

# Evalueringsadfærd: Normer og praksis hos biologiske forskere

Specialerapport ved Biologisk Institut af Iben Wiene Rathje

1. september 2009

*Vejledere: Gøsta Nachman, Biologisk Institut, og Claus Emmeche, Center for Naturfilosofi og Videnskabsstudier, Københavns Universitet*



Illustration ved Ditte Gade: Den evaluerende forsker og den producerende forsker

Denne udgave er renset af hensyn til informanternes anonymitet.

Center for Naturfilosofi og Videnskabsstudier, 2009.

URL: [http://www.nbi.dk/~natphil/prs/iwr/IWR2009\\_Speciale.pdf](http://www.nbi.dk/~natphil/prs/iwr/IWR2009_Speciale.pdf)

## Resume

Naturvidenskabelige forskere læser og producerer mange artikler, idet den enkelte forskers succes bliver målt i antal artikler i high impact tidsskrifter (Kokko & Sutherland, 1999; Lawrence, 2007; Tenopir *et al.*, 2003). Dette studie undersøger, hvordan den kvalitative evaluering af publiceret forskning udfoldes under disse forhold, og om der er normer, der motiverer forskere til som en del af forskningen at agere evaluerende i forhold til andre forskeres forskning.

Undersøgelsen bygger på interviews og indsamlet materiale fra 40 biologiske forskere i Danmark. Forskerne læser de artikler, som lægges til grund for deres egen forskning med meget varierende grundighed. Forskere, som arbejder med biologiske processer under individniveau læser sjældnere metode og statistik end forskere, som arbejder med populationer. Når forskerne har en kritik af den læste artikel, bliver kritikken meget sjældent kommunikeret på en måde som er tilgængelig for forfatteren eller andre potentielle læsere. Forskerne opsøger sjældent andre fagfællers kritik af det læste gennem de tilgængelige databaser.

Ifølge kvalitative interviews i undersøgelsen er der ingen entydige normer i forskersamfundet for evaluering af fagfællers publicerede forskning. Det står derfor klart, at manglende motivation er en grund til at publiceret forskning sjældent bliver evalueret. Derfor foreslås et evalueringssystem, som vil gøre det muligt at bruge, dokumentere og dermed måle den evaluerende adfærd. Dette vil gøre at evaluerende adfærd belønnes i bibliometriske statusopgørelser på linie med publikationer og citationer, jævnfør citationsundersøgelser og H-index.

## Abstract

Scientists read and produce numerous articles, and the success of an individual scientist is measured by the number of papers published in high impact journals (Kokko & Sutherland, 1999; Lawrence, 2007; Tenopir *et al.*, 2003). This study investigates how qualitative evaluation and critique is realized when scientists read papers published by other scientists, and whether there are norms motivating active evaluative behaviour as an integral part of research work.

This study builds upon interviews and written material gathered from 40 researchers within biology in Denmark. It is shown that scientists read papers used for their own research with highly varying degrees of thoroughness. Scientists working with biological processes below the level of the organism more rarely read the methods and statistics sections than scientists working with populations. When a scientist expresses critique of a read paper, it is infrequently communicated in a way that makes it accessible to the

author or potential readers. Scientists seldom search accessible databases for other experts' critique of the articles they use.

Qualitative interviews showed that there are no explicit univocal norms within the research community regarding how to evaluate articles published by peers. It seems clear that missing incentives is a reason why evaluation of published research is rare. Therefore a system of evaluation is proposed that would make it possible to use, document and thereby measure evaluative behaviour. Thus, evaluation and error correction could be recognized in scientometric analyses and rewarded on a par with citations, H-index, etc.

## Indhold

<a href="#">Resume</a> .....	2
<a href="#">Abstract</a> .....	2
<a href="#">Indledning: Hvorfor publicere?</a> .....	6
<a href="#">Læsevejledning:</a> .....	8
<a href="#">Metode</a> .....	9
<a href="#">Kvantitativ undersøgelse</a> .....	10
<a href="#">Koncept</a> .....	10
<a href="#">Udtrækning af informanter</a> .....	11
<a href="#">Interviewenes udførelse</a> .....	12
<a href="#">Statistisk behandling</a> .....	12
<a href="#">Kvalitative interviews</a> .....	13
<a href="#">Artikel-eksempler</a> .....	13
<a href="#">Teori</a> .....	14
<a href="#">Mertons CUDOS-normer fra 1942 til 2008</a> .....	14
<a href="#">Afprøvede alternativer til det nuværende tidsskriftsystem</a> .....	15
<a href="#">Resultater</a> .....	17
<a href="#">Kvantitativ undersøgelse</a> .....	17
<a href="#">Læsning</a> .....	17
<a href="#">Forskerens vurdering af artiklen</a> .....	21
<a href="#">Kommunikation</a> .....	23
<a href="#">Statistisk kunnen</a> .....	25
<a href="#">Opsummering af de kvantitative resultater:</a> .....	26
<a href="#">Justering af tidsskriftsystem</a> .....	27
<a href="#">Kvalitativ undersøgelse af normer</a> .....	28
<a href="#">Uklare normer for læsning:</a> .....	28
<a href="#">Ingen normer for evaluering af referencers referencer</a> .....	29
<a href="#">Generelt ingen normer for opsøgning af publiceret kritik af referencer</a> .....	30
<a href="#">Normer for kommunikation af kritik</a> .....	30
<a href="#">Citatanalyse af citater fra forskernes publikationer</a> .....	32
<a href="#">Opsamling på citatanalyse</a> .....	35
<a href="#">Diskussion</a> .....	36



<a href="#">Motivation i et tidsskriftsystem med bibliometriske effektivitets målinger</a> .....	36
<a href="#">Forventning til læsning - forskel på molekylærbiologer og økologer</a> .....	38
<a href="#">Den fundne læseadfærd i perspektiv</a> .....	39
<a href="#">Communisme - vidensdeling i praksis</a> .....	40
<a href="#">Indførelse af personlig motivering for at kommunikere evaluering ud</a> .....	40
<a href="#">Forbedringsmuligheder af metoden</a> .....	41
<a href="#">Litteraturliste</a> .....	42
<a href="#">Bilag</a> .....	45

## Indledning: Hvorfor publicere?

I det publiceringssystem, vi har nu, er der en nærværende og håndgribelig motivation for at publicere i "high impact" tidsskrifter (Bordons *et al.*, 2003; Kokko & Sutherland, 1999), nemlig karriere og forventninger om høj produktivitet, mens det egentlige overordnede formål med at publicere, at bidrage til den fælles viden, kan glide lidt i baggrunden. Motivation for at publicere bliver stærkere og stærkere, som fokus på bibliometriske indikatorer vokser i takt med, at det elektroniske tidsskrift system bliver mere og mere ene om at varetage kommunikation af naturvidenskabelig forskning. For at kunne vurdere konsekvenserne af dette, er det vigtigt at undersøge forskeres ageren i systemet og deres normer for konkret handling. Publiceringspresset tilsidesætter logisk anden form for videnskabelig praksis, herunder stillingtagen og måske i særlig høj grad kritisk evaluering af andres forskning til fordel for selve opgaven med at publicere artikler. Derfor har jeg fundet det vigtigt at undersøge forskernes praksis for evaluering og normerne for evalueringsadfærd, for at undersøge om der her er et modspil til det voksende publiceringspres.

Jeg har i denne undersøgelse interviewet forskende biologer om deres ageren i tidsskriftssystemet som læsere. Jeg har kvantitativt undersøgt, hvordan forskerne læser og vurderer de artikler, de finder i tidsskrifter på nettet, samt i hvilken grad og på hvilken måde de kommunikerer deres kritik ud, og på denne måde deltager i en kvalitativ feedbackproces. Endvidere har jeg kvalitativt undersøgt normerne for evalueringsprocessen, og i et studie af indgivne publiceringer undersøgt forskernes praksis for skriftlig kritik. Jeg bruger i denne opgave ordene evaluering, feedback og kritik nogenlunde synonymt. Deri ligger, at kritik er at forstå som evaluering eller feedback i systemet, mens den for den enkelte læser vil formes som en kritik. Den enkelte forsker tænker altså ikke "min evaluering af dette vil være", men snarere "jeg har en kritik af dette". Hvis kritikken bliver kommunikeret ud, er den så en del af den feedback, som forfatteren får, og en del af den evaluering, som den publicerede forskning er genstand for. Heraf fremgår også hvorfor jeg bruger alle tre ord, de er naturlige i hver deres kontekst, selvom jeg bruger dem om det samme fænomen. Alle tre ord kan i øvrigt forstås på andre måder, men jeg bruger dem om dele af den samme proces: En forsker finder noget, som vedkommende mener, skulle have været gjort på en anden måde. Man kan godt kalde det "negativ kritik", men jeg vil kalde det konstruktiv kritik, idet man har noget at bidrage med.

Denne opgave handler på sin vis også om sandhed i biologisk forskning. For mig er graden af evaluering af forskellige udsagn ensbetydende med grad af sandhed, idet jeg mener, at hvis metoder og påstande bliver diskuteret vidende mennesker i mellem, så er der større grad af sandhed i de konklusioner, som fremkommer. Jeg har igennem mit studium ikke kunnet undgå at stille spørgsmålstejn ved den viden, jeg

er blevet præsenteret for. Ofte ikke fordi den viden, som præsenteres, forekommer usandsynlig. Men snarere fordi den bliver præsenteret som indiskutabel, hvilket den ofte åbenlyst ikke er.

Jeg er derfor blevet interesseret i forskernes egen refleksion, og i hvilken grad de diskuterer deres forskning med hinanden. Her tænker jeg ikke på resultaterne men på selve forskningen og dens grundlag. Jeg har den antagelse, at hvis publiceringssystemet opfordrede til en højere grad af evaluering og refleksion mellem fagfæller (nære og fjerne), så ville vi bevæge os langsommere og mere sikkert mod en forskning, som er mere sand. Der er lige nu ansat så mange forskere, som hele historien har rummet. Vi har udviklet et publiceringssystem, hvor succes er antal af publicerede af artikler kombineret med kvantitative citeringer af disse. Ingen kredit gives i systemet for at forholde sig til andres forskning, ej heller for at evaluere den, med undtagelse af nødvendige citeringer til undersøgelser, man bygger oven på og med undtagelse af muligheden for at publicere reviews af et helt område. Derfor ser jeg en fare for, at vi bevæger os hen imod et forskersamfund, som producerer en masse publiceringer, som ikke bliver evalueret kvalitativt mellem fagfæller. Det overordnede formål med at forskning publiceres er til dels at udbrede kendskabet til den viden, som forskningen frembringer. Men da det ikke kun er konklusionen, som bliver publiceret, men også metode og baggrund, så må en lige så vigtig del af det overordnede formål med at publicere også være, at forskningen kan blive evalueret og eventuelt brugt som inspiration til ny forskning og vidensdannelse.

Den virkelighed, jeg ønsker at beskrive, er den som den almindelige biologiske forsker arbejder i, af Kuhn kaldet normalvidenskaben (Kuhn, 1962). Derfor har jeg udtaget informanter tilfældigt ud af alle forskende biologer på Sjælland. Dette er utraditionelt i forhold til andre videnskabsteoretiske og historiske udredninger, der ofte forsøger at beskrive de særligt store fremskridt i forskningen eller den store mands forskning. Der hvor det er lykket at sætte en dagsorden, eller opdagelserne har været sensationelle. Den tilfældige udvælgelse er en mulighed, som netop opstår, fordi den virkelighed jeg vil undersøge er definerbar. Det er altid subjektivt, hvilke eksempler man fremdrager i en historisk gennemgang, fordi man har et meget selekteret datamateriale til sin rådighed, og fordi man netop eksplicit udvælger personer med succes. Det har jeg ikke gjort i denne undersøgelse.

## Læsevejledning:

Denne opgave er umiddelbart opbygget meget traditionelt og følger videnskabelige standarder for empiriske artikler. Men på grund af det faglige indhold bryder jeg alligevel med traditionen på følgende måde.

Jeg har i afsnittet Resultater påbegyndt det, som man i en traditionel naturvidenskabelig artikel kalder Diskussion. De variable, jeg bruger, og de resultater, som fremkommer, kan uden den nødvendige forklaring og diskussion virke misvisende. Derfor har jeg tilladt mig at tage forskud på diskussionen undervejs. Tilsvarende har jeg valgt at påbegynde diskussion af teorien i afsnittet Teori. Jeg mener at begreberne skal forklares for at give mening i sammenhæng med mine resultater. I afsnittet Diskussion samler jeg Teori og Resultater, og diskuterer mine resultater i den kontekst, jeg har fremlagt i de foregående afsnit.

## Metode

Mit fokus er forskeren som læser, og forskeren som feedbackaktør i forhold til andre forskeres publicerede forskning. Derfor vil jeg kvantitativt undersøge forskerens grad af læsning af artiklernes forskellige afsnit, forskerens vurdering af det læste, samt forskerens eventuelle kommunikation af feedback til forfatteren af artiklen eller andre aktører indenfor feltet. Herudover er forskernes indtryk af normerne for læsning og evaluering, samt forskernes praksis i at publicere kritik undersøgt kvalitativt. Jeg undersøger således med både kvalitative og kvantitative metoder forskernes kvalitative evaluering af publiceret forskning. Jeg har udtaget en tilfældig stikprøve af biologer til at hjælpe mig med at analysere emnet. Jeg har i min tid som biologistuderende oplevet en tydelig opdeling af biologer, alt efter om de arbejder med populationer (kaldet økologer) eller under individniveau (kaldet molekylærbiologer). Det er derfor nærliggende for mig at sammenligne de to grupper, selvom der naturligvis i nogle tilfælde er problemer med at placere biologer i enten den ene eller den anden gruppe, fordi nogen arbejder i grænseområdet eller på tværs af opdelingen. Jeg har skelnet mellem om man arbejder empirisk over eller under individniveau. Ingen har givet udtryk for at de ikke forholder sig til empirisk arbejde.

