

## **ABSTRACT**

In this assignment we study correlations between scientific articles and their references. The study is related to bibliometrics and uses the method citationsanalysis. It consists of an empirical and a theoretical part. The empirical part consists of 45 selected articles and is performed in Dialog database 7, 34 and 439.

With results from the co-citationanalysis we explore the differences between three selected fields within the natural sciences, social sciences and humanities. We analyze and discuss these results in the light of theory and concepts related to citer motivations, genre analysis, and general differences in research fields. And we determine whether these results can be used to assess the relevance of the references to the citing article. Our main purpose is to clarify tendencies and differences within the fields.

# INDHOLD

INTRODUKTION:.....	4	fælles
Problemformulering.....	5	fælles
METODE.....	6	
Empirisk metode.....	6	CPJ
Teoretisk metode.....	8	LSR
TEORI.....	10	
Citationer og Co-citationsanalyse .....	10	
Citationer og referencer.....	10	CPJ
Citationsanalyse .....	11	CPJ
Problemer ved citationsanalyse.....	13	CPJ
Karakteristik af de tre store videnskabsområder .....	17	LSR
Citationsadfærd.....	19	LSR
Garfield (1965) .....	19	LSR
Small(1978).....	20	LSR
Brooks (1986) .....	21	LSR
Case & Miller (2011).....	22	LSR
Hvornår skal man citere? .....	24	CPJ
Citationskontekst.....	26	
Hvor & hvorfor? - Den videnskabelige artikel som genre.....	26	LSR
Hvordan? – præsentation af referencer .....	28	CPJ
Hvornår? – Hvad betyder referencernes alder?.....	30	CPJ
Hvor mange? – om antal forekomster .....	33	CPJ
PRÆSENTATION AF RESULTATER .....	35	fælles
Karakteristik af keredokumenterne.....	39	CPJ
Filosofi .....	39	CPJ
Fysik .....	40	CPJ
Sociologi.....	40	LSR
ANALYSE OG DISKUSSION .....	42	
Filosofi .....	45	CPJ
Fysik .....	46	CPJ
Sociologi.....	48	LSR
Fejlmargin.....	49	fælles
Hvad citeres hvor?.....	50	LSR

Betydningen af ORL .....	52	CPJ
Valg af referencer.....	53	fælles
Standardisering vs. obliteration .....	56	CPJ
I videnskabsteoretisk perspektiv.....	57	fælles
KONKLUSION .....	60	fælles
Metodekritik.....	62	fælles
KILDEFORTEGNELSE .....	63	

## INTRODUKTION:

Det er et velkendt fænomen, at man citerer andres værker i sit eget videnskabelige værk. Hvad er rationalet for denne praksis? Er det konventionelt forventet? At hvis der ikke er et vist antal af referencer, så er den viden, der formuleres ikke acceptabel? Er det en stilistisk metode til at vise, hvilken genre vi har med at gøre? Eller er det en etisk forpligtelse – en anerkendelse af tidligere værker, eller en tilkendegivelse af andres intellektuelle ejendom? Bliver citationer brugt som et overtalelsesredskab, sådan at læseren ud fra autoritative eller velkendte referencer, skal blive overbevist om rigtigheden af den nyopdagelse viden? Eller bruges citationer som redskaber til at skabe et 'research-space', hvor ingen før har sat deres fødder?

Dette er spørgsmål, der gennem mange år har floreret inden for feltet bibliometri, specielt i forhold til de grundlæggende metoder publikationsanalyse og citationsanalyse, men også i forhold til forskellige aspekter af begrebet citationsadfærd. Citationer opstår altid i en kontekst, og det er derfor problematisk at give et entydigt svar. Med denne opgave vil vi ikke kaste os ud i at besvare alle disse spørgsmål, men de bliver taget i betragtning i forhold til vores undersøgelse, hvori vi vil gennemgå forskellige aspekter og tendenser, der kan være med til at kaste lys over, hvorvidt nogle referencer kan betragtes som værende af højere relevans for det citerende dokument end andre.

Til at undersøge dette gør vi brug af en co-citationsanalyse, men ikke i almindelig forstand. Vi vil derimod foretage den på dokumenter der allerede er koblet til form af direkte citering. Nærmere betegnet vil vi undersøge, hvordan artikler er co-citerede med de referencer, som forskeren selv har valgt at bruge i artiklen og dermed opnå en udefrakommende vurdering på forholdet mellem dem.

Dette vil vi gøre inden for tre forskellige felter, der i denne opgave repræsenterer de tre store videnskabelige felter: naturvidenskab, samfundsvidenskab og humaniora.

Ud fra forskellig teori om citationsanalyse generelt, citationsadfærd i særdeleshed, genreteori om den videnskabelige artikel og forskellige traditioner inden for diverse felter, vil vi undersøge, om der kan ses nogle tendenser til, hvordan der citeres. Både i forhold til den udregnede co-citationsfrekvens, antal gange referencen forekommer i artiklen, og om den citerede artikels alder har nogen betydning i forhold til, hvordan den optræder i artiklen.

### *PROBLEMMFORMULERING*

- ***Hvordan er videnskabelige artikler inden for tre forskellige fagdiscipliner co-citerede med deres egne referencer?***
- ***Hvilke forhold, i forbindelse med citationsadfærd, karakteriserer de tre videnskabsområder?***

Derudover arbejder vi ud fra følgende underspørgsmål:

- Kan man ud fra denne form for co-citationsanalyse, vurdere den enkelte references relevans for det citerende dokument?
- Hvordan gør forskere brug af referencer i videnskabelige artikler?

Og hypoteser:

1. De referencer der er mest co-citerede med vores kerneartikler, vil forekomme i begyndelsen af dokumentet.
2. Der vil kunne identificeres forskelle inden for de tre felter, blandt andet på referencernes alder, antal forekomster og den procentvise co-citationsfrekvens.

## **METODE**

For at kaste lys over de problemstillinger vi ønskede at undersøge, havde vi brug for en metode til at fremfinde de referencer, vi selv havde en hypotese om, ville være de mest relevante. Her valgte vi at foretage en co-citationsanalyse, for at finde frem til, hvordan artikler er co-citerede med deres egne referencer. De referencer, der havde flest co-citationer med artiklen, var dem vi antog for at være mest relevante.

### *EMPIRISK METODE*

Empirien bestod af 45 artikler der repræsenterede tre forskellige fagdiscipliner. 15 artikler inden for fysik, 15 inden for filosofi og 15 inden for feltet sociologi. Baggrunden for dette valg er, at vi ønskede at sammenligne forskellige forskningsfelter. Vi valgte med udgangspunkt i Ken Hylands(1999) undersøgelse at søge artiklerne i de tidsskrifter, der her er udpeget til at være de 10 førende inden for deres felt. Artiklerne blev herefter fundet i citationsbasen Web of Science, hvor vi gennemførte en kombineret søgning af alle dokumenter produceret i et af disse tidsskrifter inden for en periode fra 1995 – 2005, vi udvidede dog søgesættet fra filosofi med 10 år (1985-2005). Vi endte således med tre forskellige søgesæt. Derefter afgrænsede vi søgningerne til at bestå udelukkende af artikler, da vores formål er at sammenligne videnskabelige artikler, og vi ikke ønskede at undersøge reviews eller andre former for dokumenter. Herefter valgte vi at fokusere på de mest citerede, for at artiklerne kunne have citationer nok til at gennemføre co-citationsanalysen. Dernæst var det vigtigt at artiklerne inden for hvert felt havde samme antal referencer, for at gøre dem sammenlignelige. Vi besluttede af referencelisterne måtte variere med maksimalt 4 referencer for stadig at kunne opretholde et vist antal citationer. Derefter gik vi manuelt søgeresultaterne igennem og fandt frem til de 15 artikler der først imødekom disse kriterier.

Efter udvælgelsen af artiklerne anvendte vi databaseværten Dialog til at foretage de søgninger, der var nødvendige for en brugbar empiri. Søgningerne på artiklerne fra feltet fysik blev foretaget Science Citation Index, database nr. 34, søgninger i forbindelse med artiklerne fra sociologi blev foretaget i Social Science Citation Index database 7 og de filosofiske artiklers søgninger blev gennemført i

Arts&Humanities database 439. Disse er alle citationsindekser, hvilket er påkrævet for at kunne gennemføre en co-citationsanalyse.

Første skridt var at skabe en søgestreng for den pågældende kerneartikel, hvor denne fremgik som en reference. Dette blev gjort ved at 'expande' forfatternavn, udgivelsesår og tidsskriftets navn som 'cited reference' og derefter vælge den rette referencestreng. Vi gjorde eksempelvis brug af denne kommando (filosofi artikel 7):

```
E CR = 'WALDRON J, 1992, V103, P14, ETHICS'
```

Derefter søgte vi efter den samme artikel, men denne gang ikke som en reference:

```
S au=waldron j and jn=ethics and py=1992
```

Hvis der var flere søgninger der matchede ovenstående, så fik vi dem enten vist ved at bruge kommandoen 'type' eller brugte et ord fra titlen som søgningssuffix (/ti).

