.

Implikationer for ph.d.-uddannelse/vejledning

Grundlæggende kan undersøgelser af bedømmelsers indhold vise hvilke kriterier der kan/bør/skal kommunikeres til de ph.d.-studerende før de afleverer.

Specielt må de gennemgående kritikpunkter (her især bidrag, litteraturinddragelse, diskussion) være arbejds punkter.

De 12 ikke-omtalte, fraværende eller ikke-ekspliciterede bidrag i dette samples 41 afhandlinger kan ses som en pædagogisk udfordring med at fremhjelpe forskningsbidrag i ph.d.-forløb og afhandling, gennem vejledning og de øvrige muligheder for opponering og sparring som skriversne kan tilbydes.

De mange kritikpunkter af udeladelser af værker eller hele vidensområder som ifølge bedømmelsesudvalgene savnes, rejser spørgsmålet om mangler er udtryk for bevidste, men ukommenterede fravalg, eller om de er udtryk for ukendskab til de pågældende resurser, måske manglende informationssøgningskompetence. Kritiske kommentarer til litteraturgrundlaget for halvdelen af afhandlingerne åbner for implikationer i forhold til vejledning, forskningssammenhæng, kurser og skriversnes egen litteratursøgning og -inddragelse. Hvis der skal være bedre overensstemmelse mellem bedømmernes forventninger til scholarship og repræsentationen af scholarship i afhandlingerne, så må mange kandidater arbejde mere i bredden og i dybden med forskningsoversigt og med kvalificering af og diskussion med litteratur som de inddrager eller afgrænser sig fra. Eller bedømmerne må nedjustere deres forventninger.

I den sammenhæng er det et paradoks at der igennem flere år var ringe søgning til ph.d.-skolens kurser i informationssøgning som blev aflyst og ikke længere udbudt på grund af den ringe interesse, mens der på den anden side var utilfredshed med hjælp til litteratursøgning, og bedømmerklager over huller i litteraturkendskabet (dokumenteret i en intern undersøgelse PhD Experiences 2010). Det ser ud som om mange ph.d.-skrivere undervurderer betydningen af et omfattende litteraturarbejde og bør gøres opmærksom på det.

Også kritikernes topscorer, manglende diskussioner, diskussioner der ikke bliver ført igennem og konkluderet på osv., må være et arbejds punkt ikke kun for ph.d.erne, men også for deres potentielle diskussionspartnere: vejledere, institutkolleger, opponenter og allehånde reviewere undervejs. Diskussionen i en afhandling skal ses i lyset af forskningsoversigten og dens "gaps". De ph.d.-studerendes resultater skal kvalificeres i forhold til andres bidrag og til det hul i forskningen som bidraget vil fylde ud. Skal der her være bedre overensstemmelse mellem bedømmernes forventninger til diskussionsvillighed og hvad der faktisk realiseres, så skal enten mange flere afhandlingsskrivere diskutere mere i afhandlingen, eller bedømmerne skal nedjustere deres forventninger.

Muligheder for i ph.d.-forløbet at foregribe diskussionspunkter udover i vejledning og evt. i fremlæggelser i forskningsmiljøerne er således færre i denne kontekst end i mange ph.d.-miljøer. I solist-miljøer bliver det især vigtigt med mulighed for et præ-defens og med vejleders/bivejleders/forskningsmiljøets rolle i at diskutere. Her bliver forskningsmiljø vigtigt for at udvikle kompetencer til at skrive sig ind i "Det store billede" i eget forskningsfelt.

Implikationer for bedømmelser

Vi har set at i en række bedømmelser kommenteres afgørende kriterier for ph.d.-afhandlinger ikke. Et eksempel er at 11 bedømmelser ikke nævner metode i afhandlingen, og kommenterer eller vurderer således ikke metodebrug som kriterie for forskningskompetence. Som læser ved man ikke om metode ikke kommenteres fordi den "ikke er der", dvs. er implicit, eller om den ikke kommenteres fordi den er OK og ikke påkalder sig opmærksomhed. 7 bedømmelser kommenterer slet ikke på forskningsoversigt. Igen: Er der ikke en forskningsoversigt, eller er den bare upåfaldende? Det afspejler at bedømmelserne ikke er formaterede tekster. Det giver bedømmerne frihedsgrader, men de svingende kriterier har også ulemper for dem der skal orientere sig efter kriterier, og for dem der skal læse og bruge bedømmelserne. Fx kan afhandlingsskriverne og senere ansættelse ikke regne med at en bedømmelse altid vil give et billede af hvordan en given afhandling vurderes ud fra et kendt sæt kriterier for forskningskompetence.