Min undersøgelse består af en metodetriangulering, som belyser mit emnefelt med tre forskellige metoder. Dette er valgt i erkendelsen af, at i hvert fald menneskelige interaktioner sjældent lader sig indfange fuldt ud i lyset af en enkelt metode (Holstein, 2001), mens tre forskellige metoder antages at bidrage fra hver deres vinkel, selvom menneskelige interaktioner ikke vil kunne indfanges fuldt og helt uanset hvor mange metoder, der anvendes. Triangulering er fordelagtig i min undersøgelse, fordi flere metoder har egenskaber, som vil være en fordel at benytte i min analyse. Derudover ligger i triangulering en ontologi, som på nogle måder ligner den naturvidenskabelige: Ideen om at der eksisterer en sand virkelighed (Schwandt, 2007). Denne indkredses bedst ved at benytte flere relevante metoder, idet en enkelt metode ofte vil give et begrænset billede af denne virkelighed.

Trianguleringen i min undersøgelse vil bestå af følgende metoder:

- 40 interviews med foruddefinerede svar-kategorier (spørgeskemaet findes i bilag 1). Det høje antal interviews blev valgt ud fra et ønske om at finde konkrete tendenser i forskernes publiceringsadfærd. Forskellen mellem økologer og molekylærbiologers adfærd og en indflydelse af forskernes alder blev testet (Siegel & Castellan, 1988).
- 23 korte kvalitative interviews med fire spørgsmål om normer. Denne metode blev benyttet for at analysere normernes eventuelle modvægt til det publiceringspres, som ligger på forskere i dag. Disse interviews analyseres kvalitativt med udgangspunkt i fremstillingsmåden beskrevet af Dahler-Larsen (2003).

- Citater fra 11 forskeres evalueringsadfærd i publikationer. Denne metode er valgt ud fra overvejelser om, at menneskers adfærd og interaktion med deres miljø bedst undersøges med direkte observationer (Kvale, 2003). Da de direkte observationer af selve adfærden ikke er mulig i dette tilfælde vælges en analyse af produktet af deres adfærd, forskernes egne publikationer.

Interviewene tog i gennemsnit omkring en time. De interviews, hvor forskeren deltog i alle elementer af undersøgelsen (kvantitativt skema, inklusive et lille spørgeskema om deres peer review adfærd, de kvalitative interviews og indsamling af eksempler på en publiceret kritik) tog alle over en time.

Interviewene varierede meget i længde, fra 27 minutter til 2 timer og 55 minutter (de fleste dog lidt over en time) eftersom informanternes måde at kommunikere på, erfaring, stressniveau og interesse for emnet havde stor indflydelse på interviewets forløb. Det kvantitative spørgeskema om evalueringsadfærden blev prioriteret højest i forhold til de øvrige to metoder i trianguleringen. Jeg havde en ambition om ikke at blive afvist angående den kvantitative del af undersøgelsen for at opretholde den tilfældige udtrækning. Dette havde biologerne stor forståelse for, da de jo selv ofte er meget afhængige af tilfældighed i prøveudtagning i deres eget arbejde. En enkelt person havde dog held til at undslippe. En single kvinde uden netværk med tvillinger på 2, hvis bil lige var brudt sammen, mente jeg ikke, jeg kunne overtale på anstændig vis. Ud af en stikprøve på 40 antages det ikke at være en væsentlig fejlkilde i min undersøgelse. De kvalitative normspørgsmål blev prioriteret efter de kvalitative svar. 25 informanter deltog i dette. Den endelige beslutning om normspørgsmålenes udformning blev først besluttet efter de første 8 interviews. Derfor deltager de første 7 interviewede ikke i denne del.

Forud for interviewene udførte jeg to pilotinterviews for at teste mit design. Jeg havde her mest fokus på det kvantitative spørgeskema, da det var vigtigst at få et højt antal kvantitative svar med uændret spørgeskema. Som interviewene skred frem ændrede jeg fokus for de kvalitative interviews, og jeg startede derfor først det kvalitative survey ved informant nummer 8.

## ***Kvantitativ undersøgelse***

### **Koncept**

Denne metode er udviklet efter idé fra en interviewundersøgelse, jeg var med til at udarbejde for Danmarks Fiskeriundersøgelser i 2004-2006 (Rathje, 2006) om fiskeriteknologi. Jeg interviewede i den forbindelse 50 fiskere, og kunne se nogle muligheder i at benytte en lignende metode til at spørge forskere om deres egen evalueringsadfærd. Man kan i begge tilfælde få et estimat direkte fra informanterne, et estimat, som ingen ellers kan give, og som ikke kan måles ellers. Estimatet er et skøn, men i begge tilfælde er det nok så vigtigt, at informanterne er de primære aktører i den virkelighed, man undersøger. Hvis man ændrer



fiskeriteknologien efter fiskernes egne skøn, får man en bedre modtagelse, end hvis det er efter biologernes skøn. Ligeledes mener jeg, at man i regulering af tidsskriftssystemet vil få en bedre modtagelse hos de forskende biologer, hvis man gør det ud fra forskernes egne skøn, end hvis man gør det ud fra videnskabsteoretikeres eller bibliometri-eksperters skøn. Og hvad er så mest sandt? Jeg mener, at i de tilfælde, hvor informanterne kan svare frit uden hensyn til forventninger om i deres svar at skulle leve op til bestemte normer eller reguleringer, så må det mest pålidelige svar komme fra praktikerne (henholdsvis fiskerne og forskerne). Dette kræver, at der ikke er nogle for informanten betydningsfulde normer om adfærden indeholdt i det emne de bliver interviewet.

## **Udtrækning af informanter**

Jeg interviewede 40 tilfældigt udvalgte forskende biologer på større forskningsinstitutioner på Sjælland (KU, DMU, Novo med flere). Jeg udvalgte dem tilfældigt ved at lave en liste over alle de forskende biologer, jeg kunne finde telefonnumre på, på Sjælland. Det blev desværre en lidt skævt fordelt stikprøve med hensyn til ansættelsessted, idet firmaer ikke lægger profiler af deres ansatte ud på internettet, og dem, som jeg har været i kontakt med mente ikke, det kunne lade sig gøre at udlevere dem til mig. Begrundelsen var i Novos tilfælde, at det ikke kunne lade sig gøre at lave udtræk i deres database ud fra uddannelsesbaggrund eller forskningsområde. Ud fra min research lader det til, at private firmaer primært har forskere ansat til at løse teknologiske problemer. Jeg har ikke medtaget forskere med et eksplicit teknologisk fokus, så ud fra denne betragtning mener jeg ikke, at der forholdsmæssigt er mange privatansatte, der burde have været med i udtrækningen. Fejkilden vil være størst, når det handler om molekylærbiologer, idet de fleste firmaer med biologer ansat til at forske beskæftiger sig med humanbiologi.

Jeg startede med en liste på 362 biologer, som jeg opdelte i to grupper: Økologer og molekylærbiologer. På denne måde kunne jeg sørge for, at jeg fik to lige store stikprøver af henholdsvis økologer og molekylærbiologer. Der var 167 økologer og 195 molekylærbiologer. Jeg sorterede dem i tilfældig rækkefølge indenfor hver stikprøve, så jeg kunne starte med at ringe til dem fra en ende af. Før jeg ringede, slog jeg dem altid op på internettet hvis muligt, for at kontrollere at de lå indenfor min målgruppe, altså at de var biologer, som forskede i biologi. Et væsentligt problem var, at det indenfor molekylærbiologi er svært at skelne mellem biokemikere og biologer. Jeg valgte konsekvent at udelade biokemikere, med mindre de havde et rendyrket biologisk fokus og havde en ph.d. i et biologisk og ikke biokemisk emne. Jeg mener, at der er en klar skillelinie mellem biologer og biokemikere, fordi biokemikere fokuserer på kemi og biologer på biologi, det er min erfaring, når jeg læser publikationer og taler med forskere. Naturligvis arbejder biologer og biokemikere ofte sammen, så deres emneområder er ikke nødvendigvis adskillelige. Der er således både biokemi i fagfelter, som beskæftiger sig med økologi, og fagfelter, som beskæftiger sig

med molekylærbiologi. Men det betyder ikke, at biokemikere og biologer er det samme, ligesom statistikere og biologer heller ikke er det samme, selvom de arbejder sammen. Det har naturligvis været svært at skelne skarpt, for der er jo også tværfaglige personer, ligesom skillelinien mellem økologer og molekylærbiologer også er meget flydende. I skelnen mellem økologer og molekylærbiologer, har jeg valgt at fokusere på, om de primært arbejder under eller over individniveau, når jeg skulle gruppere dem. Jeg valgte at lade være med at udelade dem, som arbejdede i grænseområdet, da jeg havde et ønske om, at de to stikprøver samlet skulle være repræsentative for hele biolog-”populationen”. Jeg har altså ikke kun spurgt ”molekylærbiologer” eller ”økologer”, men derimod alle biologer, og dernæst har jeg inddelt dem i én af to kategorier. Det skal her understreges, at en biolog, som jeg har rubriceret som ”molekylærbiolog”, ikke nødvendigvis vil bryde sig om at blive kaldt molekylærbiolog, som udgangspunkt. Men jeg har forklaret informanterne, at jeg ville inddele alle mine informanter i de to kategorier, og de har været med til at vælge hvilken kategori, de tilhørte. Hvis man kun undersøger ”rene” grupper af en bestemt biolog-stereotyp, så får man ikke de mange biologer med, som arbejder i grænseområderne.

## **Interviewenes udførelse**

Før jeg gik i gang med de rigtige interviews, lavede jeg 2 pilotinterviews og tilpassede spørgeskemaet (Bilag 1), så det i interviewsituationen var nemt at spørge korrekt og forståeligt. Min erfaring er, at det er vigtigt, at spørgeskemaet indeholder en formulering, man kan bruge direkte til at stille sit spørgsmål. Spørgsmålet må ikke være for langt, for det virker hæmmende, hvis man skal læse det op fra papiret. Jeg har i praksis i interviewene tilpasset mine formuleringer i forhold til det, vi i interviewet lige havde talt om.

Spørgeskemaet var både på dansk og engelsk. 10 informanter blev interviewet på engelsk.

Det er en vigtig præmis for metoden, at de ting, jeg spørger forskerne om, ikke er omgærdet af et tabu mod at svare bestemte svar. Det var mit indtryk, at informanterne generelt set ikke havde tænkt over deres evalueringssadfærd før, hverken systematisk og ofte heller ikke konceptuelt, og at der ikke var en forestilling om, at de kunne svare forkert på spørgsmålene. Så med hensyn til de tre metoder, der bliver brugt i trianguleringen, så har jeg stor tillid til troværdigheden af svarene.

## **Statistisk behandling**

Svarene blev tastet ind i Excel, hvor det meste af databehandling også foregik. Det havde været en del smartere at benytte et program til i hvert fald Spearmans korrelationskoefficient, da det tog vældigt lang tid at udføre disse tests (med såkaldte tights) manuelt.

I analysen af data, har jeg benyttet mig af ikke-parametriske test beskrevet i Siegel & Castellan (1988), samt en randomiseringstest, hvor man i et computerprogram (lavet til lejligheden af Gøsta Nachman) iterativt

udtager to stikprøver et vist antal gange, og sammenligner variationen mellem de to tilfældigt dannede stikprøver med variationen mellem de rigtige stikprøver. På denne måde har jeg analyseret, hvor stor sandsynligheden er for, at molekylærbiologer og økologer er udtaget af samme "population" (biologer), eller om der er signifikant sandsynlighed for, at de er to forskellige populationer (økologer og molekylærbiologer), som altså adskiller sig med henblik på de undersøgte parametre. De ikke-parametriske test blev valgt, fordi kravene til parametriske tests for det meste ikke var opfyldt af data, idet data ikke var normalfordelt eller unimodale.

### ***Kvalitative interviews***

De kvalitative interviews blev udført efter det kvantitative. Fordelen ved det var, at jeg havde fået introduceret emnet. Informanten kunne nemt forholde sig til et kvalitativt spørgsmål angående normer for den adfærd, vi netop havde talt om. Normspørgsmålene blev besvaret på omkring 10 minutter. Denne del af interviewet blev transskriberet næsten fuldt ud. I tilfælde af, at der var snak, som ikke omhandlede normerne, blev det dog udeladt af transskriptionen. Det tog omkring en arbejdsuge at transskribere de 25 norm-interviews, hvilket var mere end forventet. I de tilfælde, hvor der ikke var enighed om svarene blandt informanterne, arbejdede jeg frem mod et display (Dahler-Larsen, 2003) ved at markere udtalelser i transskriptionen, som var kendetegnende for informantens holdning. På denne måde kan man sammenligne informanterne, og undersøge om der er grupperinger af statements, som kan blive afbilledet i et display.

### ***Artikel-eksempler***

Til sidst i interviewet blev informanten, hvis der var tid til det, spurgt om hun eller han kunne komme med eksempler på, at de har evalueret en publicering kritisk i en af deres artikler. Det kunne både være indirekte eller direkte kritik. Eksemplerne blev i de fleste tilfælde markeret af informanten i en trykt udgave, så det var lige til at skrive ind for mig. Jeg har citeret de steder direkte i mit resultatafsnit. Citeringernes referencer er fjernet, for at min skrevne tekst i specialet forbliver uden spor til informanter. Dog er siderne, hvor eksemplerne er taget fra kopieret ind som bilag, for at sikre den udvalgte læsers overblik, hvilket kan fjernes, hvis opgaven udleveres til andre.