Herefter brugte vi 'rank cr'kommandoen for at få frembragt en liste over artiklens referencer. For at gemme de frembragte referencer som individuelle søgestrengene tastede vi '1-19' hvilket var det antal der fremkom i netop dette eksempel, og herefter 'exit', 'y' for at forlade 'rank'. Dernæst brugte vi kommandoen 'exs' for at kunne udføre søgningen i basen. Således havde vi en søgbar liste over hver enkelt af artiklens referencer. Så var det blot at kombinere den første søgning med hver af de enkelte referencer, på denne måde fandt vi frem til det tal, der indikerede antallet af co-citationer referencerne imellem. Denne fremgangsmåde blev udført på hver af de 45 udvalgte artikler og udgør den samlede empiri for co-citationsanalysen.

Efterfølgende placerede vi hver artikels referencer i den rækkefølge de optrådte i artiklen og tildelte derefter referencenumre.

Tallene for hver artikel varierede meget og for at få et mere præcist billede af forholdet mellem reference og kerneartikel udregnede vi den procentmæssige co-

citationsfrekvens. Til dette anvendte vi en formel præsenteret af Garfield(1980, i: Osareh,1996):

$$S = \frac{co - citationsofA + B}{(totalcitationsA + B) - (co - citationsofA + B)}$$

Endvidere gik de empiriske undersøgelser ud på at finde frem til, hvor mange gange de enkelte referencer blev nævnt i den tilhørende artikel Dette blev gjort ved manuel optælling for hver af de 45 artikler, hvorefter disse data blev placeret ud fra de tilhørende referencer. Desuden ønskede vi at undersøge artiklernes brug af orienterende referencelister (Hargens, 2000) og derfor identificerede vi disse og noterede, hvilke referencer der optrådte i sådanne lister. Afslutningsvis beholdte vi referencernes udgivelsesår som separat datakolonne for at kunne bruge disse i vores analyse.

Dette blev til sidst sammenført i tre excelark og opstilles således, at vi havde grundlag for at gennemføre en overordnet vurdering af de empiriske undersøgelser.

### *TEORETISK METODE*

Vores teoretiske udgangspunkt er blevet dannet blandt andet på baggrund af tidligere undersøgelser af citationsanalyser og citationsadfærd, som vi vil gennemgå og uddybe. Derudover vil vi tage fat i den litteratur der beskriver emnet, diskuterer det, kritiserer det og bifalder det (Garfield (1965, 1975, 1996), MacRoberts&MacRoberts (1987), Osareh (1996), Smith (1981), Voos&Dagaev (1976)).

Ud fra denne litteratur definerer vi de begreber og metoder vi anvender, og hvad de har til formål at afklare. Vi vil samtidig kort komme ind på hvilke videnskabsteoretiske synsvinkler, der ligger til grund for disse metoder. Relevante begreber og undersøgelser i forbindelse med citationsadfærd vil blive gennemgået (Brooks (1986), Case & Miller (2011), Small (1978)), da vi mener det har stor betydning for vores undersøgelse. Her kan vi ud fra vores egne resultater spore os ind på hvilke faktorer, der kan have sammenhæng med betydningen af citationsadfærd. Dette vil yderligere blive analyseret og diskuteret med forskellige



perspektiver på disciplinære forskelle (Hargens (2000), Hellqvist (2010), Hyland (1999)).

Med John Swales'(1990) teorier vil vi redegøre for hvordan videnskabelige artikler er opbygget. En gennemgang af denne teori, vil yderligere kunne anskueliggøre, hvorfor vi har valgt at opsætte en hypotese om, at de mest relevante referencer optræder første gang i artiklers introduktion.

Vi vil endvidere inddrage undersøgelser der beskæftiger sig med at redegøre for forskeres citationsadfærd i forbindelse med referencernes placering, præsentationsteknik og alder.

Efter en præsentation af vores resultater, vil vi analysere og diskutere disse i forhold til den gennemgåede teori, som således skulle lede os frem til at kunne besvare opgavens overordnede problemstilling i en afsluttende konklusion.

# TEORI

## *CITATIONER OG CO-CITATIONSANALYSE*

Dette afsnit indeholder en gennemgang af begreberne citationer og referencer, samt citationsanalyse, som vil bidrage til forståelse af de bibliometriske metoder, vi gør brug af og som vi analyserer og diskuterer senere i opgaven. Der vil specielt være fokus på co-citationsanalyse, som er opgavens overordnede metode.

### CITATIONER OG REFERENCER.

Når en videnskabelig artikel produceres, tager den udgangspunkt i tidligere forskning på det område, den emnemæssigt forholder sig til. Artikler er generelt dannet på et eksisterende intellektuelt grundlag og kan ikke stå alene. Der bruges tidligere publiceret forskning til at underbygge argumenter, samt skabe ny viden (Smith, 1981), (Osareh, 1996).

I denne proces skaber forskere en form for indirekte videnskabelig kommunikation. Forfatteren af artiklen bruger henvisninger til tidligere værker og på denne måde repræsenterer referencelisten i et hvert videnskabeligt dokument forfatterens beslutninger om relationer, der opstår mellem det værk, der bliver til og værker af andre forfattere (Osareh, 1996). Det er vel at mærke udelukkende den skrivende forfatters synsvinkel på relaterede dokumenter, der kommet til udtryk ved at gennemse referencelister.

Eftersom referencelisten som udgangspunkt skal bruges til at klarlægge den intellektuelle baggrund for et værk, kan denne være af stor betydning, da alle udsagn og påstande, der optræder i et videnskabeligt værk skal kunne bekræftes, støttes op om, og derved kunne testes og verificeres. Den forskning, der beskrives i videnskabelige dokumenter, har ofte en teori eller et budskab som gerne skulle berige læseren og pointere nye ideer, opdagelser eller synsvinkler, og her er det vigtigt, at referencerne er korrekte og gives af reelle årsager.

Når en referenceliste produceres, giver forfatteren en optimalt set en reference til hvert enkelt værk benyttet i dokumentet; det enkelte dokument i referencelisten modtager derimod en citation fra det citerende dokument. Forskellen på

referencer og citationer er lille, men betydelig. Den mest simple forklaring er, at en reference den anerkendelse man giver et andet dokument, og en citation er en anerkendelse man modtager. (Osareh, 1996).

### CITATIONSANALYSE

Citationsanalyser handler grundlæggende om at tælle, hvor mange gange og hvorledes en forfatter, et dokument etc. er citeret. De tidlige citationsstudier var baseret på referencelister fra artikler, fra et relativt lille antal tidsskrifter og blev udgivet i trykt form. Ved hjælp af deelektroniske citationsindexer, er det i dag muligt at lave citationsanalyse på meget store mængder af data og sammenligne hele nationer og forskningsfelter (De Bellis, 2009).

Det er indekseringen af dokumenters referenceliste, der er grundlaget for citationsanalysen. De søgbare referencelister kan således bruges både retrospektivt og prospektivt, det er altså muligt både at finde en citeret samt en citerende artikel.

Citationsanalyser er en del af bibliometrien, der beskæftiger sig med at undersøge forholdene mellem citerede og citerende dokumenter (Smith, 1981). Antallet af referencer mellem dokumenter kan siges at afspejle, hvilken grad af tilhørsforhold dokumenter har til hinanden. Bellardo (1980-81 I: Osareh, 1996) forklarer det som en kvantitativ metode til at måle forholdet mellem videnskabelige forestillinger. Antagelsen af, at højt citerede dokumenter er vigtigere eller mere betydningsfulde end lavt citerede dokumenter, er essentiel for citationsanalyser, da dette danner grundlag for at udføre denne slags undersøgelser. Samme princip gælder forfattere og tidsskrifter, jo oftere de er citeret, jo mere indflydelsesrige betragtes de for at være. Denne opfattelse af citationer virker umiddelbart logisk. Hvis forskning ofte bliver anvendt af andre forskere, må man antage, at den har indflydelse på mange andre værker. Når dette er tilfældet, må den forskning, der har modtaget mange citationer, efterlade os med et indtryk af, hvad der er anerkendt. Derudover kan man i mange tilfælde bruge citationsanalyser som en indikator for emnemæssig beslægtethed mellem bl.a. dokumenter og forfattere, samt på hvilket niveau forskere, lande og så videre er placeret i forhold til hinanden.

Nicola De Bellis (2009) mener, at man kan betragte referencer som en slags ukontrollerede emneord, der skaber relationer mellem dokumenter, forfattere etc. Han pointerer, at en citation ikke bare er en kobling mellem dokumenter, men en kobling mellem ideer og argumenter, og derfor er vigtige redskaber til at klarlægge videnskabelige strukturer.