Undersøgelsen kan pege på behov for eksplicite læringsmål for ph.d.-studiet som der bedømmes i forhold til, og for at forskerskoler/ph.d.-administrationen sørger for opstilling af dimensioner som bedømmerne *skal* kommentere på, og som skal kommunikeres ud med autoritet og mere ensartet udformede bedømmelser.

De tidligere nævnte australske forskere påpeger i deres seneste artikel (Bourke & Holbrook, 2013, s. 415) at der nu (som noget nyt) er et stigende krav fra australske universiteter til deres bedømmelsesudvalg om (lokalt specificerede) kriteriebaserede bedømmelser. Forskerne foreslår at disse kriterier baseres på

"... obtaining more information from examiners, who effectively set thesis standards of what is accepted and what is not, this practice can only assist the assessment process, and should be seen as positive for assessment reliability and for providing greater guidance to supervisors and future candidates."

Bedømmelse af ph.d.-afhandlinger er potentielt en detaljeret feedback og feedforward til både de individuelle skrivere og til forskningsmiljøerne. I at gennemlæse og analysere samlinger af bedømmelser ligger forskerskolelederes og dekaners mulighed for at tilrettelægge og løbende justere forskeruddannelser efter hvor læringsbehovene faktisk befinder sig. De faciliteres hvis bedømmelserne altid skal forholde sig

til samme grundtemaer, som fx sammenhængen mellem problemstilling, materialer og metoder, som foreslået af Schaumburg-Müller, 2010, s. 2.

I Australien er bedømmelsen (skrevet af bedømmelsesudvalg på tre, som her) formatteret til en maks. tre siders tekst som vinder i overskuelighed over de meget lange detaljerede og implicit argumenterende, men ofte ikke eksplicit kriterielle, danske bedømmelser som vi har set i nærværende undersøgelse.

Der er al mulig grund til at investere i at alle gennemførte ph.d.-forløb fører til en bedømmelse der kan bruges fremadrettet, og ikke kun til en bagudrettet opsummering af hvad der kunne have været bedre. Og til at tilrettelægge forskeruddannelse og vejledning så bedømmelsernes optik på afhandlingerne forløbene inddrages.

Lotte Rienecker er cand.psych., universitetspædagogisk specialkonsulent ved LearningLab DTU, har gennem mange år arbejdet med universitetspædagogik på bl.a. adjunktspædagogikum, undervisningskurser for ph.d.-studerende, vejledningskurser for især specialevejledere på humanistiske og samfundsvidenskabelige uddannelser. L.R. var i 16 år leder af Akademisk Skrivecenter, KU, og har arbejdet med akademisk skriftlighed og feedback i forhold til studerende og vejledere i over 20 år og er (med)forfatter og redaktør bl.a. sammen med Peter Stray Jørgensen til en lang række lærebøger, bl.a. Den gode opgave (4. udg. 2012) Universitetspædagogik (2013, Samfundslitteratur).

Peter Stray Jørgensen er cand.phil. i dansk sprog, kommunikation og litteratur. P.t. forlagskonsulent og universitetspædagogisk freelancer. Tidligere specialkonsulent på Akademisk Skrivecenter, Det Humanistiske Fakultet, Københavns Universitet (1996-2011) Tidligere redaktør på DUT. PSJ har undervist og skrevet om bl.a. akademisk skriftlighed, genrer, vejledning, studieteknik og akademisk sprog. Har senest arbejdet med universitetspædagogik på kurser for adjunkter, ph.d.er og vejledere. Har sammen med Henrik Galberg Jacobsen skrevet Håndbog i Nudansk (6. udg. 2013), sammen med Lotte Rienecker Den gode opgave (4. udg. 2011), sammen med Lotte Rienecker og Thomas Harboe Vejledning (2005). Senest redigeret og skrevet Universitetspædagogik (2013) sammen med Lotte Rienecker, Gitte Holten Ingerslev og Jens Dolin.