# Teori

## *Mertons CUDOS-normer fra 1942 til 2008*

For at sætte min undersøgelse i en kontekst, har jeg opsøgt litteratur, som beskæftiger sig med normer for evaluering i naturvidenskabelig forskning. Det meste videnskabssociologiske og -filosofiske arbejde beskæftiger sig primært med værdier eller ukonkrete normer (værdinormer) for videnskabelig praksis. Det gælder Merton (1942), som introducerede CUDOS normerne og nyere empiriske studier som Macfarlane & Cheng (2008) og Prpi\_ (1998). Alle fokuserer de på værdier og ikke handling. Der er naturligvis en sammenhæng mellem værdier, normer og handling, og det er også interessant at undersøge forskeres idealer og værdier. Jeg undersøger konkrete handlinger og normer for konkrete handlinger. Jeg bruger begrebet norm om dét, der er et flertal for, at man bør gøre. Altså ikke en gennemsnitlig praksis, ikke et ideal, men en norm for hvad man bør gøre i praksis.

De konkrete handlinger og normer, som jeg undersøger, relaterer sig til værdi-normer som Communisme og Organiseret Skepsis (Merton 1942). Communisme skal ikke forstås i den politiske betydning af Kommunisme, men som en norm for vidensdeling. En norm om deling af viden kombineret med en organiseret skepsis, må føre til at også kritik af metode og resultater bliver tilgængeligt for alle på en struktureret facon. Jeg benytter i denne sammenhæng den almene forståelse af Mertons Organiserede Skepsis (wikipedia.com; Andersen *et al*, 2006), som lægger op til at skepsissen også skal gælde internt i den videnskab, som ikke er influeret af andre institutioner som kirke og stat. Jeg mener, at denne forståelse er en tolkning, som ikke kan gøres udelukkende ud fra Mertons egen udlægning af Organiseret Skepsis. Merton har efter min vurdering et stramt fokus på skepsis overfor kirke og statens indvirkning på videnskaben. Merton selv argumenterede altså ikke i *The Normative Structure of Science* (1942) for en skepsis udøvet som praksis intern indenfor videnskaben selv, kun i tilfælde, hvor ikke-videnskabelige (ikke empiriske eller logiske) betragtninger udøvede indflydelse på videnskabens udfoldelse. Man kan argumentere for, at en logisk følge heraf er, at også videnskaben skal evalueres mellem fagfæller, fordi netop inkonsistens fra empiri og logik vil give grundlag for kritik. Men dette hviler på, at der er en kraftig motivation eller drivkraft for den enkelte til at forholde sig kritisk til fagfællers arbejde, hvilket man ikke kan tage for givet.

Som Kuhn argumenterede for (Kuhn, 1973) uden dog at fokusere på normer, så er teori- og paradigmevalg ikke kun en rationel proces, da værdierne vægtes i mod hinanden, og sommetider er i modstrid. Sådan kan Originalitet, som organiseret Skepsis senere blev suppleret med (Kalleberg, 1997), virke modsat

Communisme og Organiseret Skepsis, ved at prioritere den originale forskning frem for den evaluerende, idet man, hvis man vil engagere sig i at udvise skepsis og evaluere, må nedprioritere den nye og originale forskning. I denne modsatrettede virkning, kan eksterne værdier, som samfundets og kapitalens indflydelse (i forskernes verden ofte udmøntet i et pres for at publicere) have stor indflydelse på den enkeltes adfærd.

### ***Afprøvede alternativer til det nuværende tidsskriftssystem***

I dette afsnit vil jeg gennemgå nogle af de alternativer til peer review processen, som allerede er afprøvet i forskellige tidsskrifter. Jeg vil komme med eksempler på de biologiske tidsskrifter, jeg kender, som tager del i den udvikling. Alle alternativer fremkommer som et svar på kritik af det eksisterende peer review system, hvor man kan sætte spørgsmålstejn ved:

-Communismen, idet man skal have log in for at downloade artiklerne, og kritik ikke bliver publiceret men udvekslet imellem udvalgte

-Disinterestedness, fordi peer reviewerne har mulighed for beskytte deres og deres nære kollegers artikler fra angreb

#### Open peer commentary

BBS (Behavioural and Brain Science) har open peer commentary, hvor der efter artiklen (som da allerede har været igennem et redaktionelt review) er en åben kommentar blog, hvor udvalgte specialister kan kommentere på artiklen, og forfatterne kan svare. Kommentarerne vises i forlængelse af artiklen og har karakter af en ofte uoverskuelig udførlig diskussion.

#### Open peer review

En måde, at imødekomme kritik af den traditionelle peer review proces er, at peer review processen åbnes op i en eller anden grad, således at peer revieweren ikke er anonym overfor forfatteren. Nogle steder bliver peer review korrespondancen offentliggjort. Flere tidsskrifter har taget hul på dette. Biology Direct bruger en version, hvor hele peer review processen foregår online, hvor alle kan følge med hele vejen.

#### "Klik op"-kommentar

Biology Direct har også et system, hvor læsere kan komme med korte kommentarer, som man kan klikke op ved siden af linket, hvor man downloader hele artiklen. Kommentarsystem fungerer godt til at give præcis kritik, både ris og ros. Dette system ligner et system, jeg gerne vil foreslå, hvilket jeg vil komme tilbage til i diskussionen.

### Open access

For at sikre, at så mange forskere som muligt har mulighed for at evaluere de primære kilder, er det en fordel med open access, tidsskrifter, hvor man ikke skal betale for at downloade artikler. Plos One er et eksempel herpå, og Biology Direct har også open access.

### Faculty of 1000

En af mine informanter er fakultets medlem af Faculty of 1000, som er en database af anmeldelser, som fakultets medlemmer udfører. Anmeldelserne består af en rating ("F1000 Factor") og en klassificering som i grupperinger som "New finding" eller "Important confirmation" samt en kvalitativ evaluering. Databasen er mest anvendelig til at holde sig opdateret, i det anmeldelserne ikke er linket til tidsskrifterne. Det er altså en ekstra opgave for den læsende forsker at søge i databasen.



## Resultater

Mine resultater vil her blive fremlagt hierarkisk, som mit spørgeskema er opbygget (bilag 1). Det vil sige, at jeg først gennemgår mine resultater vedrørende biologernes læsning af artikler, dernæst deres vurdering af det de læser, og dernæst deres kommunikation af deres vurdering. På denne måde har det ene spørgsmål i spørgeskemaet dannet vejen for det næste, hvilket jeg også håber, det vil, for læseren af denne opgave.

Den statistiske behandling er foretaget med randomiseringstest for forskel mellem to stikprøver, Spearmans korrelationskoefficient for korrelation mellem to variable i en given stikprøve, og Sign test for at undersøge om en bestemt variabel i en stikprøve generelt har en højere værdi end en anden.

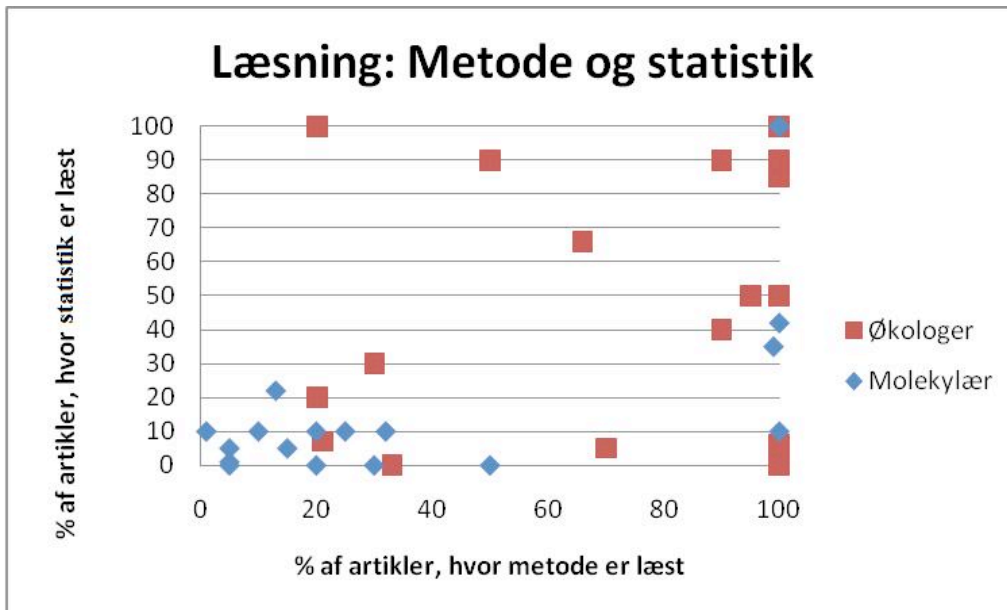
Statistiske resultater vil blive anført som signifikante, hvis sandsynligheden er mindre end 5 % ( $\alpha=0,05$ ) for en fejlagtig forkastelse af 0-hypoteserne. Jeg tester forskelle mellem "økologer" og "molekylærbiologer" indenfor de forskellige spørgsmål, og korrelationer mellem de forskellige adfærdstyper. De vigtigste resultater vil fremgå af underoverskrifterne.

### *Kvantitativ undersøgelse*

#### **Læsning**

Dette afsnit handler om, hvor ofte forskerne læser henholdsvis metode, statistik og relevante referencer til artiklerne på en måde, hvor de efterfølgende kunne tage stilling til, om det var udført på en måde, som informanten kunne stå inde for.

Jeg har i nedenstående graf plottet svarene for metode og statistik, så både en mulig sammenhæng mellem, hvor ofte man læser metode og statistik, vil kunne ses, og også en forskel på molekylærbiologer og økologer vil kunne ses.



Figur 1: Informanternes læsning af statistik i de artikler, som informanten læser og overvejer at bruge til planlægning af egen forskning eller som reference, plottet mod informanternes læsning af metode. Bemærk at man ikke på plottet, kan se de steder, hvor flere informanter har svaret på samme måde med hensyn til både statistik og metode. Der er således 4 molekylærbiologer, som har svaret, at de læser både metode og statistik afsnit 100 % af gangene.

#### Forskel på molekylærbiologer og økologer

Det mest i øjenfaldende på figur 1 er en forskel på molekylærbiologer og økologer. Molekylærbiologerne læser i mindre grad både statistik og metode, hvorimod økologernes adfærd er meget spredt fordelt.

Der er ved en randomiseringstest signifikant ( $\alpha=0,05$ ) forskel på økologer og molekylærbiologers adfærd vedrørende læsning af både metodeafsnit og statistikafsnit.

#### Molekylærbiologers læsning af statistik og metode korreleret

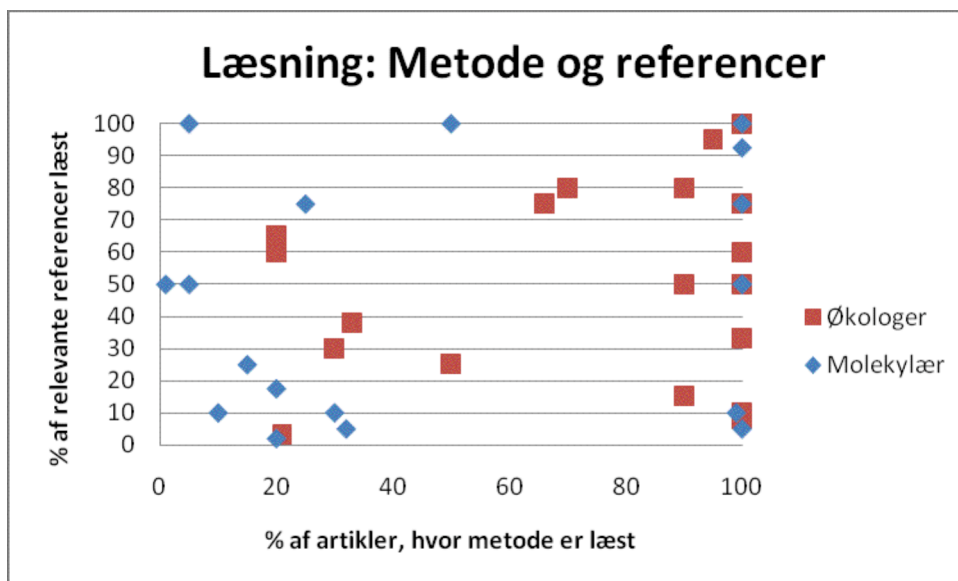
Det ses også, at der er en tendens til, at de samme personer læser en lav procentdel af både artiklernes metodeafsnit og statistikafsnit. Da økologer og molekylærbiologer har vist sig forskellige med hensyn til læsning, er der lavet korrelation for de to grupper separat. Der er signifikant sammenhæng mellem molekylærbiologers læsning af statistik og metode, mens der ikke er en signifikant sammenhæng mellem økologers læsning af metode og statistik (Spearman rankorder). Dette, mener jeg, betyder, at økologer har forskellige fokuspunkter, mens det for molekylærbiologerne mere er et spørgsmål om generel grundighed. Grunden til dette er, at økologers statistiske metoder oftere end molekylærbiologers er meget specialiserede og kræver meget opmærksomhed, mens molekylærbiologers statistik oftere er mere standardiseret og generel. Hvis man går ud fra at noget er gennemprøvet og anerkendt, kan det måske

være svært at motivere sig for at gennemgå det kritisk. Man kan foreslå en anden betydning: At statistikken skulle være sværere at adskille fra metode for molekylærbiologerne, idet det er svært at skelne præcist mellem om eksempelvis antal kontrolforsøg hører med til metode eller statistik. Min skelnen har konsekvent ligget mellem den del af teksten, som beskriver indsamling af data (metode), og behandling af data efter indsamling (statistik). Det er dog ikke sikkert, at alle informanter har kunnet skelne, men dette vil snarere forurene mine data end indgå som systematisk fejl, idet læsning af eksempelvis set up med hensyn til kontroller i dette tilfælde vil tælle med i enten metode eller statistik, og således ikke nødvendigvis forstærke en korrelation mellem dem.

Bemærk at man ikke på plottet, kan se de steder, hvor flere informanter har svaret på samme måde med hensyn til både statistik og metode. Der er således 4 molekylærbiologer, som har svaret, at de læser både metode og statistik afsnit 100 % af gangene.

#### Ingen forskel på molekylærbiologer og økologer angående referencelæsning

Følgende er et tilsvarende plot for reference versus metode.