Der findes to former for citationsanalyser, den første kaldes bibliografisk kobling. Her undersøges hvor mange artikler, der bygger på samme artikelgrundlag, det vil sige deler referencer. Jo flere referencer dokumenterne har til fælles, jo højere bibliografisk koblingsfrekvens har de. Det antages til en vis grad at dokumenter, der bygger på samme tidligere forskning, må have en vis emnemæssig lighed, eller i hvert fald deler rødder inden for samme felt.

En anden metode er co-citationsanalyse. Her måles, hvorvidt dokumenter optræder i den samme referenceliste. Hvis to dokumenter er co-citerede skal begge dokumenter være nævnt i et tredje dokument's referenceliste. Jo flere gange de bliver det, jo højere co-citationsfrekvens har de. På denne måde kan co-citationer også være en indikator for lighed mellem dokumenterne.

Co-citationsanalyser kræver, for at analysen skal være succesfuld, at der gøres brug af relativt højt citerede dokumenter. Et dokument kan af god grund ikke være co-citeret med et andet, uden selv at være blevet citeret. Dette medfører, at der ikke kan foretages co-citationsanalyser på meget nye udgivelser, da disse endnu ikke har haft mulighed for at blive citeret i andre dokumenter

Henry Small undersøgte i 1973, hvordan co-citationer adskiller sig fra bibliografiske koblinger, og hvilken relation de har til fordel at måle. Hans resultater indikerede at bibliografiske koblinger ikke i samme grad var i stand til at finde frem til den emnemæssige lighed mellem dokumenter som co-citationsanalyser formåede. Han konstaterede dog, at en stærk co-citationsfrekvens dels afhænger af emnelighed, men også af dokumenternes fælles idegrundlag.

Nicola De Bellis (2009) påpeger sammenhængende, at dokumenter udelukkende co-citerede, fordi andre forskere har skabt enighed om betydelig forskning inden for grænserne af deres felt, og den intellektuelle baggrund der gøres brug af. Han

mener, at co-citationsanalyser i høj grad afslører de vigtigste kilder, koncepter, empiriske undersøgelser og metoder og dermed bedre kan afdække mønstre i forskningsfelter og fremhæve den overordnede videnskabsstruktur.

Citationsanalyser har yderligere den fordel ved at de, i forbindelse med tid og trend samt udvikling inden for forskningsfelter, kan siges at være variable og fremadskuende (Small 1973);(De Bellis 2009)

Selvom analyserne giver forskellige resultater, må det dog til en vis grad antages, at jo højere frekvenserne er, både inden for bibliografisk kobling og co-citationer, jo mere har dokumenterne til fælles. Dette til trods for, at der ikke eksisterer enighed om hvad dette fælles grundlag indebærer. Man kan dog pointere, at den forbindelse co-citationsanalyser danner, ikke afhænger af forfatterne til de citerede dokumenter, men derimod er en kobling, der afhænger af de citerende dokumenters forfattere. Det er dermed ikke sagt, at den ene metode er bedre end den anden, men det kan være værd at overveje hvilken, der bedst vil kunne bidrage til det man ønsker at belyse.

### *PROBLEMER VED CITATIONSANALYSE.*

Her vil de to videnskabsteoretiske positioner inden for citationsteori blive beskrevet i forhold til, hvilke problemer der kan være ved citationsanalyser.

Selv med de ovenstående aspekter taget i betragtning, virker citationsanalyser dog stadig til at være et yderst brugbart redskab inden for bibliometri, men der er dog en række problemer i forbindelse med metoderne. MacRoberts og MacRoberts (1989) har opstillet en liste over, hvilke problemer der kan opstå i forbindelse med citationsanalyser. De pointerer først, at det er vigtigt at være opmærksom på, at al kritik må sættes op imod det teoretiske grundlag.

*"Consequently, all the problems we discuss must be considered against the background of the theory that inform most types of citation analysis. This is the normative theory, which maintains that bibliographies are lists of influences and that authors cite in order to give credit where credit is due(...)"*

(MacRoberts&MacRoberts, 1989: 342)

De mener, at det er denne opfattelse af den videnskabelige kommunikationsproces der ligger til grund for citationsanalyser. Den er oprindeligt formuleret af Robert K. Merton(1973 I:De Bellis, 2009) og kaldes den videnskabelige norm. Teorien indebærer som citatet indikerer, at citationer generelt er en troværdig indikator for, hvilke andre værker der har haft indflydelse på det citerende værk. Merton påstår ikke, at forskere mekanisk følger en række normer i deres produktion af forskning. Han fastholder dog stadig, at alle forskere, til en vis grad, opererer ud fra en række etiske forskrifter. Disse normer opdeles i fire overordnede værdier som kaldes "the CUDOS norms":

COMMUNISM: Videnskabelige resultater skal være offentlige og kunne benyttes af alle. På baggrund af, at al forskning udspringer af tidligere forskning, har ingen mere ret til denne end andre.

UNIVERSALISM: tilsigter at al viden må holdes op imod et sæt objektive forudbestemte kriterier og gøres universal. Samtidig må al sandfærdig forskning betragtes ens uanset baggrund

DISINTERESTEDNESS: Al viden skal opstå uden nogen former for økonomisk gevinst eller personlige fordele for forskere. Forskning skal kunne overvåges og testes for at mindske sandsynligheden for svindel, manipulation og autoritært misbrug.

ORGANIZED SKEPTICISM: al videnskabs validitet skal kunne bekræftes ved at dets metoder og resultater holdes op imod logiske og empiriske beviser.

(De Bellis, 2009)

Den normative citationsteori tager udgangspunkt i disse normer og bygger på, at videnskab er en institution grundlagt af interne belønninger og godkendelser, hvor forskere menes at veksle informationer i form af publikationer, til gengæld for anerkendelse i form af citationer. Citationer er altså, ifølge den normative teori, et redskab til at betale de værker man intellektuelt har benyttet tilbage i form af citationer. (MacRoberts&MacRoberts, 1987)

En anden videnskabsteoretisk tilgang til citationer er den socialkonstruktivistiske citationsteori, som bygger på, at forskere bruger citationer som overtalelsesredskab. Her er det ikke værdien eller indholdet af publikationen, der

tages højde for, men det citerede værks forfatter og dennes placering i det videnskabelige hierarki, der bruges til at overtale læseren om deres påstandes troværdighed.

MacRoberts og MacRoberts liste retter en del kritik mod en ensrettet opfattelse af citationer. De pointerer, at der er flere forhold i forbindelse med citationsanalyser der bevirker, at de ikke er troværdige.

Kritikken kan deles op i to dele. Den første del er rettet mod de teorier og antagelser, der ligger til grund for citationsanalyser. Den anden del koncentrerer sig om de data man anvender, samt deres funktion og kvalitet.

*Event-data problems of citation analysis:*

1. *Formal influences not cited*
2. *Biased citing*
3. *Informal influences not cited*
4. *Self-citing*
5. *Different types of citations*
6. *Variations in citation rate related to type of publication, nationality, time period, size and type of speciality.*
7. *Technical limitations of citation indices and bibliographies*
  - a. *Multiple authorship*
  - b. *Synonyms*
  - c. *Homonyms*
  - d. *Clerical errors*
  - e. *Coverage of literature.*

(MacRoberts&MacRoberts, 1987: 343)

Det første punkt omhandler hvorvidt referencelister er udtømmende. Undersøgelser af referencelister har givet MacRoberts og MacRoberts det resultat, at kun en forholdsvis lille procentdel af den forskning, der rent faktisk bliver brugt, også bliver nævnt i referencelisten.

Det næste punkt behandler den problematik, at citationer gives fordomsfuldt og derved skaber skævheder. Dette kan være fordomme der behandles endvidere af

(Camacho-Minano, M. & Nunez-Nickel, M. 2009), som vi kommer ind på senere. Her henvises der blandt andet til Matthæus-effekten (Merton 1968); til den der har, skal mere gives; dette uddybes senere i opgaven. Derudover kan der være faktorer som kollegial citation og smiger, der gør, at forskere citerer med bagtanker.

Punkt tre omhandler, ligesom punkt et, at forskning der har haft indflydelse, ikke bliver citeret, her menes der dog den viden som ikke er nedfældet, også kaldet 'tacitknowledge', dette indebærer samtaler og diskussioner om emnet med andre forskere, men også teknisk hjælp, der muliggør undersøgelserne. MacRoberts og MacRoberts pointerer dog, at det endnu ikke vides, hvor stor betydning denne uformelle viden har for citationsanalyser og om det overhovedet er muligt at gøre op i en referenceliste.

Punkt 4-6 omhandler problematikker ved de mange forskellige typer af citationer, inklusiv sevl-citationer, samt kulturelle og disciplinære forskelle i citationspraksis. I punkt syv med dertilhørende underpunkter tages der højde for tekniske problemer og begrænsninger, der kan opstå i forbindelse med citationsanalyser.