Bilag 1: Pædagogiske perspektiver af en analyse af 41 ph.d.-bedømmelser fra et humanistisk fakultet.

Oversigt over de vigtigste temaer i kategorierne

Temaer i kategorierne

På baggrund af de konkrete analyser af bedømmernes kommentarer, har vi udlæst særligt fremtrædende og hyppige temaer inden for hver kategori. Disse er:

Bidrag til

1. Empirisk, faktuel viden
2. Teorier, fortolkningsmåder, fortolkninger, vinklinger
3. Fagets historieskrivning
4. Metodiske angrebsvinkler, metodeudvikling
5. Faglige/offentlige debatter og til problemløsninger til forskning og/eller praksis.

Indledning

1. Emne, fokus, problemstilling/-formulering, præcision, relevans
2. Begrundelser, motiveringer af valg, relevans
3. Begrebsdefinitioner.

Forskningsoversigt

1. Litteraturkendskabets bredde, udeladelser og huller
2. Relevans af de inddragne kilder
3. Kildebrug.

Metode og metodediskussion/-kritik

1. Hensigtsmæssighed
2. Redegørelse og begrundelse for metoden og metodevalg
3. Systematik, grundighed og præcision i opbygningen og brugen
4. Diskussion og kritik af metoden.

Empiri

1. Omfang af empirien, herunder indsamlingen
2. Dokumentation og systematik, herunder overblik og præcision
3. Relevans, begrundelse og diskussion, herunder udvælgelse og afgrænsning.

Teori

1. Relevans, hensigtsmæssighed og brug.
2. Systematik, grundighed og præcision (begrebsdefinitioner)
3. Overblik, indsigt, sammenhæng (mellem teorier indbyrdes og empiri)
4. Begrundelser for valg, diskussion af teorierne.

Analyse

1. Karakteren af analysen (omfang, kompleksitet, detaljering)
2. Udførelsen (grundighed, konkretion, præcision, systematik)
3. Indsigtsfuldhed
4. Sammenhæng mellem analyse, teori og empiri – eller ofte: mangel på samme
5. Overdrevent teoridrevne analyser (kun negative kommentarer).

Diskussion

1. Overbevisningskraft som resultat af diskussion.
2. Inddragelse af indvendinger, debatter, positioner i en diskussion.
3. Diskussionernes nuancering og grundighed.
4. Gennemførelse af diskussion til en konklusion.
5. Diskussionens relation til teori, positioneringer og brug til at positionere skriveren i forskningslandskabet.
(4 og 5 kommenteres der stort set kun på når de savnes!)

Konklusion, resultater

1. Overbevisningskraft, kvalificering og begrundelse, evt. forbehold /ikke overbevisende konklusioner
2. Tilstedeværelsen af opsummering, sammenfatning, syntetisering
3. Interessant, tankevækkende, værdifuld.

Perspektivering

1. Perspektiver for forskning
2. Perspektiver for faglig, professionel brug og offentlig debat.

Positionering af skriveren

1. Skriveren som fagperson, især i de positive kommentarer
2. Bias, tendentiøsitet, manglende kritisk forholden sig (i de negative kommentarer).

Argumentation

1. Velargumenteret, (solidt) dokumenteret, overbevisende, eller det modsatte
2. Systematik, klarhed, og mangel på samme.

Struktur

1. Klarhed, overblik
2. Intern sammenhæng (mellem indhold og struktur), progression.

Sprog og formalia

1. Velformulering, klarhed, præcision.
2. Retskrivning, korrektur
3. Formalia, især mht. litteraturreferencer og -lister.
4. Læselighed, skemaer, illustrationer, herunder præsentation og forklaring.
5. Levende, personlig, elegant formulering.

Det er en liste over opmærksomhedspunkter både for den ph.d.-studerende og for vejlederen.