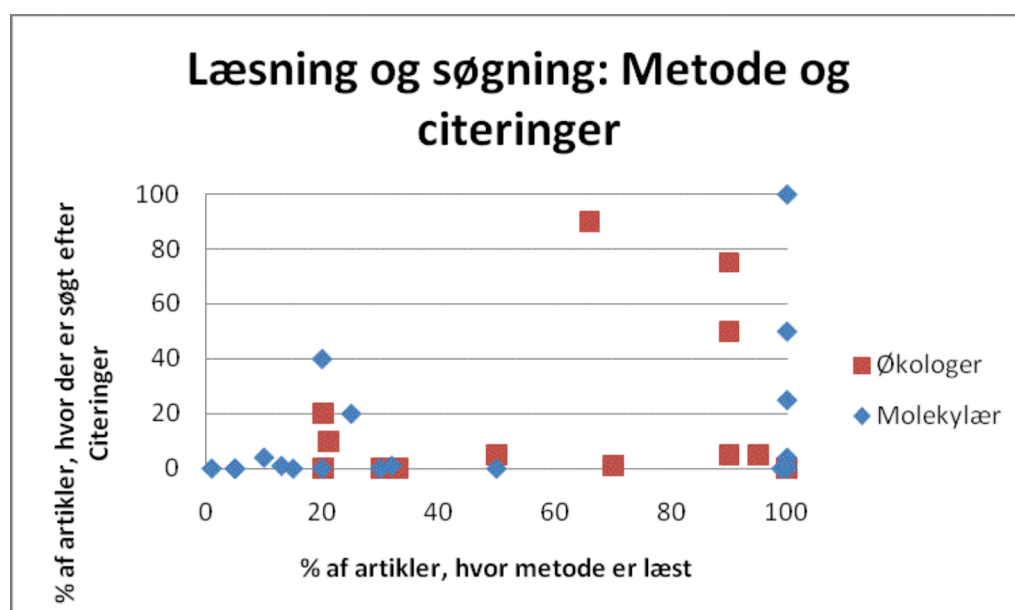
Figur 2: Informanternes læsning af "relevante referencer" i de artikler, som informanten læser og overvejer at bruge til planlægning af egen forskning eller som reference, plottet mod informanternes læsning af metode.

Her ses, at der ikke er en sammenhæng mellem, hvorvidt man læser referencerne i artiklerne, hvis man læser metodeafsnittet. Det ses også, at der er en meget stor spredning af, hvor meget informanterne skønner, at de læser de relevante referencer. Samme tendens gør sig gældende, hvis man plotter referencer mod statistik. Der er ikke signifikant forskel på molekylærbiologer og økologers adfærd med

hensyn til læsning af referencer. Spørgsmålet om læsning af relevante referencer var et vanskeligt spørgsmål, hvilket vil blive gennemgået i afsnittet om forbedringsmuligheder.

#### Over halvdelen af informanten opsøger aldrig kritik via citeringssystemet

Følgende graf er et tilsvarende plot for søgning af citeringer versus metode. Altså hvor ofte informanten søger efter citeringer via en database, når vedkommende læser en artikel, de overvejer at bruge som grundlag for forskning eller som reference.



Figur 3: Informanternes søgning af citeringer af de artikler, som informanten læser og overvejer at bruge til planlægning af egen forskning eller som reference, plottet mod informanternes læsning.

Der er 21 informanter, som opgiver, at de aldrig søger i en database efter artikler, som har citeret en given artikel, det ses ikke tydeligt på grafen.

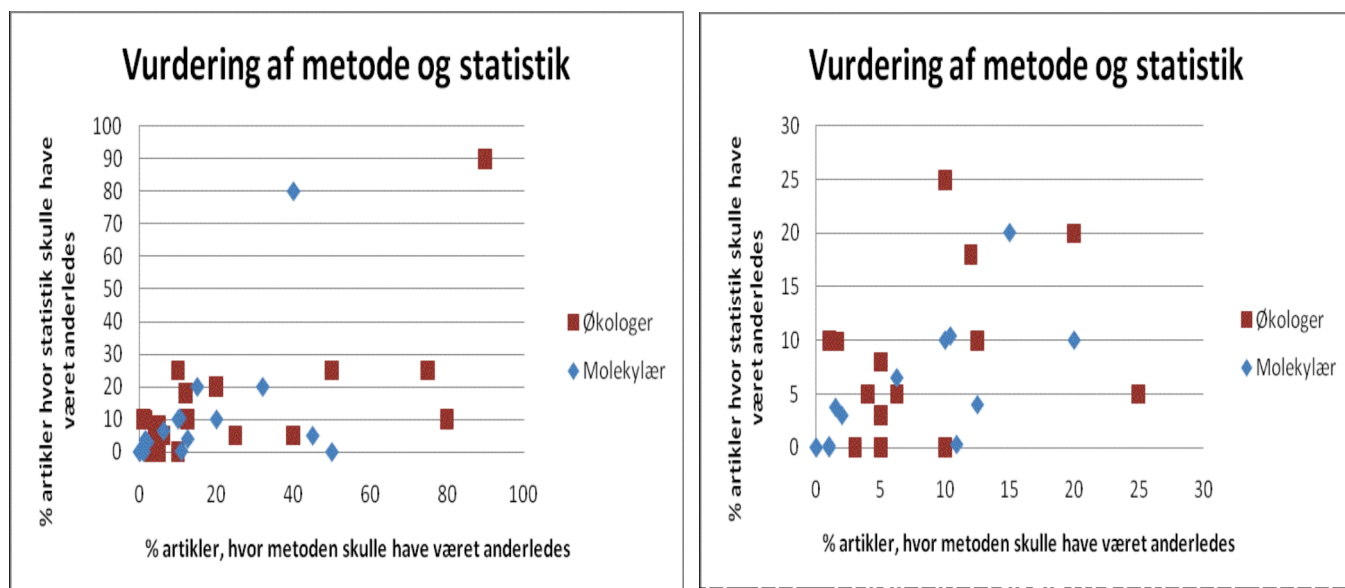
Der er ikke signifikant forskel på molekylærbiologer og økologers adfærd angående, hvor meget de opsøger citeringer af den aktuelle artikel for at gennemgå dem kvalitativt. Kvantitative status-evalueringer af, hvor meget den pågældende artikel er citeret, er ikke medtaget, omend det ofte var således spørgsmålet i første gang blev forstået.

Da der er signifikant forskel på de to stikprøver med hensyn til læsning af metode, testes korrelationen for henholdsvis økologer og molekylærbiologer særskilt. Der er ikke signifikant korrelation mellem metode og søgning af citerende artikler med Spearmans korrelationskoefficient.

## Forskerens vurdering af artiklen

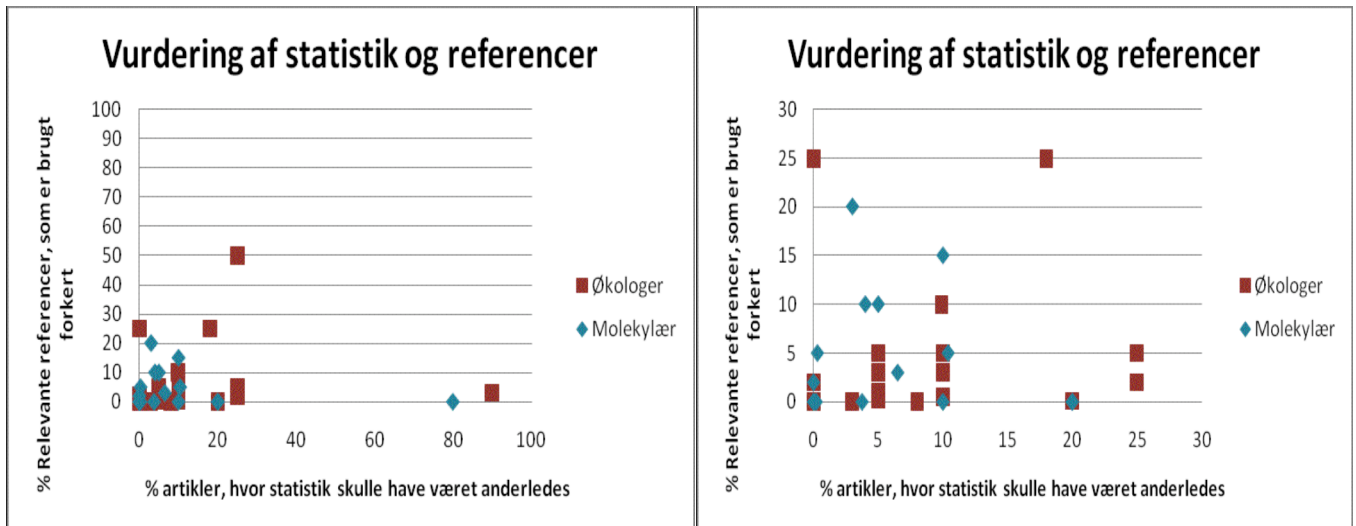
Det næste spørgsmål i interviewene var, hvor ofte vedkommende vurderede at henholdsvis metode, statistik og referencer, skulle have været gjort anderledes. Kun den andel af kritikpunkter, som af informanten blev vurderet til at have potentielle effekter for den biologiske konklusion, er medtaget. Spørgsmålets formål var primært, at forberede det næste spørgsmål, således at informanten var klar over hvilken type kritikpunkter, vedkommende skulle reflektere over. Det bemærkes, at en vurdering af hvornår noget skulle have været gjort anderledes, samt hvor ofte det så har potentiel betydning for den biologiske konklusion, er meget vanskeligt at opgøre kvantitativt for informanterne. Det har heller ikke været interessant for mig at kvantificere hvor mange "fejl" der er, da det slet ikke kan opgøres på denne måde. Afsnittet er alligevel taget med, for at udforske eventuelle forskelle mellem molekylærbiologer og økologer, i det noget af variationen læseadfærden muligvis kunne forklares ved en forskel i oplevelse af kvalitet af det læste.

### Økologer og molekylærbiologer vurderer ikke artikler forskelligt



Figur 4a: Informanternes skøn, hvor ofte de vurderer, at metode/statistik efter informantens mening skulle have været udført anderledes end det fremgår af de læste artikler.

Figur 4b: 30%\*30%kvadratet forstørret, så variationen kan ses.



Figur 5a: Informanternes skøn over, hvor ofte de vurderer, at det læste statistikafsnit eller brug af en vigtig reference gør anledning til at vurdere, at statistik efter informantens mening skulle have været udført anderledes eller at referencen er brugt på en forkert måde.

Figur 5b: 30%\*30%kvadratet forstørret, så variationen kan ses.

2 har svaret, at de aldrig finder noget i metodeafsnittet, som skulle have været gjort anderledes, og som har potentiel betydning for den biologiske konklusion. De fleste ligger mellem 1 og 15 % (figur 4a og b)

Tilsvarende har 2 svaret at de aldrig finder noget i statistikafsnittet, som skulle have været gjort anderledes efter deres mening. De fleste ligger mellem 1 og 10 % (figur 5a og 5b).

Der er ved randomiseringstest ikke forskel på økologer og molekylærbiologer, angående hvor ofte de vurderer, at noget skulle have været gjort anderledes, når de læser henholdsvis metodeafsnittet eller statistikafsnittet eller gennemgår referencer i en artikel.

Der er korrelation mellem informantens vurdering af statistikken og informantens vurdering af metode. Det kan skyldes flere af følgende faktorer:

- Nogle informanter har ikke kunnet adskille metode og statistik, fordi det for dem har været det samme, og det har været svært at adskille svarene konsekvent for mig.
- Nogle informanter kan være mere kritiske end andre.
- Nogle informanter eller små fagområder har en højere grad af tvivl angående deres metoder og statistik.

Man kan ikke vurdere entydigt, hvad korrelationen betyder, og dette vil blive diskuteret under afsnittet diskussion.



Der er en ikke signifikant (randomiseringstest,  $p < 0.1$ ) tendens til at også referencevurdering og metode er korreleret. Der er ikke en tendens til at statistik og reference er korreleret ( $p > 0.25$ ).

#### Ikke signifikant tendens til at metode bliver hårdere vurderet end statistik

Man ser på figur 4a en tendens til at metode bliver vurderet som kritisk oftere end statistik. Denne tendens er ikke signifikant (Twotailed Sign test bruges). Det kan skyldes, at Sign testen ikke vægter store forskelle højere end små forskelle. En stærkere tendens ses på grafen for vurdering af metode og referencer. Denne tendens er signifikant. Ligeledes kunne man så forvente, at statistik oftere blev vurderet som kritisk end referencer, men denne tendens er ikke signifikant ( $p = 0,08$ ).