MacRoberts og MacRoberts(1989) afviser blankt den mertoniske opfattelse af den normative citationsteori og påstår, at citationsanalyser ikke er andet end ukorrekte gengivelser af det egentlige forhold mellem dokumenter og forfattere. De opfatter citationsanalyser som fuldstændig misvisende billeder af videnskabelige strukturer og forhold (MacRoberts&MacRoberts, 1987) (De Bellis, 2009: 261)

Deres undersøgelser er dog lavet på meget små mængder af data, og kræver videre efterforskning for fuldstændigt at kunne afvise citationsanalyser som metode til blandt andet bedømmelse af forskning. Deres opgør med den teoretiske baggrund bliver sat yderligere i perspektiv i forbindelse med forskeres citationsadfærd.



## *KARAKTERISTIK AF DE TRE STORE VIDENSKABSOMRÅDER*

I dette afsnit vil vi beskrive citationsmønstre og traditioner inden for naturvidenskaben, humaniora og samfundsvidenskab.

De tre forskellige felter, filosofi, fysik og sociologi, som vi har valgt at lave vores undersøgelse ud fra, repræsenterer de tre store videnskabsområder, nemlig henholdsvis humaniora, naturvidenskab og samfundsvidenskab. (Herefter kan forkortelserne HUM, NAT og SAM forefindes i opgaven.) Vi finder det derfor relevant at identificere forskellige reference- og citationsmønstre inden for de forskellige discipliner.

Hyland (1999) undersøger, hvordan forskellige akademiske citationspraksisser medvirker til konstruktionen af viden i form af fx videnskabelige artikler. Hans studie baserer sig på citationskontekstanalyse inden for otte videnskabelige discipliner. Det er på baggrund af denne artikel, at vi har udvalgt de tre felter, der er genstand for vores analyse.

Han identificerer store disciplinære forskelle, både i hvordan forfattere citerer, og hvordan de bruger referencerne i teksten – hvis de gør det. Inden for HUM og SAM bliver der brugt væsentligt flere referencer end i NAT, og i humaniora er de mere tilbøjelige til at integrere referencerne i teksten. Samtidig argumenterer han for, at forskellene bunder i epistemologiske og sociale konventioner inden for de forskellige felter.

Ligesom Swales (1990), beskriver han brugen af integrerede kontra non-integrerede referencer. Integrerede referencer er dem, hvor den citerede forfatters navn optræder som en del af en sætning (som den foregående sætning). Og non-integrerede referencer henviser til forfatteren, enten ved hjælp af en parentes (som denne: Hyland 1999: 344) eller en nummereret fodnote<sup>1</sup>.

Derudover nævner Hyland også, at man i HUM og SAM kunne finde referencer der i teksten blev beskrevet som utroværdige. Negative referencer blev altså ikke fundet inden for NAT.

---

<sup>1</sup> Som denne: Swales (1990) side 146

Björn Hellqvist (2010) har skrevet en artikel om 'Referencing in the Humanities', hvori han beskriver, hvad der karakteriserer den humanistiske referencetradition. Derved får han også ridset op, hvordan det forholder sig inden for SAM og NAT. For det første er publikationsformen inden for humaniora, præget af markant flere monografier og antologier – og langt flere udgivelser på lokalsprog. Udgivelser fra et felt som fysik vil primært være på engelsk, da det i mange sammenhænge regnes for at være videnskabens sprog. HUM og SAM har derimod et bredere publikum, og kan derfor også blive læst uden for det videnskabelige samfund.

*"Citations are used as rhetorical devices to signify the context of a statement, or to act as a symbol for a theory or method, but to a lesser extent as a tool for turning statements into facts."*

(Hellqvist 2010: 313)

Som citatet viser, så ser Hellqvist citationer som et retorisk greb, eller et 'concept symbol' (hvilket vi kommer mere ind på senere), og i mindre grad som et redskab til at omdanne påstande til fakta, som oftere vil være referencernes funktion inden for de naturvidenskabelige felter. For de humanistiske videnskaber, vil der typisk heller ikke være nogen distinktion mellem fiktion og non-fiktion – begge dele bruges på lige fod. Ikke mindst fordi det ofte handler om at fortolke og forholde sig til ældre tekster, især inden for filosofi.

Gennem videnskabens historie har originalitet været en afgørende faktor for videnskabelige gennembrud. Men igen er der forskelle mellem de forskellige videnskaber. I humaniora kan en ny tilgangsvinkel, en ny metode fra et andet felt, eller en ny teori være lige så original eller opsigtsvækkende, som nye empiriske opdagelser i naturvidenskaben. Derfor ser man også i humaniora en mere promiskuøs citationsadfærd, hvor sammenhængen mellem to tilsyneladende separate fænomener skaber originalitet. Det betyder, at der citeres bredt inden for meget forskellige felter.

Referencerne i humaniora (stadig ifølge Hellqvist 2010) er ofte integrerede i teksten og skaber selve konteksten omkring forfatteren og dennes tekst – lige så vel som de også kan anerkende tidligere forskning (og andre tekster). På grund af

den diskuterende og argumenterende facon i den humanistiske tradition, bliver referencerne også mere tvetydige og kontrastfyldte og ofte sat op imod hinanden, også som negative referencer.

Generelt er citationsanalyser oftest blevet foretaget inden for det naturvidenskabelige felt, da der her er en højere koncentration af udgivelser på artikelbasis, som derfor bedre er dækket i citationsindexerne.

Også Smith(1981) nævner problematikken omkring variationer inden for forskellige felter. Dette giver problemer når man ved brug af citationsanalyse skal finde et tværfagligt sammenligningsgrundlag imellem to eller flere felter. Hun henviser til Marcia J. Bates (1980, I:Smith, L., 1981), der pointerer, at der er forskel på, hvor stort antallet af forskere er inden for felterne, og dermed også forskel på hvor mange potentielle "citers" der er for et givent dokument.

De her fremsatte påstande om forskelle inden for de forskellige discipliner, er i øvrigt også nævnt i Larivière et al. (2006) og i Glänzel&Schoepflin (1998)

### *CITATIONSADFÆRD*

Det er vanskeligt at tale om citationsanalyser og måske endda en mulig citationsteori, uden også at beskæftige sig med citationsadfærd. Der er nemlig mange forskellige grunde til at forskere citerer hinanden – og sig selv. I dette afsnit vil vi se på mulige grunde til at citere, som de er opstillet af bl.a. Garfield (1965), Small (1978), Brooks (1986) og Case & Miller (2011)

#### GARFIELD (1965)

Mange faktorer kan være i spil, når en forsker -bevidst eller ubevidst- vælger at referere til et andet værk. Ifølge Garfield (1965) bruges citationer til at understøtte eller fremsætte dokumentation for specifikke udsagn. Han mener også at citationer og referencer er korte repræsentationer af de dokumenter, som de identificerer.

Garfields liste med 15 mulige grunde til at en forfatter citerer en anden, er velkendt inden for dette felt, men her følger den alligevel:

1. *Paying homage to pioneers*
2. *Giving credit for related work (homage to peers)*

3. *Identifying methodology, equipment, etc.*
4. *Providing background reading*
5. *Correcting one's own work*
6. *Correcting the work of others*
7. *Criticizing previous work*
8. *Substantiating claims*
9. *Alerting to forthcoming work*
10. *Providing leads to poorly, disseminated, poorly indexed or uncited work*
11. *Authenticating data and classes of fact - physical constants, etc.*
12. *Identifying original publication in which an idea or concept was discussed*
13. *Identifying original publication or other work describing an eponymic concept or term as, e.g., Hodgkins's Disease, Pareto's Law, Friedel-Crafts Reaction, etc.*
14. *Disclaiming work or ideas of others (negative claims)*
15. *Disputing priority claims of others (negative homage)*

(Garfield, 1965: 80)

Denne liste er blevet flittigt brugt og videreudviklet, hvilket også vil kunne ses senere i dette afsnit.

### SMALL(1978)

Small (1978) argumenter også, lige som Garfield, for at citationer kan ses som repræsentationer af dokumenter, men han har en anden indgangsvinkel, der ikke er modstridende med de ovennævnte grunde, men derimod en udvidelse, og en ny vinkel. Small introducerer nemlig begrebet "concept symbol" om en citation eller reference.

Det symbolske i en citation/reference kan ses ud fra to vinkler. Den første, og mest åbenlyse er at referencen (forfatter, årstal, titel, tidsskrift, nummer, sider) repræsenterer det fysiske dokument der citeres i sig selv. Den anden, og mere interessante grund ifølge Small, er at det citerede dokument er et symbol for det koncept, altså den idé, metode eller det grundlag som det citerede dokument bygger på. Denne sidste forbindelse er metaforisk, hvilket betyder at der i ekstreme tilfælde ikke behøver at være nogen som helst forbindelse mellem det citerede og citerende dokument. Men forholdet mellem det, som dokumentet

symboliserer og det rent fysiske dokument, kan sammenlignes med den første vinkel. Altså at det, som dokumentet symboliserer, repræsenterer det fysiske dokument, på samme måde som de bibliografiske oplysninger gør.