Genkommende kriterier på tværs

En metaanalyse af bedømmelseskommentarerne viser at der inden for temaerne er tværsgående kriterier i alle kategorier. Det er markant at der i de fleste kategorier i vurderingerne nævnes

- relevans (brugbarhed i undersøgelsen)
- præcision
- systematik
- begrundelse for alle "handling", især af valg som den ph.d.-studerende foretager.

Litteratur

- Biggs, J. & Tang, C. (2011). *Teaching for Quality Learning at University*. Maidenhead: Open University Press.
- Bourke, S. (uden årstal, formentlig 2007). PhD thesis quality: The views of examiners. Keynote address. Stellenbosch University, South Africa.
<http://academic.sun.ac.za/chaec/conf/post/Bourke%20PhD%20thesis%20quality.pdf> (tilgået 13.6.2013).
- Bourke, S. & Holbrook, A. (2008). PhD thesis quality: Predicting examiner recommendation as one measure of thesis quality. Kiley, M. & Mullins (Eds.): *Quality in postgraduate research: Research education in the new global environment*. G. Canberra: CEDAM, ANU.
- Bourke, S. & Holbrook, A. (2013). Examining PhD and research masters theses. *Assessment & Evaluation in Higher Education*, 38(4), 407-416
- European Qualifications Framework <http://www.grossroads.eu/qualifications-framework-of-the-european-higher-education-area/third-cycle---phd> (tilgået 24.06.2013).
- Feak, C. B. & Swales, J. M. (2009). *Telling a Research Story: Writing a Literature Review*. Michigan Series in English for Academic & Professional Purposes. Ann Arbor: University of Michigan Press.
- Holbrook, A. (2007). Dalliance or intimate relationship? Evidence of what contributes to successful use of the literature in a doctorate. (Keynote, Postgraduate Supervision - the state of the art and the artists. April, 23-26 2007. Stellenbosch, South Africa.
- Holbrook, A., Bourke, S., Lovat, T. & Dally, K. (2004). Qualities and Characteristics in the Written Reports of Doctoral Thesis Examiners. *Australian Journal of Educational & Developmental Psychology*. Vol 4, 2004, 126-145.
- Holbrook, A., Bourke, S., Lovat, T. & Dally, K. (2004). Investigating PhD thesis examination reports. *International Journal of Educational Research*, 41, 2004, 98-120.
- Holbrook, A., Bourke, S., Fairbairn, H. & Lovat, T. (2007). Examiner comment on the literature review in Ph.D. theses. *Studies in Higher Education*. Vol 32. No. 3, June 2007, 337-356.
- Jackson, C. & Tinkler, P. (2007). *A Guide for Internal and External Doctoral Examiners*. Issues in Postgraduate Education: Management, Teaching and Supervision. 2.(2). London: SRHE.
- Lovitts, B. E. (2007). *Making the Implicit Explicit. Creating Performance Expectations for the Dissertation*. Virginia: Stylus Publishing.
- Murray, Rowena (2006): *How to Write a Thesis*. London: Open University Press.
- PhD Experiences 2010. Evaluation of the PhD education, upubliceret. Ph.d.-evalueringsrapport sponsoreret af Københavns Universitets Forskeruddannelsesråd. (Sammenskrivet af The PhD Evaluation Team: P. W. Bay & S. Klerke).

- Retningslinjer for bedømmelser af ph.d.-afhandlinger ved Det Humanistiske Fakultet, 4. nov. 2009.
- Rienecker, L. & Stray Jørgensen, P. (2011). Bedømmelser af Humanioras ph.d.-afhandlinger – en analyse af samtlige 41 bedømmelser af ph.d.-afhandlinger fra Det Humanistiske Fakultet, Københavns Universitet 2009. Upubliceret rapport. Bilag 3.
- Schaumburg-Müller, S. (2009). Retsvidenskabelige problemer i følge bedømmelsespraksis. *Ugeskrift for Retsvæsen*. 192-198.
- Trafford, V. & Leshem, S. (2008). *Stepping Stones to Achieving your Doctorate – By focusing on your viva from the start*. Maidenhead: Open University Press.