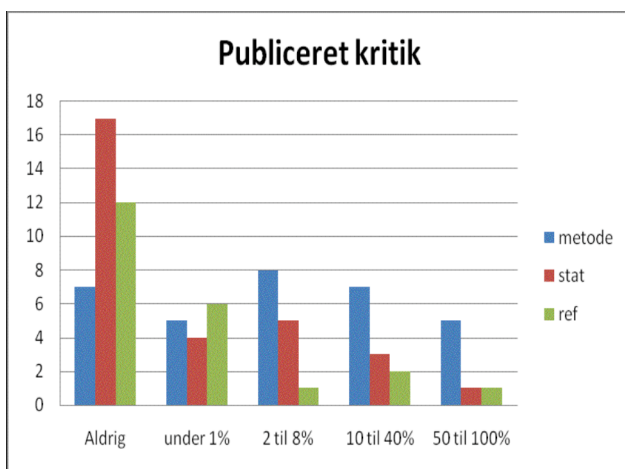
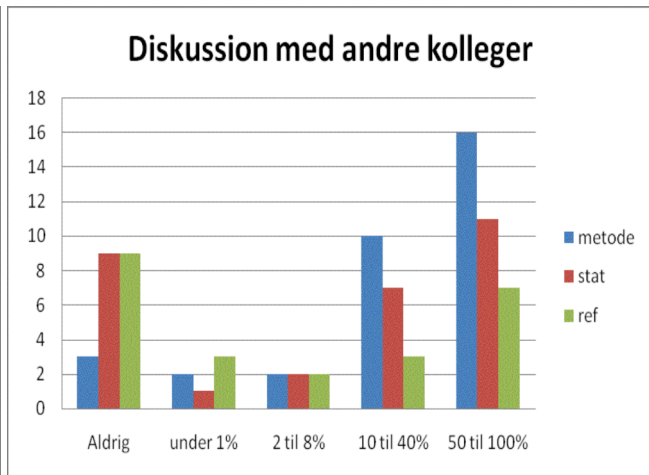
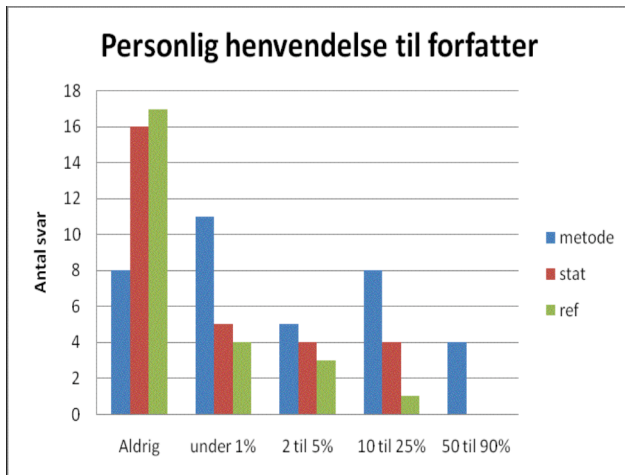
#### Metodeproblem med hensyn til referencernes betydning

Det er vigtigt at lægge mærke til, at her kun er medtaget de kritikpunkter, som er af betydning for den konkrete artikels biologiske konklusion. Betydningen af forkert brug af referencer kan have konsekvenser for den generelle forståelse af biologi uden nødvendigvis at influere på konklusionen i den aktuelle artikel de forefindes i. Det må siges at være anderledes med kritikpunkter vedrørende statistik og metode, som jo har en direkte konsekvens for artiklens egen konklusion.

### **Kommunikation**

Det næste spørgsmål i spørgeskemaet var, hvor ofte informanten gav evalueringen videre, såfremt kritikpunktet var af potentiel betydning for den biologiske konklusion. Svarene blev delt op i forskellige evalueringsmetoder:

- Personlig henvendelse til forfatteren (herunder også tilfældigt møde ved konferencer)
- Diskussion med øvrige kolleger
- Publicering af kritik med reference til den pågældende artikel



Figur 6a-c: Evaluering af de artikler, hvor informanten skønner, at noget skulle have været gjort anderledes. X-aksen er ikke kontinuert og angiver andel gange, den enkelte forsker skønner at udføre evalueringsformen. Y-aksen angiver antal svar indenfor det pågældende interval angivet på X-aksen.

### Metode bliver evalueret mere end statistik og referencer

Figurene 6 a-c viser alle sammen, at metodeafsnittet er den del af artiklerne, som i størst udstrækning bliver evalueret i tilfælde af fejl eller uenigheder. Jeg har valgt at dele søjlerne op i "naturlige opdelinger" for at øge det visuelle overblik. Det er en opdeling, man bruger meget i Geografiske Informations Systemer for netop at vise variationen i data visuelt overskueligt. Man deler der, hvor data selv har huller, og fordeler så vidt muligt data i lige store grupper. Bemærk at den horisontale akse således ikke er kontinuert eller på en konsekvent kvantitativ skala. Der er med Sign test en signifikant forskel på både graden af alle tre former for kommunikation af evaluering angående metode og henholdsvis statistik og referencebrug. Biologer evaluerer altså i højere grad metode end statistik og referencer. Det bemærkes, at informanternes svar kun

gælder tilfælde, hvor de vurderer, at der er en potentiel konsekvens for den biologiske konklusion. Der er altså ikke grund til at tro, at det er en vurdering af vigtigheden af metode frem for statistik, som ligger til grund for denne forskel.

Der er ingen signifikant forskel på graden af videregivelse af evaluering for henholdsvis referencer og statistik. Den ikke signifikante tendens (omkring  $p=0,30$  med den konservative Sign test) går i retning af en større evaluering af statistik.

#### Diskussion med kolleger mest almindeligt

På figurerne 6 a-c ses også en forskel på, hvor ofte de tre kommunikationsformer benyttes. Diskussion med andre kolleger end forfatteren er det mest almindelige, mens særligt statistik og referencer sjældent bliver evalueret på en form, som er tilgængelig for forfatteren af artiklen. Dette bekræftes også statistisk, da der ved Signtest er en signifikant forskel på diskussion med andre kolleger og de andre to kommunikationsformer, mens der er en emne-afhængig forskel mellem publiceret kritik og personlig henvendelse til forfatteren. Der er en ikke-signifikant tendens til at metodekritik i højere grad publiceres end den bliver kommunikeret personligt til forfatteren. Med hensyn til statistik er der slet ikke en forskel, mens der med hensyn til referencer er en signifikant tendens til at kritikken publiceres snarere end at blive kommunikeret personligt.

#### Generelt ingen forskel på molekylærbiologer og økologer angående kommunikation

Randomiseringstests viser, at der langt fra er en signifikant forskel på molekylærbiologer og økologer med hensyn til om de kommunikerer deres kritiske evalueringer af metoden samt referencebrugen ud. Det samme gør sig gældende med hensyn til statistik. Personlig kontakt vedrørende evaluering af statistik er signifikant forskellig mellem økologer og molekylærbiologer. Det kan ikke afvises, at det er et tilfælde, da forskellen mellem de to stikprøver udgøres af fire personer ud af 30.

### **Statistisk kunnen**

#### Statistiske evner influerer på vurderingen af statistikken samt publicering af kritik

Som tillæg til den øvrige undersøgelse, og fordi jeg var nysgerrig, spurgte jeg informanterne om, hvor gode de selv synes, de er til statistik. Der er ikke forskel på, hvor meget man læser statistikken, alt efter hvor god man anser sig selv for at være i forhold til det ideelle. Der er derimod en signifikant forskel på, hvor ofte de vurderer, at noget af statistikken skulle have været gjort på en anderledes måde. Dem, som vurderer sig selv som værende gode nok i forhold til det ideelle indenfor deres felt, mener, at der oftere er noget, som skulle være gjort anderledes. Der er ikke en signifikant tendens til, at de "gode nok" oftere kommunikerer

med deres kolleger eller henvender sig personligt til forfatteren om statistikken. Der er derimod en signifikant tendens til, at de "gode nok" oftere publicerer kritik end de "ikke gode nok", i de tilfælde at de har en kritik at publicere.

### Både økologer og molekylærbiologer udfordret angående statistik

Der er ikke forskel på, hvordan molekylærbiologer og økologer vurderer deres statistiske viden. I stikprøverne var der halvt af hvert, som vurderede sig som henholdsvis gode og ikke gode nok i forhold til det ideelle indenfor det område, de arbejder med. Dette er vigtigt i forhold til forståelsen af, hvorfor molekylærbiologer og økologer læser statistik i forskelligt omfang. Det tyder ikke på, at statistik indenfor molekylærbiologi ikke kan opfattes som så gennemprøvet og trivielt, at der ikke er nogen grund til at bruge tid på det.

### **Opsummering af de kvantitative resultater:**

Da jeg havde skrevet resultatafsnittet og læste det igennem, følte jeg trang til en mindre teknisk opsummering af de mest interessante resultater i min undersøgelse. Jeg kan derfor forestille mig, at mine læsere måske har den samme trang, så her følger en mere flydende opsummering af de væsentligste konklusioner.

#### Læsning:

Økologer er meget forskellige, når det handler om, hvor meget de læser både metode og statistik afsnittene. De ligger helt spredt fordelt i spektret, dog er der ingen, som angiver at de læser metode mindre end 20 % af gangene.

Molekylærbiologer er også forskellige, men de fleste angiver, at de læser metodeafsnittet og statistikafsnittet i under 30 % af artiklerne. Det er de samme molekylærbiologer, som adskiller sig med hensyn til statistik og metode, og der er da også en korrelation mellem de to variable for molekylærbiologernes vedkommende. Dette kan tyde på, at økologer er mere specialiserede i enten statistik eller metode, mens det for molekylærbiologerne handler mest om grundighed, eller hvorvidt der er en norm om eller en grund til at man skal eller ikke skal læse metode- og statistikafsnittet. Forskerens vurdering af sin statistiske kunnen lader ikke til at influere på læsningen af statistik.

Man skal ikke lægge for meget i besvarelsen af læsning af relevante referencer, idet det var en meget vanskelig udregning for informanterne at gøre. Citeringerne derimod var lettere at opgøre. Over halvdelen af informanterne undersøger aldrig, om der er en kritik af en artikel, de skal bruge, mens resten af

informanterne er spredt fordelt, flest dog under 20 %. Biologer opsøger altså sjældent evalueringer af den forskning, de bygger videre på, via databaser på nettet.

### Forskerens vurdering af artiklen

De fleste besvarelser angående, hvor ofte forskeren vurderer, at noget skulle have været gjort anderledes, ligger mellem 1 og 15 % for metode og 1 og 10 % for statistik. Der er ikke forskel på økologer og molekylærbiologer i denne sammenhæng. Der er ikke signifikant forskel på, hvorledes statistik og metode bliver vurderet. Statistik bliver vurderet mere kritisk af forskere med større tiltro til deres egen statistiske viden. Referencer bliver sjældent vurderet til at have været brugt forkert på en måde, hvor det har betydning for den biologiske konklusion. Her ligger de fleste mellem 0 og 5 %. 10 informanter i alt vurderer, at en forkert brug af en reference har betydning for den biologiske konklusion i en artikel.

### Kommunikation:

Biologer kommunikerer generelt i højere grad kritik af metode end kritik af statistik og referencer.

Diskussion med andre kolleger end forfatteren til artiklen er den mest anvendte evalueringsform. Særligt statistik og referencer bliver sjældent evalueret på en form, som er tilgængelig for forfatteren af artiklen.

Der er generelt ikke forskel på, hvor meget økologer og molekylærbiologer kommunikerer deres kritik.

Forskere med stor tiltro til deres egen statistiske viden publicerer oftere deres kritik af statistik end forskere med mindre tiltro til deres egen statistiske viden.

## **Justering af tidsskriftsystem**

I dette afsnit har jeg valgt at bruge tal i stedet for at skrive tallene ud, idet det øger overblikket over de tal, som sammenlignes.

For at teste et alternativ til det eksisterende evalueringssystem for publiceret forskning, har jeg spurgt forskerne, hvad de ville synes om et publiceringssystem, hvor der var mulighed for at få publiceret kommentarer med et synligt link fra den pågældende artikel, netop der i artiklen, hvor kritikken relaterer sig til. Kommentarerne skulle i den foreslåede model være godkendt af editor. 27 informanter svarede, at de synes, det ville være en god idé, mens 9 svarede, at de ikke synes, det var en god idé, og 2 svarede måske. Altså er der en rimeligt stor overvægt af informanter, som synes ideen er værd at prøve.

Begrundelserne for ikke at være interesseret i det nye system er følgende:

-Det fungerer fint, som det er. Man kan kommentere i slutningen af sin egen artikel

-Man skal hellere læse artiklen objektivt.

-Mangel på tid og ressourcer

-Generel skepsis angående kvaliteten af kommentarerne.

30 informanter svarer endvidere, at de ville læse andres kommentarer, hvis systemet eksisterede, mens kun 5 siger, at de ikke ville læse dem. 9 ville ofte skrive kommentarer selv, mens 13 ville gøre det sjældent og 14 ikke vil skrive kommentarer. Det er altså meget forskelligt, hvor meget folk selv ville skrive kommentarer, og det er værd at bemærke.

29 af forskerne mener, det er en god idé at indføre, at man skal kunne downloade data og udregninger i sammenhæng med publicerede artikler. 6 mener ikke, at man skal kunne downloade den slags materiale. De begrundet det primært med, at det er for besværligt. En enkelt har oplevet, at data bruges på en fejlagtig måde, fordi nogle har lavet analyser af det uden at have forstået usikkerhederne ved data. 15 informanter oplyser, at de bruger eller ville bruge muligheden (hvis den var der) for at downloade ekstra materiale ofte. 20 informanter ville gøre det en sjælden gang imellem, mens 2 aldrig ville bruge det.

## ***Kvalitativ undersøgelse af normer***

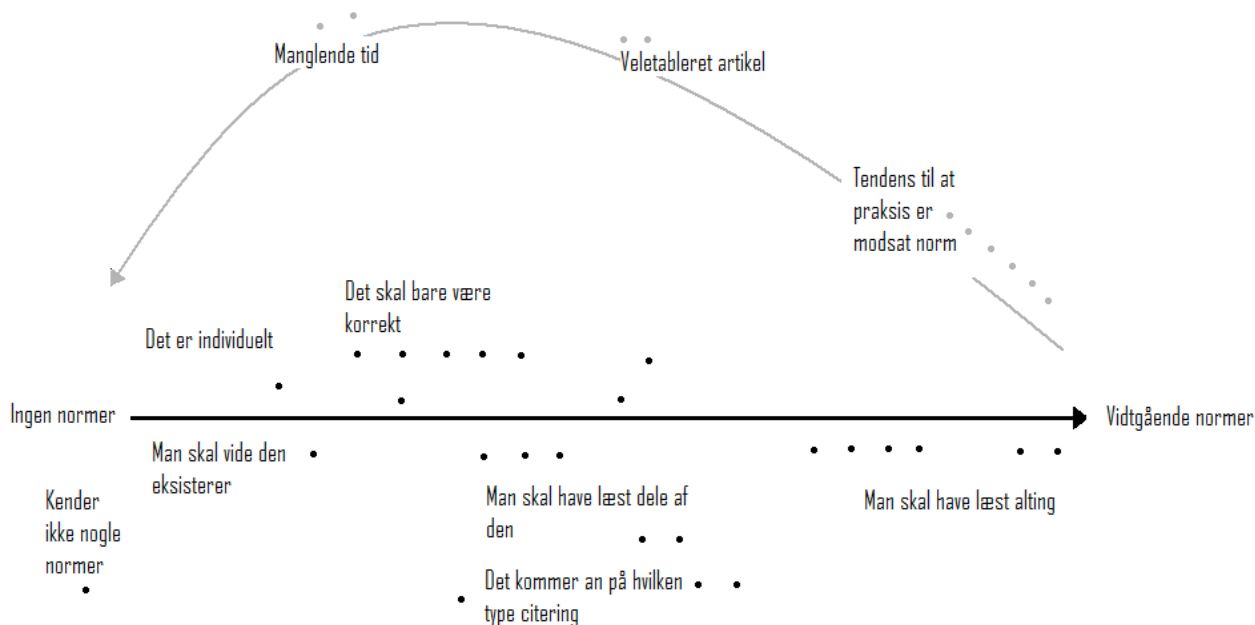
For at belyse en mekanisme, som kan have en indflydelse på de parametre, jeg kvantitativt har undersøgt, har jeg spurgt informanterne om, hvilke normer, de mener, der er i det forskningsmiljø, de er en del af. De fulde transskriptioner findes i bilag 2.