Selv en liste af værker (som senere af Hargens (2000) defineres som en ORL – Orienting Reference List), der opremses som en bibliografi i en artikel, indikerer, at disse andre værker har et fælles værdigrundlag/ metode/teori, der skaber en kontekst omkring den nærværende undersøgelse/artikel.

Højt citerede dokumenter kan få en status som 'standard symboler' inden for forskellige felter.

*"A theory of citation practice, if such a theory is possible, must take account of the symbolic act of authors associating particular ideas with particular documents. The analysis has shown that when a scientist cites, he or she is creating a link between a concept, procedure, or kind of data, and a document or documents. In some cases, the association of idea and document is well established by uniform practice within the community."*

(Small, 1978: side 337)

Ideen om citationer som symbol for det dokument de repræsenterer, er ikke modstridende med andre grunde til at citere, det er blot en udvidelse heraf. Konkluderende kan det siges at Small ser citationer som fortolkninger af det citerede værk.

Da Small er den første til at foreslå, at citationer kan fungere som 'concept symbols', er hans egen artikel med tiden blevet til et standard symbol, for citationer som repræsentationer for de ideer, koncepter, metoder, osv – som det fysiske dokument indeholder.

### BROOKS (1986)

20 år efter Garfield og knap 10 år efter Small, opsatte Terrence Brooks (1986) også en liste, men med kun syv punkter, som dog tydeligt er inspireret af Garfields liste.

I Brooks' undersøgelse blev 20 akademikere interviewet omkring deres grunde til at give i alt 437 referencer. Størstedelen af referencerne var ikke kun givet ud fra et enkelt punkt på listen, faktisk var hele 70,7 % af referencerne motiveret af en kombination af flere forskellige punkter.

Brooks' syv punkter er som følger:

1. *Currency (up-to-dateness)*
2. *Negative credit*
3. *Operational information*
4. *Persuasiveness*
5. *Positive credit*
6. *Reader alert*
7. *Social consensus*

(Brooks, 1986: 35)

Undersøgelsen viste, at mange af citationerne var en kombination af især punkt 4: '*Persuasiveness*' og derudover punkt 1, 5 og 7. Brooks deler da også de syv motiver fra listen op i tre grupper, hvor gruppe 1 består af punkt nummer 4, 5, 1 og 7, med 4 som det dominerende punkt. Forfattere bruger tilsyneladende et eller flere af disse motiver til at gøre opmærksom på den seneste, vigtig eller 'must-read' forskning. Negative citationer har hos Brooks fået sin egen gruppe, og han nævner at hans undersøgelse faktisk validerer MacRoberts&MacRoberts påstand om at negative citationer også spiller en rolle, da de i Brooks' undersøgelse ofte ledsages af en mere positiv reference. Gruppe 3 består af de resterende punkter 3 og 6, og betegnes som 'servicereferencer', hvor forfatteren gør opmærksom på baggrundslitteratur. Ifølge denne undersøgelse citerer man primært for at overbevise andre om vigtigheden af sit arbejde.

Brooks mener selv, at han er den første der undersøger de komplekse grunde til at citere empirisk og opfordrer til yderligere undersøgelser (s. 36), hvilket flere da også har taget til efterretning.

#### CASE & MILLER (2011)

Blandt andet en meget ny undersøgelse fra Case & Miller(2011), der undersøger, hvorvidt bibliometrikere citerer anderledes end andre akademikere, og

konkluderer at der findes 20 grunde til at citere, og at disse 20 kan inddeles i syv overordnede faktorer.

1. *Negative citationer – kritik af andres arbejde; uoverensstemmelser eller lignende.*
2. *Personlig indflydelse - det citerede dokument har stor betydning for forfatterens valg af metode eller tænkemåde.*
3. *Sociale grunde – citerer en mulig anmelder; legitimering af ens eget arbejde*
4. *Klassiske citationer – konceptuelle, fundamentalreferencer*
5. *Kreative – der fx inspirerer til videre forskning, eller skaber bro mellem to forskellige forskningsfelter*
6. *Demonstrerende – man viser kendskab til feltet (hænger til en vis grad sammen med de sociale grunde); citerer en fremstående forsker eller et indflydelsesrigt tidsskrift.*
7. *Review – omformulerer ældre undersøgelser; gør opmærksom på andre undersøgelser*

Case & Miller (2011): 427-428

Listen er en tydelig videreførelse af både Garfield og Brooks' lister.

Mange af de adspurgte indikerer, at den enkelte reference er en 'concept marker', altså et symbol for et koncept, der dermed kan repræsentere en hel genre, en række af velkendte teorier, metoder og opdagelser, hvilket også Small (1978) argumenterer for.

White (2001) har udført en undersøgelse, hvis formål var at afdække tendensen til at forskere reciterer sig selv og andre over tid. Der skelnes mellem en individuel forskers citationsimage og citationsidentitet. Et citationsimage indebærer hvilke andre forskere denne er co-citeret med og afspejler hvordan forskningsfeltet opfatter dem, samt hvilke andre forskere de sidestilles med. Dem man selv citerer i sine egne publikationer over tid skaber derimod en forskers citationsidentitet. Forskere vil typisk recitere sig selv og andre over tid, et mønster, der måske/måske ikke, er bevidst. Recitationsanalyser er oftest set inden for humaniora og samfundsvidenskaberne, da naturvidenskabelige værker ofte er

lavet af forskningsgrupper og derfor har mange forfattere, hvorved der ikke kan redegøres for, hvem der er ansvarlige for hvilke referencer.

Han redegør, i forhold til citationsadfærd, for at citationer altid opstår i kontekst, hvor han opstiller 4 typer citationer. Disse adskilles ved at være positive dybdegående eller overfladiske vs. negative dybdegående eller overfladiske. (for uddybning se White, HD. 2001) Denne kontekst bør tages i betragtning, når man vurderer forskeres citationsimage og -identitet.

Som afslutning på dette afsnit om citationsadfærd, kan det konkluderes at på trods af, at der ikke findes et regelsæt for, hvornår man skal citere, så vidner både Garfield, Brooks og Case & Millers lister om, at der alligevel findes en vis konsensus omkring, hvornår og hvorfor man citerer. Whites undersøgelse vidner derimod om, at det er meget individuelt, og at enhver forsker vil have sit eget citationsimage og sin egen citationsidentitet.

### *HVORNÅR SKAL MAN CITERE?*

De ovenstående afsnit er blot et udpluk af undersøgelser der har beskæftiget sig med, *hvordan* forskere citerer hinanden – og sig selv. Artiklen "When to Cite" af Garfield (1996) beskæftiger sig derimod med, *hvornår* man bør citere. Et område som stort set ikke er blevet undersøgt, hvilket Garfield finder meget bemærkelsesværdigt, i forhold til, hvor meget citationsadfærd ellers er blevet diskuteret inden for videnskabssociologien.

Flere steder nævner Garfield (1975, 1996) en opgave, som han gentagende gange har bedt sine nye studerende om at udføre. En artikel helt uden referencer uddeles og de studerende skal så angive, hvor i dokumentet de mener, at der bør være en reference. Gennem tiden har resultatet varieret fra 15 til 75, medens det antal referencer Garfield selv havde planlagt var omkring 45. Dette eksempel viser tydeligt, at der er stor forskel på, hvornår den enkelte mener, at der er behov for en reference. Men det interessante i denne forbindelse er, igen ifølge Garfield selv (1975), at de fleste studerende er endt med at foreslå et antal omkring netop 45.



Dette kan blandt andet skyldes, at der er stor forskel på, hvad der af den enkelte forsker anses for almen viden, og derfor ikke behøver at blive citeret. Erfarne forskere vil ofte prioritere originalitet og kreativitet, hvorimod mere uerfarne forskere vil benytte sig af oversigtsartikler og sekundære referencer (Seglen,1996).

I den forbindelse pointerer Garfield (1985), at det er vigtigt for den enkelte forsker at finde den rette balance af referencer, der også varierer væsentligt, afhængigt af, inden for hvilken fagdisciplin man befinder sig. Han gør opmærksom på, at der i den forbindelse er mange der udgiver overproportionerede referencelister og ligeså mange udgives mangelfulde.

Ifølge Garfield (1996) har også Matthæus-effekten en indflydelse på citationsadfærd. Metodologiske dokumenter forefindes ofte på lister over højt citerede dokumenter. Dette kan i nogle tilfælde ske på bekostning af mindre kendte værker, der måske i virkeligheden er mere relevante for den pågældende forskning. Dette resulterer i en skævhed i forhold til, hvem der krediteres for forskningsresultater. Der er, ifølge Merton, en tendens til, at man i et dokument lægger mærke til en i forvejen kendt forsker, men ignorerer mindre kendte medforfattere. (Merton, 1968).

Smith (1981) har fremsat fem formodninger der understøtter brugen af citationer og referencer.

1. God citationsadfærd fordrer at forfatteren henviser til alle dokumenter, vedkommende har brugt.
2. Der er sammenhæng mellem antallet af citationer og dokumentets kvalitet.
3. Det er de bedst mulige værker der citeres.
4. Et citeret dokument er emnemæssigt relateret til det citerende dokument.
5. Alle citationer vægter lige meget. Med undtagelse af selvcitationer.

I forhold til punkt tre, spiller tilgængelighed en stor rolle, hvilket også nævnes af Camacho-Minano, M. & Nunez-Nickel, M. (2009).

Der findes altså ingen universel guide til, hvornår man skal citere. Muligvis er citationsidentiteten noget der grundlægges allerede i løbet af ens skoletid, eller tillæres ud fra den citationsadfærd, der afkodes og opleves gennem læsning af videnskabelige tekster.

### *CITATIONSKONTEKST*

I dette afsnit vil vi se på andre forhold der spiller ind på de referencer der gives i videnskabelige artikler. De første afsnit omhandler den videnskabelige artikel som genre, dernæst vil vi beskrive hvordan referencer kan optræde på tekstniveau. De sidste afsnit drejer sig om referencernes alder og hvor mange gange de forekommer i teksten.

### HVOR & HVORFOR? - DEN VIDENSKABELIGE ARTIKEL SOM GENRE

I nogle tidsskrifter er der tradition for, at artiklerne er opdelt efter formatet IMRaD, der står for Introduction, Method, Results and Discussion. De artikler vi har brugt i vores empiri, er dog ikke opstillet ud fra de nævnte kriterier, men må formodes at have de samme indholdselementer alligevel. Swales (1990) analyserer disse forskellige dele af en videnskabelig artikel på tekstniveau, og hans bog "*Genre Analysis – english in academic and research settings*" danner baggrund for det følgende.

INTRODUKTIONER er kendt for at volde forfattere/forskere store problemer (Swales 1990). Selvom det er det første, der møder læseren (hvis der da ikke er et abstract), så skrives introduktionen oftest til sidst. Der skal her foretages nogle valg omkring, hvad der er vigtigt at fremhæve, hvilken baggrundsviden skal læseren have inden den fortsatte læsning, og hvor direkte skal indgangsvinklen være. Swales redegør for en analyse udført af Zappen ([1983], i: Swales 1990), der viser at introduktioner oftest indeholder: et mål (hvad man vil opnå med artiklen); nuværende kapacitet (det bedste man kan gøre på det pågældende tidspunkt); et problem; en løsning (løsningsforslag i den nærværende artikel) og evalueringskriterier (hvordan kan det bruges).

Swales selv har lavet en model for videnskabelige artiklers introduktioner, som han kalder for CARS-modellen (Creating A Research Space). Den består af tre

forskellige 'moves' og herunder forskellige 'steps', som altså handler om at skabe et undersøgelsesområde ved hjælp af disse forskellige sproglige handlinger. Modellen er udledt af mange års erfaring inden for akademisk skrivning, men kan også bruges som et redskab til at skrive introduktioner. Selve modellen er vedlagt som bilag nr. 9, men går i korte træk ud på at etablere et territorium, hvor man kan forsikre læseren om områdets væsentlighed og/eller lave generaliseringer inden for området og/eller anmelde tidligere undersøgelser. Dernæst gælder det om at etablere en niche ved enten A) at lave modpåstande, eller B) indikere et 'gap' (altså et 'hul' i den eksisterende forskning), eller C) stille spørgsmål, eller endelig D) for at fortsætte en forskningstradition. I det sidste 'move' indtager man sin niche ved at anføre sit formål eller fremlægge nuværende forskning og meddele om principielle opdagelser og endelig kan man vælge at forklare artiklens struktur.

METODEAFSNITTET i en videnskabelig artikel vil der (alt efter fagområde) være en mere eller mindre uddybning af, hvilken metode der er brugt i den pågældende undersøgelse. Der vil ofte være en hel del implicite faktorer, fx hvis en metode bare er nævnt ved navn og ikke har en specifik reference, så vil det oftest betyde, at det er almenviden inden for området og derfor ikke behøver en eksplicit forklaring. Dette er et problem, fx hvis man er ny inden for feltet, eller opererer inden for et andet diskursfællesskab, da man så ikke uden videre vil kunne finde den pågældende metode i et opslagsværk. Ofte vil metodeafsnittet også tage læserens tavse viden for givet, og bruge fagudtryk i sine sætninger, så det for en 'uindviet' vil lyde som det rene volapyk. Her kan også nævnes punkt 3 på MacRoberts&MacRoberts liste, at man 'glemmer' at citere 'informalinfluences'. Som med alt andet, når det drejer sig om videnskab, så er der selvfølgelig store forskelle inden for forskellige discipliner og inden for fx humaniora kan metoden have en helt anden form og berettigelse.

RESULTATER, DISKUSSION & KONKLUSION. Hvor introduktion og metode er helt faste elementer i en videnskabelig artikel, uanset fagområde, så er Resultater, Diskussion & Konklusion ikke nødvendigvis en fast bestanddel. Selvom IMRAD-modellen er en god retningslinje for, hvordan en artikel er opbygget, så kan disse tre delelementer være sat sammen på kryds og tværs og have forskellige andre synonymer som overskrift. Og der kan være stor forskel på, hvad der bliver

beskrevet i de forskellige afsnit. Swales begiver sig ud i nogle forsigtige hypoteser, da området her ikke er særlig godt undersøgt. Men han skriver, at et resultatafsnit, kan spænde lige fra at være 20 siders tekst, til at rumme en række grafer eller tabeller, ledsaget af en opremsende tekst.

Swales skriver afslutningsvis i kapitlet, at den viden der findes om den videnskabelige artikel som genre, slet ikke er komplet. Bestemte karakteristika kan siges at gå igen inden for de forskellige sektioner, men der er stadig meget at undersøge. Men baseret på den forskning der er udført indtil videre, som i mindre grad er gennemgået i det foregående, finder han det overraskende, at de store tekstlige forskelle ikke ligger i introduktion og diskussionsafsnittene, men derimod i metoden og resultaterne, som kan have meget forskellig udformning.

Der kan tilmed være en sammenhæng mellem hvor i artiklens afsnit referencerne forekommer og hvilken forbindelse til det citerede dokument de derved repræsenterer. Det er på baggrund af dette, at hypotese 1 udspringer.

#### HVORDAN? – PRÆSENTATION AF REFERENCER

Charles Bazerman(2000) argumenterer, for at referencer der ikke er blevet oblittereret i forskningsfeltet, ofte optræder i artiklens introduktion med den funktion at 'etablere litteraturen'. Dette vil sige at tilbyde læseren et overblik over al den litteratur på området, der både kan give et historisk tilbageblik, samt gøre opmærksom på vigtige pointer inden for emnet - ifølge forfatteren vel at mærke. Disse bliver samlet og nævnt i artiklens begyndelse, men sjældent behandlet særlig dybt.

Lowell L. Hargens(2000) introducerer begrebet 'Orienting Reference List' (herefter betegnet som ORL), der er en lignende liste, som indeholder referencer, der skal skabe en ramme om den pågældende forskning og samtidig retfærdiggøre den videnskabelige position. Denne ORL defineres som en referenceliste bestående af minimum 3 forskellige dokumenter, der af den citerende forfatter refereres for at eksemplificere en tilgang, tema eller metode og ikke specifikke resultater. Disse ORLs viste sig at være mest almindelige inden for de adfærdsmæssige videnskaber, der også havde tendens til at overcitere fundamentale værker inden for feltet. De ORLs der forekom inden for de naturvidenskabelige genre, havde oftere tendens til

at bestå af nyere referencer end i de andre områder. Samtidig kom det frem, at mange af de fundamentale referencer i ORLs var nævnt som generaliserede og standardiserede referencer inden for det pågældende emne og altså blev citeret for deres 'mainknowledgeclaim' eller som 'concept symbol'. Forskellen er at disse ORLs som udgangspunkt kan forekomme alle steder i en videnskabelig artikel. Brugt optimalt ses de som et godt redskab for forfatteren, blandt andet til at sætte læseren ind i en kontekst, der skaber bedre forståelse for forfatterens synspunkt.

Ken Hyland (1999) ser ligesom Swales (1990), citationer som et stærkt retorisk redskab i videnskabelige artikler. Han mener at disse retoriske redskaber bliver brugt på forskellig måde inden for forskellige fagdiscipliner. Dette begrundes med andre måder, hvorpå forskere konstruerer andre dele af den videnskabelige artikel, i kontekst af det videnskabelige bagland og deres traditioner og kommunikationsmønstre. Han pointerer, at til trods for disse indforståede regler der eksisterer i både forskningsfelter og tidsskrifter, er det samtidig det individuelle valg, der kan afsløre og signalere forskellige meninger.