### **Uklare normer for læsning:**

Første spørgsmål i den kvalitative del af interviewet var: Er der nogle normer for, hvor meget man skal læse de artikler, man refererer til?

I nedenstående figur 9, har jeg forsøgt at afbilde de 25 kvalitative svar grafisk på en linie, som går fra, at informanten oplever, at der ikke er nogen normer, til at der er vidtgående normer. Ud over den lineære bevægelse har jeg valgt at afbilde mere komplekse svar vertikalt. Der er stor forskel på, om man mener normen er, at det bare skal være korrekt, hvordan du opnår den sikkerhed er dig frit for (kan også være mundtlig eller skriftlig andenhånds information), eller om man mener, at man selv skal have læst alle artikler fra A til Z. Hver informant er repræsenteret ved en prik, som er placeret der, hvor jeg har vurderet, de passede bedst. De er tættest ved den kvalitative kommentar, som beskriver dem bedst, men er samtidigt placeret i forhold til den lineære graduering og de øvrige kvalitative kommentarer. To informanter ligger mellem statementet om, at man skal have læst dele af den og statementet om, at det

kommer an på hvilken type citering, da de har givet begge udsagn. Det grundlæggende indtryk er, at de fleste mener, at der er nogle normer, men at det er meget forskelligt, hvor vidtgående normerne er. De fleste informanter har også virket ret usikre på, om der egentlig er normer, så alt i alt er min konklusion, at der ikke er klare normer på området. Der er ikke forskel på økologer og molekylærbiologer.



Figur 9: Informanterne er hver repræsenteret med en sort prik (tæl her ikke de grå prikker ved den buede linie med) Den sorte, lige horisontale linie repræsenterer et spektrum af informanternes normer, hvor dem placeret til venstre i figuren udtrykker de mindst vidtgående normer, og dem repræsenteret til højre i figuren udtrykker de mest vidtgående normer for læsning af artikler. Figuren har ingen "y-akse", variationen vertikalt er således kvalitativ med statements, som informanterne er placeret i forhold til. Den grå buede linie og de grå prikker er en illustration af de mekanismer, som kan gøre at normerne i praksis er nærmere venstre side, end de har oplyst. De grå prikker angiver hvor mange af informanterne, som har givet det pågældende statement. Alle informanter er hver repræsenteret ved en sort prik. Herudover er nogle af informanterne repræsenteret ved en eller flere grå prikker.

## Ingen normer for evaluering af referencers referencer

Et tilsvarende spørgsmål blev stillet angående referencerne i de artikler, man refererer til: Er der nogen normer for hvor meget, man skal læse referencerne i de artikler, man refererer til? Her er tendensen klar, der er ikke nogen norm, for hvor meget man skal læse artiklerne, som artiklerne, man læser, refererer til.

## Generelt ingen normer for opsøgning af publiceret kritik af referencer

Et tilsvarende spørgsmål blev stillet for at belyse, om der er nogle normer for, hvor opsøgende man selv skal være for at opfange andres evalueringer af det, man refererer. Dette er et meget essentielt spørgsmål for min undersøgelse. Spørgsmålet lød: "Er der nogle normer for, om man skal tjekke citeringer af de artikler, man refererer til?" Ofte udbygget med en forklaring for at imødegå den misforståelse, at mit fokus skulle være en kvantitativ evaluering i form af en citationsanalyse, som de fleste forskere har større fokus på.

Tendensen i svarene er klar, der er ingen normer for, om man skal undersøge, om der efterfølgende er kommet en kritik af den artikel, man refererer. Tre informanter modsiger flertallet og har en forståelse af en norm på området. En af disse, siger:

"**Ja, selvfølgelig.** Det er alfa mega! Noget andet er så, hvis man laver fejl, men bestræbelserne er uomgængelige. (...) Og det er jo meget nemmere at gøre, end det var i gamle dage. Der sad man med de store tykke bøger med numre, og citationsindex, som stod på kilometervis af hylder, det var et enormt arbejde at sidde med. Nu kan man klikke sig ind, og hvis man ikke gør det, så **svigter man sine kolleger.**" Dette bakkes op i blødere udgave af en anden informant.

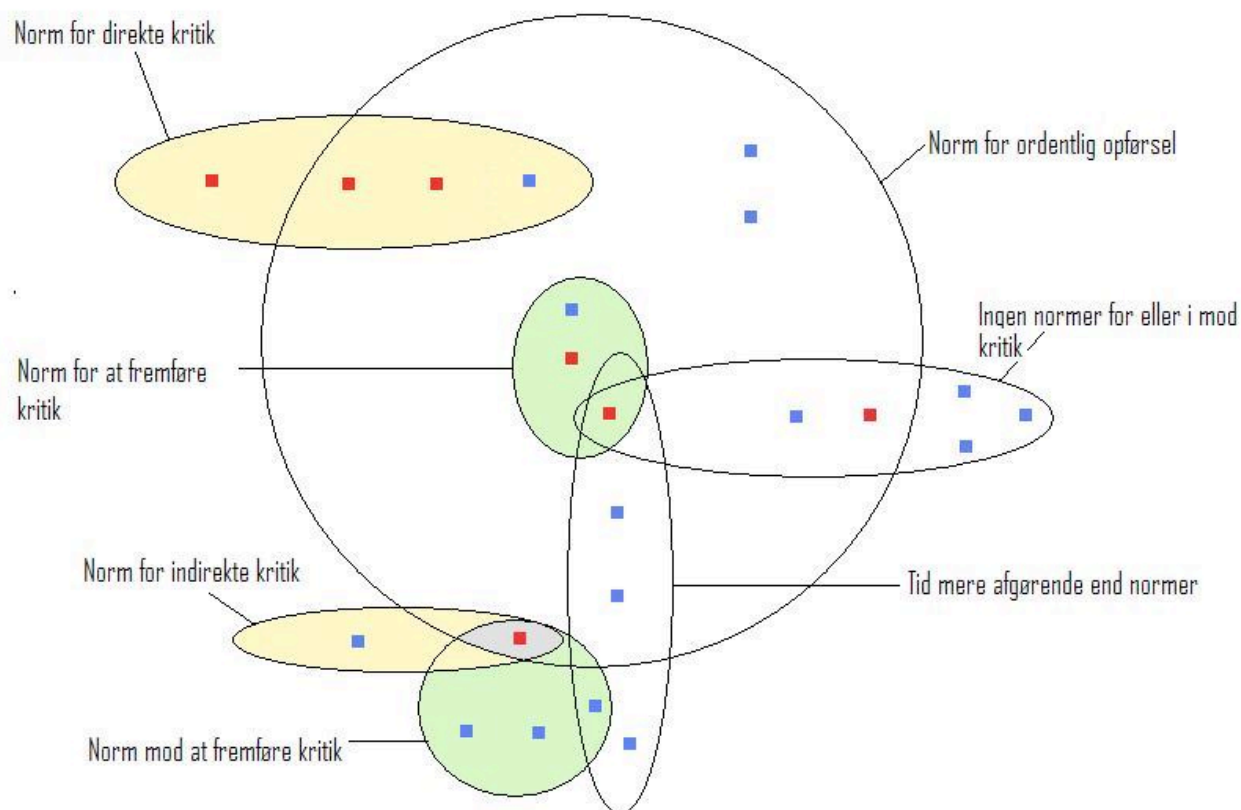
En anden påpeger, at der på PubMed, popper en blok op, hvis der er en tilbagetrækning, og det er vigtigt at tjekke den, hvis det sker. En af dem, som ikke mener, der er en norm, påpeger, at man ikke finder erratum på Web of Science.

Ellers er det klare svar, at der ingen norm er for det, men fem nævner selv, at de synes, det er en god idé at gøre det, mens en anfører, at det vil være urealistisk at gøre på grund af manglende tid.

## Normer for kommunikation af kritik

På figur 10 ses en illustration af, hvorledes informanterne fordeler sig på spørgsmålet, om der er nogle normer for, hvornår eller hvordan man skal kritisere andres artikler. 7 økologer og 15 molekylærbiologer har svaret på dette spørgsmål. Få prikker kan ikke tages som udtryk for, at det ikke er en gældende norm, da ikke alle har svaret på alle forhold. Informanterne har nogenlunde selv kunnet bestemme, hvad de ville fokusere på i deres svar.





Figur 10: Firkanterne repræsenterer hver en informant, de røde er økologer og de blå molekylærbiologer. Cirklerne repræsenterer statements, som informanterne indenfor cirklen har givet udtryk for. At en informant ikke har ytret et bestemt statement betyder ikke, at den pågældende ikke støtter det pågældende statement, idet ikke alle informanter har udtalt sig vedrørende alle statements. De to farver på cirkler, gul og grøn indikerer, at der her er modstridende statements, som kan sammenlignes: De omhandler altså det samme, men modsiger hinanden.

Den største enighed går på, at man skal kritisere på en ordentlig måde, hvis man kritiserer. 13 informanter gør opmærksom på dette forhold med følgende adjektiver, man helst skal opfylde: Sober, høflig, konstruktiv, friendly, saglig, velbegrunder, pænt, lødigt og objektiv. Det må ikke være for hårdt, ubehageligt, aggressivt, og frem for alt må det ikke virke, som om kritikken er personlig.

På figur 10 ses endvidere, at 6 informanter mener, at der ikke er en norm om, hvorvidt man bør kommunikere sin kritik ud. 4 personer mener, at der er en norm imod, at man kommunikerer sin kritik ud, 3 mener der er en norm for. 4 anfører tid, som et argument for ikke at gøre det, altså en grund modsat en norm. Jeg har endvidere forsøgt at få at vide, om der er en norm for, at man skal kunne læse direkte ud af

teksten, at forfatteren mener, at der er noget, som er gjort på en forkert måde. 2 informanter anfører, at det bør være på en indirekte måde, mens 5 anfører, at det skal være på en direkte måde.

(Følgende er ikke afbilledet på figuren, da det ikke kunne lade sig gøre at få relevante informanter fanget ind af cirklerne, uden det blev meget uoverskueligt.) 2 informanter mener, at det kan skade ens karriere at kritisere andres forskning. I modsætning til dette påpeger 2, at det også kan gavne karrieren. Yderligere 3 mener, at det er bedre at komme med et positivt eksempel på, hvorledes man skal gøre i stedet for at kritisere andre. 4 informanter mener, at man bør kommunikere kritik ud, hvis kritikpunktet har en stor betydning for feltet, altså en norm for i visse tilfælde at kommunikere sin kritik ud i andre tilfælde ikke.

## Citatanalyse af citater fra forskernes publikationer

Herunder følger formuleringer fra publicerede evalueringer, som 11 informanter selv har indgivet efter opfordring. De vendinger, som formulerer selve kritikken er markeret med fed. Nogle af de mere indirekte kritikker vil være svære at forstå uden helheden, men det er også en pointe i sig selv. Sammenhængen vil delvist fremgå af bilag 3, hvor hele den side, hvorfra citatet stammer er kopieret.

### Informant 1

En informant har i samme publikation et eksempel på en direkte og en indirekte kritik. Han forklarer i interviewet ikke, hvorfor han i det første tilfælde vurderer, at det er en god idé at komme med en direkte kritik, men han forklarer hans tanker bag ved den indirekte kritik.

”Citaterne vil blive læst op til Siv.” Eksempel 1.

Dette er en direkte kritik, med klar reference til dem, han mener, gør det forkert.

I samme publikation er et eksempel på en formulering informanten oplyser, at han har tænkt længe over, fordi han vil ”vække dem, som var faldet i søvn, når de kom til den side”.

” Citaterne vil blive læst op til Siv.)” s. 183 eksempel 1.

Han gør ikke direkte opmærksom på, at resultaterne i de omtalte studier er upålidelige på grund af den store variation i det målte stofskifte. Han anfører blot, at variationen er ”overraskende”.

Adspurgt om han ikke burde skrive et sted, at det er en misforståelse at tolke resultaterne på den måde siger han: ”Nej, det gør jeg ikke, jeg skriver til kloge folk. De kan sige, nå, er det overraskende, hvorfor er det det, jamen det er det sgu da også, det har vi da aldrig tænkt over før.(...) Du må regne med, der er jo noget diplomati i det her.”

Følgende er ligeledes markeret som værende en kritik eller evaluering af andres arbejde:

” Citaterne vil blive læst op til Siv.”