Hyland (1999) undersøger i denne forbindelse ikke hvor, men hvordan referencerne placeres og i hvilken sammenhæng. I forhold til det empiriske grundlag blev allereferencer i første omgang talt og allerede dér, viste der sig store forskelle i det gennemsnitlige antal af anvendte referencer. Sociologi havde flest, filosofi havde over gennemsnittet og fysik det færreste antal referencer. Generelt viste alle discipliner, på nær filosofi, en præference for non-integrerede referencer. Her tages også forbehold for, at naturvidenskabelige tidsskrifter generelt har et krav om nummererede referencer. Hyland (1999) argumenterer for, at de integrerede referencer bringer mere fokus på forfatteren og kan bruges som en form for dialog i argumentering og analyse, hvor de non-integrerede tager mere afstand og afpersonaliserer det anvendte materiale. Desuden fandtes der klar tendens i, hvilke verber de forskellige discipliner anvendte til at introducere deres referencer. Inden for de bløde videnskaber, var det primært kommunikerende verber så som; siger, påstår, synes, diskuterer, forklarer. Hvor de mere hårde videnskaber brugte flere neutrale verber, såsom: viser, rapporterer, udvikler, beskriver. Han lægger mærke til en tendens specielt inden for filosofi, hvor forskere oftere blotter sig akademisk og forholder sig yderst subjektivt og

fortolkende til de anvendte referencer. Hvorimod forskerne inden for fysik mere kamouflerer deres egne fortolkninger i forskningsprocessen.

### HVORNÅR? – HVAD BETYDER REFERENCERNES ALDER?

En undersøgelse som både beskæftiger sig med citationsadfærd, men også med referencernes alder, er Shadish, Tolliver, Gray & Gupta (1995). Denne undersøgelse, som er foretaget i begyndelsen af 90'erne på feltet psykologi, bygger på en kvalitativ spørgeskemaundersøgelse, hvor de citerende forskere bliver bedt om at svare på, hvorfor de har citeret den pågældende artikel/forfatter.

Shadish et al. udleder to vigtige pointer om højt citerede værker, nemlig at de af andre forskere bliver set som betydningsfulde, og at de er af tilsyneladende høj kvalitet. (s. 485)

De finder også en tendens til at citationer til 'gamle' værker (der kommer ikke en nærmere definition), ofte gives fordi de enten er:

- Eksemplariske, forbilledlige (altså et muligt 'concept symbol')
- De bliver opfattet som forældede (og bliver citeret fordi forfatteren bygger videre på noget, eller kommer med en forbedring af fx en metode)
- De har været/ er genstand for kritik (og den citerende artikel retter op på denne kritik, eller fortsætter kritikken)
- De er højt citerede (og der derfor kan være en forventning om, at dette værk 'skal' citeres inden for det pågældende felt)

Det er mindre sandsynligt, at hvis en reference er gammel, at den så er central for nye forskningsområder (hot topics). Tydeligvis er ikke alle gamle citationer højt citerede eller forbilledlige – nogle er bare gamle.

Når ældre værker er højt citerede, kan det alene være på grund af deres rolle som forbilleder (exemplars) der skal anerkendes, selvom de menes at være metodisk eller teoretisk forældede, og ikke af høj kvalitet efter moderne standarder. Omvendt er højt citerede nyere værker, ofte af høj kvalitet og genstand for både ros og kritik fra kollegaer. (s. 488)

Mange citationer er ikke udelukkende lig med kvalitet, af to grunde:

1. Citationer måler andet end kvalitet. (fx forbilledstatus)

2. Ikke alt kvalitetsforskning er højt citeret. (s 492)

Det er lige værd at nævne, at negative citationer ikke i denne undersøgelse bliver fundet i en udstrækning, der kan argumentere for, at det skulle give højt citerede artikler.

Lowell Hargens (2000) har tidligere undersøgt forskellige aspekter ved forskeres brug af litteratur, inden for forskellige fagdiscipliner. Her undersøges om der er forskel på, hvilken litteratur der gøres brug af, og der skelnes mellem litteratur fra forskningsfronten og fundamental litteratur inden for feltet. I undersøgelsen bliver der fremlagt to former for prototypiske strukturer.

Den første, hvor forskere har tendens til at citere nyere publicerede dokumenter og ignorere det ældre og fundamentale. Dette vil ofte være inden for fagdiscipliner, hvor fortidens viden er så inkorporeret i nyere forskning, at man ikke finder det nødvendigt at nævne de fundamentale kilder. Og forskere inden for disse felter har fokus på at bygge videre på den nyeste forskning. I denne sammenhæng kan nævnes Garfieldsobliterations fænomen (Garfield, 1975), der først bliver behandlet af Merton (1965) i form af den velkendte frase "*Standing on the shoulders of giants*". Fænomenet skal betragtes som forklaringen på, hvad der sker når metoder, begreber eller ideer bliver institutionaliseret. Jo mere almindelig kendt noget bliver, jo mindre vil det blive citeret. (Som eksempel bruges Arkimedes' konstant, bedre kendt som  $\pi$ ). Dette opfatter Garfield som den ultimative anerkendelse en forsker kan opnå.

I den anden struktur har forskere omvendt tendens til at fravælge det nyere forskning og fokusere på det tidlige og fundamentale forskning. Denne struktur vil ofte eksistere inden for felter, hvor der sættes perspektiver på spørgsmål formuleret i klassiske tekster, eller hvor der fortolkes på ældre værker. Desuden kan der typisk være tale om en kendt metode, der udforskes i nye sammenhænge (Hargens 2000).

Undersøgelsens resultat viser, at de naturvidenskabelige fagdiscipliner har tendens til at falde ind under den første kategori, hvor de adfærdsmæssige

videnskaber afspejler den anden struktur. Den humanistiske disciplin kan hverken siges at være det ene eller det andet. Det konkluderes, at fagområder der har tendens til at overcitere fundamentale dokumenter, gør det for at overbevise læseren, det vil sige det eksisterende forskningsmiljø, om vigtigheden og rigtigheden af deres egen forskning. Hvorimod de der har tendens til at overcitere dokumenter fra forskningsfronten for specifikke detaljer eller metoder, derved skaber en indforstået enighed om, hvad der er relevant på området. Hargens (2000) skaber disse prototyper og diskuterer resultaterne med udgangspunkt i andre tendenser og strukturer der eksisterer inden for forskellige fagdiscipliner, hvilket vi vender tilbage til senere.

Andre forskere der har beskæftiget sig med at undersøge alderen på anvendte referencer er Oppenheim og Renn (1978). I deres artikel argumenteres for en lang række årsager til, at det oftest er de nyere dokumenter der bliver citeret. Obliterations fænomenet nævnes her igen, og derfor finder de det relevant at undersøge, hvad årsagen til at visse ældre højt citerede dokumenter bliver ved med at modtage citationer.. Undersøgelsen viser at omkring 40 % af referencerne til de ældre dokumenter udelukkende var givet for 'historisk baggrund', hvilket inkluderer 'givinghomage to pioneers' og det at angive relevant baggrundslæsning. Susan E. Cozzens (1985) gennemgår i sin sammenligning af videnskaber, hvordan brugen af en reference kan ændres over tid. Her skelnes mellem at referere til et værks 'mainknowledgeclaim', 'metode' eller 'peripheralknowledgeclaims'. Hendes undersøgelse gik altså ud på at finde ud af, hvorvidt årsagen til de mange citationer ændrede sig over tid, og blev foretaget på en naturvidenskabelig tekst og en tekst fra sociologien. Det viste sig, at der var klare ændringer i citationerne til den naturvidenskabelige artikel. I den første årrække blev den citeret meget for detaljer og metoder i forbindelse med artiklen, og senere hen blev den næsten udelukkende citeret for hovedideen eller opdagelsen bag artiklen.

Det er dette fænomen som Merton og Zuckerman (ifølge Cozzens, 1985) referer til som kodifikation: "*konsolideringen af empirisk viden til kortfattet og indbyrdes afhængige formuleringer*" (s. 128: vores oversættelse). Altså overgangen fra nye empiriske opdagelser til en formulering af en generel ide i takt med tiden, og



dermed også en nuance af citationer som 'concept symbols' som Small (1978) betegner det.

Den anden tekst som Cozzens undersøgte, viste derimod ikke den store ændring i konteksten af citationer over tid, men dette skyldes - ifølge forfatterne - en manglende opmærksomhed omkring artiklens egentlige hovedpointe (Cozzens, 1985). Der gøres opmærksom på, at kodifikation ikke er et begreb, der direkte kan bruges i forbindelse med citationskontekstsanalyse, og at det altid vil være afhængigt af, hvorledes de citerende forskere vælger at tage imod den viden, der præsenteres i det citerede dokument.

#### HVOR MANGE? – OM ANTAL FOREKOMSTER

Det har længe været et spørgsmål til diskussion og undersøgelse om citationer i sekundære ressourcer som eksempelvis SCI har den samme værdi eller om nogle citationer er mere 'værd' end andre. I et paper fra 1976 undersøges det empirisk af Henry Voos og Katherine S. Dagaev (Voos&Dagaev, 1976) om der er, eller burde være, forskel på citationers værdi, afhængigt af hvor i den videnskabelige artikel den befinder sig, og hvorvidt der henvises til det samme værk fra referencelisten flere gange i løbet af artiklen. Deres formål er at forsøge at gøre brugen af citationer mere præcis og anvendelig. Deres undersøgelse bygger dog kun på optællinger af citationer i fire værker ud af de 30 mest citerede fra 1970-1973. På daværende tidspunkt var de tilgængelige citationsindekser alle i papirform, og citationsanalyser blev udført ved manuel optælling. To antagelser var (og er stadig) gældende, nemlig at alle citationer har den samme værdi ved deres første optræden i dokumentet og flere henvisninger til samme reference i samme dokument, ikke tæller mere end én citation.