Læser man citatet i dette lys, kan man forestille sig, at det første statement er en kritik af de referencer, som gives i næste sætning. Det vil altså være muligt at forstå hans evaluering, hvis man i forvejen kender de to referencer eller efterfølgende læser dem. (Eksempel 2)

#### Informant 2

”Citaterne vil blive læst op til Siv.”. (Eksempel 3)

Her er en relativ klar kritik af en metode, og der gives reference, som benytter denne. Det er dog ikke tydeligt, at det er de konkrete referencer, som benytter netop denne metode, det kunne også have været andre publikationer, som omtaler samme problem. Denne misforståelse er mulig, da kritikken ikke er direkte. Altså: Det er ikke direkte nævnt, at disse publikationer konkret ignorerer disse faktorer, og at dette kan medføre en dårlig forvaltning af fiskeriet.

#### Informant 3

”Citaterne vil blive læst op til Siv.” (Eksempel 4)

Eksemplet indeholder en klar metodekritik med reference til pågældende artikel.

#### Informant 4

” Citaterne vil blive læst op til Siv.” Eksempel 5 s. 804

Publikationen skaber et link mellem andres generelle metodekritik til informantens eget felt. Kritikken gives direkte.

#### Informant 5

”Citaterne vil blive læst op til Siv.” (Eksempel 6)

Publikationen gør her opmærksom på, at nogle bestemte forskere var først med sekvensering af neuropeptider. Det gør han uden at citere dem, som har en anden opfattelse af det. Ved interviewet forklarede han, at en anden forsker fejlagtigt havde fået en nobelpris for at være den første. Informanten gjorde ved interviewet opmærksom på, at han ikke referer nogen, han er uenig med. Han skriver kun, hvordan det er, ikke hvordan det ikke er.

#### Informant 6:

" Citaterne vil blive læst op til Siv." (Eksempel 7)

Meget klar og direkte kritik med reference.

#### Informant 7

Fra en lærebog:

"Citaterne vil blive læst op til Siv." . Eksempel 8 s. 189.

Her ses en direkte kritik med reference til artiklerne, hvis metode ifølge denne bog bygger på en cirkulær argumentation. Det er værd at bemærke, at bøger ikke er en del af referencesystemet, og således er helt uden for systemet i Web Of Science og andre databaser.

#### Informant 8

Følgende er et eksempel på evaluering, som foregår uden for det publicerede system. Følgende er publiceret i et nyhedsbrev, som ifølge min research ikke er tilknyttet Web Of Science eller en anden webdatabase.

" Citaterne vil blive læst op til Siv." (Eksempel 9)

#### Informant 9

" Citaterne vil blive læst op til Siv..

Eksemplet er en indirekte kritik af referencens laboratoriearbejde. Det følgende afsnit argumenterer for, at det er informantens resultater, som er de mest pålidelige. (Eksempel 10)

#### Informant 10

" Citaterne vil blive læst op til Siv." (Eksempel 11)

Citatet er igen en indirekte kritik. Det kan være et eksempel på en kritik, som er nem at forstå, når man er fortrolig med feltets metoder, men meget svær at forstå, hvis ikke man er. Jeg forstår den ikke helt, da det ikke fremgår tydeligt, om det så er bedst med et med en hurtig respons, eller om dette i virkeligheden er vildledende.

#### Informant 11:

Denne informant har givet information om en konference, han har arrangeret for at forsøge at vinde lydhørhed overfor en kritik af et centralt dogme indenfor hans felt. Her understregede informanten, at der ikke skulle komme noget skriftligt ud om den, at den kun var for de bedste, og dem som turde være kritiske inden for feltet. Han ville heller ikke have at noget skulle publiceres om det efterfølgende, for det ville gøre at forskerne på konferencen ikke ville turde udtale sig frit.

Det er den mest hemmeligholdte kritik, jeg indtil nu er stødt på. Argumentet for hemmeligholdelsen er, at folk ellers ikke tør deltage i konferencen, eller ikke tør ytre sig kritisk imod den almindelige opfattelse af bestemte videnskabelige forestillinger.

### ***Opsamling på citatanalyse***

Af de viste eksempler fremgår det, at der er meget forskellige måder at evaluere videnskabelige publikationer på. Eksemplerne er ikke repræsentative, da de er de eksempler, som informanterne kom til at tænke på, som et godt eksempel på, at de har udtrykt sig kritisk. Derfor er en del af eksemplerne udtryk for en tydelig kritik, men selv i dette udvalg af kritiske formuleringer ses det, at det i nogle tilfælde kan være svært, at læse og forstå en kritik, selv hvis man opsøger den aktivt.

Af de få informationer jeg fik under surveyet (6 informanter er noteret for at have udtalt sig om det), fremgår det, at den mest almindelige måde at publicere kritik på er den indirekte, hvor man altså ikke skriver "jeg mener, det er forkert", men hvor man gør opmærksom på, at der er forskellige måder at gøre det på, og eventuelt argumenterer for at den metode, man selv har brugt, er særligt god på nogle bestemte områder uden at kritisere den anden metode. Både direkte og indirekte publiceret kritik er medtaget i det kvantitative survey som værende "publiceret kritik".

## Diskussion

Selvom peer-review-systemet anses for at være en god sikkerhed mod fejl og dårlig forskning, så er der en generel erkendelse af, at der findes diskutabile data og metoder i publicerede artikler indenfor flere videnskabelige områder (Wright & Armstrong, 2008; Jasanoff, 1994; Welles, 2009; Chan *et al.*, 2008). Det fremgår også af mit resultatafsnit, for de fleste informanter svarer, at de finder ting, de mener, kan kritiseres, og som har potentielle konsekvenser for den biologiske konklusion i mellem 1 og 15 % af de artikler, de læser. Jasanoff (1994) argumenterer for, at peer review fungerer bedst inden for områder, hvor der er enighed om teorier og metoder. Jeg er overbevist om, at hun har ret, idet jeg mener, at praktiseringen af Mertons norm (Merton, 1942) om Disinterestedness i nutidig naturvidenskab er usikker (Prpi\_, 1998). Forskernes position mellem fagfæller er afhængig af, at deres forskning og dennes grundlag bliver set som en succes. I min undersøgelse fremgår det, at forskerne sjældent opsøger kritik af den artikel de læser via et citeringssystem (Web of Science, PubMed eller lignende). Det er for mig klart, at der er brug for et struktureret evalueringssystem til at tilvejebringe en relevant evaluering af den allerede publicerede forskning i direkte sammenhæng med den kritiserede artikel (Garfield, 1987; Welles, 2009).

### ***Motivation i et tidsskriftssystem med bibliometriske effektivitets målinger***

I det publiceringssystem, vi har nu, er det tydeligt, at der som nævnt i indledningen er en nærværende og håndgribelig motivation for at publicere. Denne motivation synes at blive stærkere og stærkere, som fokus på bibliometriske indikatorer vokser. Det tilsidesætter logisk en anden form for videnskabelig praksis, herunder stillingtagen og måske særlig kritisk evaluering af andres forskning. Derfor har jeg fundet det vigtigt at undersøge forskernes konkrete handlinger og normerne for konkret handling, for at undersøge et modspil eller medspil til evaluerende adfærd.

Når jeg har diskuteret mit projekt med interesserede mennesker, har de ofte sendt mig publikationer om bibliometri som forslag til litteratur. Denne form for kvantitativ evaluering kan man argumentere for er nødvendig. Jeg vil ikke gå ind i den diskussion her. Jeg fokuserer i dette studie på sandheden af den viden, vi opnår og dermed publicerer, ikke på hvem der skal have penge eller status. Når det er sagt, så må jeg forholde mig til, at når man taler om evaluering af forskning, så tænker alle bibliometri. Især må jeg forholde mig til de konsekvenser, det har for udviklingen af tidsskriftssystemet og konsekvensen for den enkelte forskers motivation. Når en stor del af motivationen for den enkelte forsker og det enkelte tidsskrift er bibliometriske målinger, så har det naturligvis konsekvenser for den måde tidsskrifter og forskere handler. Der er et pres for at publicere, og så er der et pres for at holde sig opdateret, så man citerer de rigtige, og ikke laver forskning som er en gentagelse af andres forskning. Dette vil influere på

forskernes evalueringsadfærd således, at de læser mange artikler (Tenopir *et al.*, 2003) og producerer mange publikationer med størst mulighed for at komme i et high impact tidsskrift (Kokko & Sutherland, 1999; Lawrence, 2007).

Det, som har været mit udgangspunkt for dette studie, er, hvordan den kvalitative evaluering udfoldes under disse forhold, og om der eventuelt er nogle normer, som virker imod dette publikations pres. Jeg havde en forestilling om, at det måtte være nemt at finde nogle opstillede normer for videnskabeligt arbejde, og altså også for evaluering af de antagelser, man som forsker bygger sin forskning på. Men jeg har mest fundet litteratur, som ikke beskriver normer for konkrete handlinger, men beskriver mere værdibaserede holdninger (Mcfarlane & Cheng, 2008; Merton, 1942). Den mere kritiske indstilling til forskningen selv bliver eksplicit introduceret med Kuhns Videnskabens revolutioner (Kuhn, 1962), som godt nok ikke handler om normer. Strukturen for, hvorledes revolutioner skabes eller holdes tilbage, beskrives med udvalgte eksempler, en metode man kritisk kan kalde "big mens' science", altså analysen af forskning baseres på de forskere og opdagelser, som menes at være af særlig stor betydning frem for anden forskning. Hos Kuhn skelnes mellem normalvidenskab og revolutionerne, hvor revolutionernes struktur beskrives, som værende det ubetinget mest betydningsfulde for ny erkendelse. Det interessante i denne sammenhæng er, at Kuhn gør op med forestillingen om, at videnskabskvinden/manden kun forholder sig til facts og på en objektiv måde forholder sig til dem. Her skifter fokus altså fra at være videnskaben mod religion og politik med mere til at være videnskaben overfor sig selv. Der bliver sat spørgsmålstejn ved sandheden i videnskaben. Det er i denne bølge, jeg bevæger mig, når jeg i indledningen skriver, at for mig er sandhed relativt, altså jo mere forskere forholder sig kritisk til hinandens forskning, jo mere de evaluerer hinandens metoder, desto mere sandhedsværdi lægger jeg i den. Kuhn mente, at normalvidenskab ikke stræber efter erfaringsmæssige eller teoretisk set afgørende nyheder, og at den ikke finder nogen, når den er vellykket (Kuhn 1962). Derfor fokuserer han på paradigmeskift. Jeg mener, at dette er en grov forenkling eller måske en vellykket provokation. Jeg finder normalvidenskab umådeligt interessant, og da samfundet bruger forskningsmidler på som udgangspunkt at finansiere normalvidenskab, så er det jo, fordi vi netop har en forventning og en erfaring for, at den netop bidrager med nyheder, som vi kan bruge til noget i vores samfundsudvikling. Dette forklarer min empiriske fremgangsmetode med tilfældigt udvalgte informanter. På denne måde møder jeg netop et tilfældigt udvalg af forskere, som logisk primært vil beskæftige sig indenfor normalvidenskab indenfor deres område. Jeg undersøger den faglige evaluering i normalvidenskab, fordi jeg mener, at den underbygger kvaliteten og efter min mening sandheden i den forskning, som udføres.

## ***Forventning til læsning - forskel på molekylærbiologer og økologer***

I min undersøgelse har jeg fundet forskel på molekylærbiologer med hensyn til læsning (Figur 1-3).

Molekylærbiologerne læser ikke statistik og metodeafsnit så grundigt, som økologer. Dog er ingen af grupperne konsekvent grundige med begge afsnit. Jeg havde i en af mine interviews en interessant samtale med en molekylærbiolog, som efter vedkommendes egne oplysninger var specialist i statistik i forhold til sine kolleger. Han havde en bedre baggrund til at vurdere statistikken og gjorde det grundigere end sine kolleger. Han mente, at selv om statistikken var simpel inden for hans felt, blev der ofte brugt en forkert standard error for eksempel. Ud fra dette mener jeg ikke, at forskellen i læseadfærd skal udmønte sig i forskellige anbefalinger til ændringer i systemet. Både molekylærbiologer og økologer har, om end i forskellig grad, forskere, som tager kritisk stilling til andres publikationer. Det, som er vigtigt for mig er, at påskønne denne adfærd, og jeg mener ikke umiddelbart, det skal gøres forskelligt inden for de to grupper ud fra de resultater jeg har.

Derimod kan denne undersøgelse være med til at bekræfte, at molekylærbiologers fokus på statistik er meget minimalt. Molekylærbiologer, som arbejder under individniveau, fokuserer mere på kvalitativt at forstå de biologiske mekanismer, de arbejder med, i stedet for (som økologer) at lave modeller, som kvantitativt beskriver systemer med statistik som en vigtig del af argumentationen (Winther, 2006 og Winther, 2009). Ikke desto mindre bruger de fleste molekylærbiologer statistik eller læser artikler med statistik i alligevel, hvilket kan give problemer, som beskrevet af informanten ovenfor.

I en anden kontekst kan man overraskes over, at det lader til at netop molekylærbiologer i høj grad springer læsning af statistik og metode over. Molekylærbiologerne er knyttet til den medicinske tradition, hvor man har et stort fokus på, at læsning af og søgning om referencer skal være i orden. Når man refererer til en artikel, skal man ikke bare læse abstractet, men altså også eksempelvis metode og statistik, hvilket er beskrevet i "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" (ICMJE, 2008). Man skal undgå brug af abstracts som referencer. Forfattere er ansvarlige for, at ingen af de brugte referencer er blevet trukket tilbage. En del af grunden til, at dette tilsyneladende ikke har den ønskede effekt, kan blandt andet være, at der i PubMed – databasen er en automatisk pop up funktion, hvis artiklen, man læser, er blevet trukket tilbage. Den slags teknologiske mekanismer er efter min mening en god løsning, men kan måske have den konsekvens, at den enkelte læsers opmærksomhed over for selv at bedømme referencen svækkes. Hvis der automatisk kommer et pop up, hvis artiklen bliver trukket tilbage, hvilket sker meget sjældent, så kan nogen tage det, som en slags garant for, at det, som der ikke kommer en pop up på, ikke er blevet kritiseret eller ikke har grund til at blive det. En informant i min undersøgelse gav i udtryk for denne



tankegang. Altså er det potentielle problem ved den slags mekanismer ikke det, som bliver trukket tilbage, men alt det, som kan kritiseres fundamentalt men ikke nok til, at det bliver trukket tilbage.

### ***Den fundne læseadfærd i perspektiv***

En del af mit formål med undersøgelsen var, at give en deskriptiv fremstilling af de parametre informanterne har opgivet. Det, som er ulempen, ved den type fremstilling er, at diskussionen efterfølgende kan mangle sammenligningsgrundlag, såfremt ingen andre har undersøgt samme fænomen. Det kan være svært at svare på om det er meget eller lidt at læse metodeafsnittet i eksempelvis 20 % af de artikler, man læser. Det, som er interessant i denne sammenhæng er, at man ikke automatisk kan gå ud fra, at forskere vurderer, hvorledes undersøgelsen, de refererer til eller bygger videre på, er udført (Figur 1-3). Denne adfærd er meget forskellig blandt de forskellige forskere. Dette afspejler, at der ikke er klare normer angående læsning af artikler, man vil referere til, men kan også afspejle at der er nogle faktorer, som virker modsat normerne (se figur 9). Normer kan ses som en konstruktion af almindelig videnskabelig praksis og diskussioner og forhandlinger mellem forskellige grupper (Tuunainen & Knuutila, 2008), og netop den definition viser måske, hvorfor normerne er så flydende på området. Det er mit indtryk ud fra interviewene, at hverken diskussion eller forhandling foregår. Praksis synes at være op til den enkelte forsker.

Noget af den litteratur, jeg har fundet, som beskriver omfanget af forskeres læsning, beskæftiger sig med citationsfejl. En af studierne har ud fra gentagelser af fejl (stavefejl eller lignende) i litteraturliste udregnet, at forskere kun i 20 % af tilfældene overhovedet læser noget af den artikel, de refererer, idet det er meget udbredt at citationsfejl bliver gentaget (Simkin & Roychowdhury, 2003). En grundantagelse i den metode er, at forskere ikke kopierer citeringer fra andres referencelister, hvis de er i besiddelse af eller læser originalartiklen. Da jeg havde læst artiklen, opdagede jeg, at der hverken var opgivet tidsskrift eller sidetal i udprintet. Min første tanke var, at dem kunne jeg jo finde i den publikation, hvor jeg først fandt artiklen refereret. Altså havde jeg netop udført den adfærd, som artiklen antager, er meget usandsynlig. En anden metode bliver brugt af Wright & Armstrong (2008). Wright og Armstrong har gennemgået referencerne på en artikel, de selv har skrevet, kvalitativt, altså vurderet i hvert enkelt tilfælde om referencen i pågældende artikel er brugt korrekt. Dette er en interessant metode at benytte, idet den vil belyse noget, som min undersøgelse ikke belyser, nemlig hvor godt et match der er mellem hvad forfatterne mener, der står i artiklen, og hvad dem, som bruger artiklen, mener der står. Min undersøgelse belyser i stedet, for hvor ofte forskerne som læsere mener, at en pågældende reference er brugt forkert.

## ***Communisme - vidensdeling i praksis***

Det er svært at svare på, om det er lidt eller meget at kommunikere sin kritik ud for eksempel 1 % af de gange, man finder noget, som er af potentiel betydning for den biologiske konklusion (figur 6). Men det er helt umiddelbart ærgerligt, at der ikke er en større del af den viden, som opstår ved læsning af artikler og i diskussion med nære kolleger, som er tilgængelig for andre. Husk på at tallene for publiceret kritik er inklusiv indirekte kritik, hvor det (jævnfør citatanalysen på side 30) er usikkert hvor mange af læserne, særligt de nye indenfor feltet, som forstår evalueringen. Det er et tegn på, at Mertons Communisme (Merton, 1942 og Kalleberg, 1997), forstået som at dele sin viden med andre, i hvert fald på dette område ikke er udtalt i biologisk forskning. Det, at der er så stor forskel på, hvor ofte forskerne kommunikerer statistik- og metodekritik ud (figur 6), vidner om de ikke-eksisterende normer på området (figur 10).

## ***Indførelse af personlig motivering for at kommunikere evaluering ud***

Mit formål med denne undersøgelse er ikke at kritisere forskernes adfærd, for det mener jeg ikke tjener noget formål i denne sammenhæng. Caellegh (2002) argumenterer for, at de videnskabelige samfund (societies) skal påtage sig et ansvar overfor at bestemme sig for, hvad der er god publikationspraksis, og oplære medlemmer i denne praksis. Det er nok lidt svært at få dette indført, da det er meget forskelligt, hvor organiserede forskersamfundene er, og det vil også være meget forskelligt, hvor motiverede de er for at arbejde med disse problemstillinger. I stedet vil jeg foreslå at ændre motiveringen i systemet, så publiceringssystemet i højere grad kan motivere forskerne til at dele deres evalueringer. For mig at se, betyder den lave grad af kommunikation i kombination med de uklare normer, at der skal tænkes i, at indføre en motivation for at evaluere andres artikler. Dette kan ske ved at indføre statuspoint for at publicere kritik, og gøre kritikken mere fremtrædende i citeringssystemet, som også beskrevet under afsnittet "Justering af tidsskriftsystem" med direkte link til stedet i den pågældende artikel, et rødt tal som man kan finde nogle steder i Wikipedia, som indikerer, at man kan klikke en kritik op, som relaterer sig til emnet. Et system lignende det man kan finde i Biology Direct. En sådan kritik kunne i forskellige sammenhænge (for eksempel citations- og publiceringsanalyser, opgørelser over forskningsaktivitet og CV'er) tælle som en halv eller en kvart publicering eller et andet mål, som relaterer sig til antal af publiceringer. Ved at rate det højt, ville man skabe en mulighed for at få en karriere, som i høj grad byggede på at evaluere andres forskning og komme med konstruktive bidrag. Man kunne fra tidsskrifternes side opstille klare normer, om at kommentarerne skulle være saglige (ikke personlige) og konstruktive, og en editor skulle kigge dem igennem, eventuelt sammen med et svar eller forsvar fra forfatterens side. Editoren ville her få mere magt og arbejde i forhold til peer reviewere. Det er for mig at se vigtigt, at det ikke er nære fagfæller (peers), som bedømmer hinandens kritik, men at en lidt mere udenforstående men

vidende person, som en editor typisk er, bedømmer det ud fra specifikke kriterier uden selv at være en del af uenighederne.

Hvad ville ulemperne ved et sådant system så være? Den største ulempe vil være, at tid og ressourcer vil gå fra et forskningsfelts nye forskning til at evaluere forskningen. Det vil være en vurdering, om man foretrækker det ene frem for det andet. Personligt ville jeg vælge et kommentarsystem til, ligesom flertallet af informanterne melder ud (se afsnittet "Justering af tidsskriftssystem"). Fordi der er så mange forskere ansat, så mange som der aldrig har været før gennem historien. Vores samlede erkendelse må være afhængig af, om vi formår at samle vores viden. Det er paradoksalt, at vi ikke har tid til at diskutere den viden, vi producerer. Jeg mener, at grunden til, at forskerne føler, at de ikke har tid til det, må ligge i den individuelle opfattelse eller eventuelt for fagområdets opfattelse. En opfattelse af at man har meget travlt, hvis man skal opretholde en interesse og en pengestrøm til ens område. Men for at tage stilling til den samlede vidensudvikling, må man løfte sig fra den enkelte forskers synsvinkel, og se på systemets reguleringsmuligheder. Vi har et unikt redskab i internettet, og dets mange positive muligheder bør implementeres i højere grad i tidsskriftssystemet, end de er i dag.

## **Forbedringsmuligheder af metoden**

Jeg mener, at den kombination af metoder, jeg har brugt, har et stort potentiale, til at undersøge videnskabsteoretiske problemstillinger, som relaterer sig direkte til forskernes adfærd og syn på forskning og deres arbejde. Men der er nogle ting, som jeg ville ønske, jeg havde gjort anderledes.

### Komplekse spørgsmål

Spørgsmålet angående læsning af referencer var, som følger: "Hvor ofte læser du relevante referencer til de artikler, du læser?". Det var for komplekst, det var (forståeligt nok) en for kompliceret tankerække at udføre for informanterne. Det var det også, da de skulle vurdere hvor ofte, det havde betydning for den biologiske konklusion i den artikel, referencerne var referencer i. Jeg spurgte på denne måde, fordi jeg valgte at bruge samme kriterier og spørgemåde angående metode, statistik og referencer. Det forstyrrede interviewet en del, og gør desværre at jeg ikke kan bruge disse resultater til så meget, idet det egentligt var mere interessant, om de kritikpunkter, informanterne havde angående referencerne, havde betydning for biologiske forståelse generelt.

## Litteraturliste

- Andersen, H., Emmeche, C., Norup, M. & Sandøe, P., 2006: *Videnskabsteori for de biologiske fag*, Forlaget Biofolia
- Bordons, M., Fernandez, M. T. & Gómez, I.: "Advantages and limitations in the use of impact factor measures for the assessment of research performance in peripheral country." *Scientometrics*, 53: 195-206
- Calleigh, A. S., 2003: "Roles for Scientific Societies in Promoting Integrity in Publication Ethics", *Science and engineering Ethics*, 9, 221-241.
- Chan, A., Hróbartsson, A., Haahr, M. T., Gøtzche, P. C. & Altman, D. G., 2004: "Empirical Evidence for selective Reporting of Outcomes in Randomized Trials", *JAMA*, 291, 2457-2465.
- Dahler-Larsen, P, 2003: *At fremstille kvalitative data*, Syddansk Universitetsforlag
- Fisker, K. et al., 1995: *Praktisk statistik for samfundsvidenskaberne*, Akademisk Forlag
- Garfield, E., 1987: "How to Avoid Spreading Error: Scientists Must Search for corrections", *The Scientist* 1(19):9; reprinted in: Garfield, E.: *Essays of an Information Scientist, 14: Science Reviews, Journalism Inventiveness and Other Essays*, p., 255-257. Philadelphia: ISI Press, 1991.
- Holstein, B., 2001: "Triangulering – metoderedskab og validitetsinstrument" " i Lunde, I. M. & Ramhøj, P. (eds.), *Humanistisk forskning inden for sundhedsvidenskab - Kvalitative metoder*. København: Akademisk Forlag.
- ICMJE, 2008: *Uniform requirements for Manuscripts Submitted to Biomedical Journals*. The International Committee of Medical Journal Editors, <http://www.icmje.org/>.
- Jasanoff, S., 1994: *The fifth branch, Science advisers as policymakers*. Harvard University Press.
- Kalleberg, R. 1997, "A Reconstruction of the Ethos of Science", *Journal of Classical Sociology* 7(2): 137-160
- Kokko, H. & Sutherland, W. J.: "What does impact factors tell us?", *Trends in Ecology and Evolution*, 14, 382-384.
- Kuhn, T. S., 1962: *Videnskabens revolutioner* (ny udgave ved Petersen, S. A.), København: Fremad, 1995.

- Kuhn, T. S., 1973: "Objektivitet, værdidomme og valg af teori", i Kuhn: *Videnskabens revolutioner* (ny udgave ved Petersen, S. A.). København: Fremad, 1995, 262-285
- Kvale, S., 2003: *Interview: en introduktion til det kvalitative forskningsinterview*, København: Hans Reitzels forlag
- Lawrence, P. A., 2007: "The mismeasurement of science", *Current biology* 17, R583-R585
- Macfarlane, B. & Cheng, M. 2008, "Communism, Universalism and Disinterestedness: Re-examining Contemporary Support among Academics for Merton's Scientific Norms", *J Acad Ethics* 6: 67–78
- Merton, R. K., 1942, "The normative structure of science", in: Merton: *The Sociology of Science*, ed. Storer, N. W, Chicago: The university of Chicago Press, 1973.
- Neumann, W. L., 1997: *Social research methods – qualitative and quantitative approaches*. (3rd. ed.). Needham Heights, MA: Allyn and Bacon.
- Prpi, K., 1998: "Science Ethics: A study of Eminent Scientists' Professional Values". *Scientometrics*, 43, 269-298
- Rathje, I. W., 2006: "Den teknologiske udvikling i den danske fiskeflåde, 1980 – 2004", Bachelorprojekt, Biologi, Københavns Universitet
- Robillard, R. J., 2002: "How scientists Use Journals: Electronic and Print" [Annual meeting reports], *Science Editor*, 25, 192
- Schwandt, T. A., 2007: *Dictionary of Qualitative Inquiry*. SAGE Publications
- Siegel, S. & Castellan, J. N., 1988: *Nonparametric Statistics for the Behavioural Sciences*. McGraw-Hill Book Company
- Simkin, M. V. & Roychowdhury V. P., 2003: "Read before you cite!" *Complex Systems*, 14 269-273
- Tuunainen, J. & Knuutila, T., 2008: "Determining the Norms of Science: From Epistemological Criteria to Local Struggle on Organizational Rules?" Chapter 10 in Välimaa, J. & Ylijoki, O.-H., (eds.): *Cultural Perspectives on Higher Education*. Springer.
- Tenopir, C., King, D. W., Boyce, P., Grayson, M., Zhang, Y. and Ebu, M., 2003: "Patterns of Journal Use by Scientists through Three Evolutionary Phases", *D-Lib Magazine*, 9, number 5

Welles, J. F., 2009: "Correcting Errors in Science", *Journal of Information Ethics*, 18, 16-20

Winther, R. G., 2006: "Parts and theories in compositional biology", *Biology and Philosophy* 21. 471-499

Winther, R. G., 2009: A Brief Description of Compositional and Formal Biology, upubliceret materiale.

Wright, M. & Armstrong, J. S., 2008: "The Ombudsman: Verification of Citations: Faulty Towers of knowledge?" *Interfaces*, 38, 125-139

## **Bilag**

Bilag 1: Spørgeskema til den kvantitative undersøgelse

Bilag 2: Transskriptioner af de kvalitative interviews

Bilag 3: Citatanalysens eksempler