Da videnskabelige artikler jo består af flere dele, finder de det relevant at undersøge hvor i artiklen referencen optræder og siger også, at det for brugeren kan være værdifuldt at vide, om en reference er blevet brugt flere gange i det samme dokument. De har i modsætning til vores undersøgelse taget udgangspunkt i alle de artikler, der har citeret deres fire kerneartikler, hvor vi har

taget udgangspunkt i 45 kerneartikler og undersøgt deres måde at citere på. Her finder de altså det komplette antal 'op-cits' i de citerende artikler.

Deres opdagelser viser, som de ovenstående undersøgelser også kommer ind på, at der inden for de forskellige fagdiscipliner er stor forskel på, hvor referencerne forekommer. De aner dog en tendens til at deres kerneartikler er mest citeret i introduktioner, og det er samtidig her, de har flere op-cits end i nogle af de andre steder i artiklerne. De mener ikke, at der hersker nogen tvivl om, at en forfatter eller artikel der er citeret flere gange i et dokument, har større betydning end den forfatter eller artikel som kun er citeret end enkelt gang.

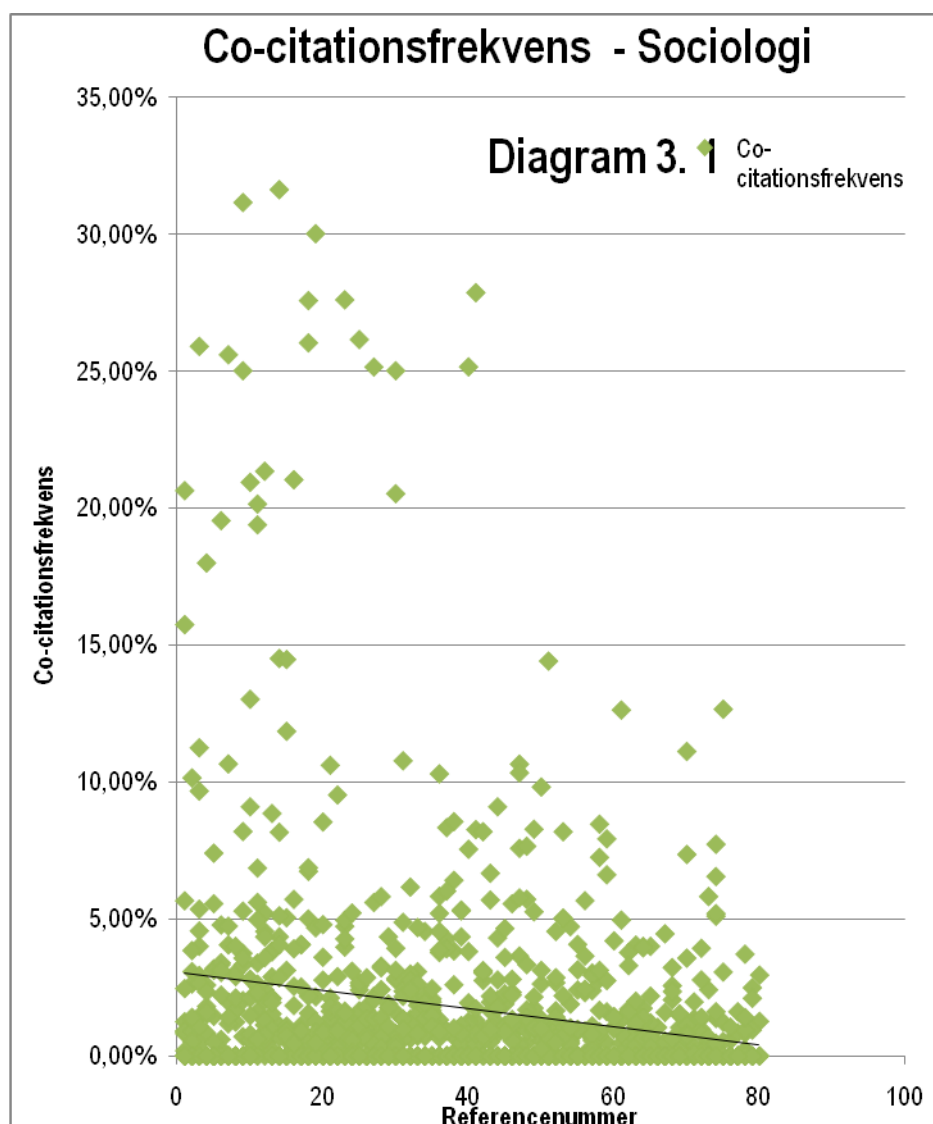
Ud fra det ovennævnte konkluderer de, at der er tydelige indikationer af, at det er muligt at beregne en citations værdi, ved at undersøge i hvilket afsnit den befinder sig og om den forekommer flere gange.

...

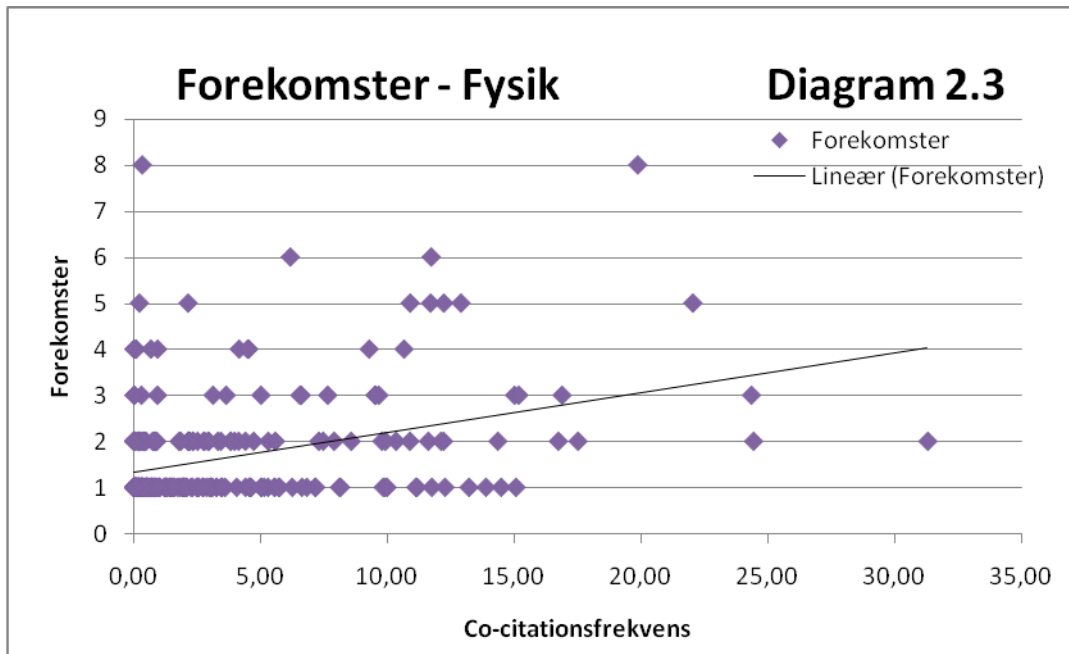
## PRÆSENTATION AF RESULTATER

Følgende afsnit indeholder en kort gennemgang af de resultater vi er kommet frem til via co-citationsanalysen samt citationskontekstanalysen. Vi har opdelt data fra hver fagdisciplin i præsentationsark og grafer, for at kunne spore eventuelle tendenser og samtidig identificere eventuelle feltforskelle. Alle diagrammerne kan yderligere ses i bilag 1.

I diagram 3.1 ses et eksempel på resultatet af co-citationsanalysen, med referencenumrene på x-aksen og co-citationsfrekvensen på y-aksen. Ved at anvende co-citationsfrekvensen, tager vi også højde for antallet af referencernes individuelle citationer og dette gør resultaterne mere sammenlignelige. Diagrammet præsenterer hvordan disse fordeler sig i forhold til artiklens opbygning.



I diagrammerne nedenfor kan ses en oversigt over antallet af forekomster(y-aksen) for referencerne i forhold til den procentmæssige co-citationsfrekvens (x-aksen), hvilket viser hvor mange gange referencerne optræder i forhold til hvor højt co-citerede de er.



Samme antal forekomster(y-aksen) ses i diagram 2.2, men i forhold til hvor i artiklen de optræder, på y-aksen hvor referencenumrene er placeret. Det er altså muligt at se hvordan de referencer, der har flere forekomster er placeret i artiklens opbygning.

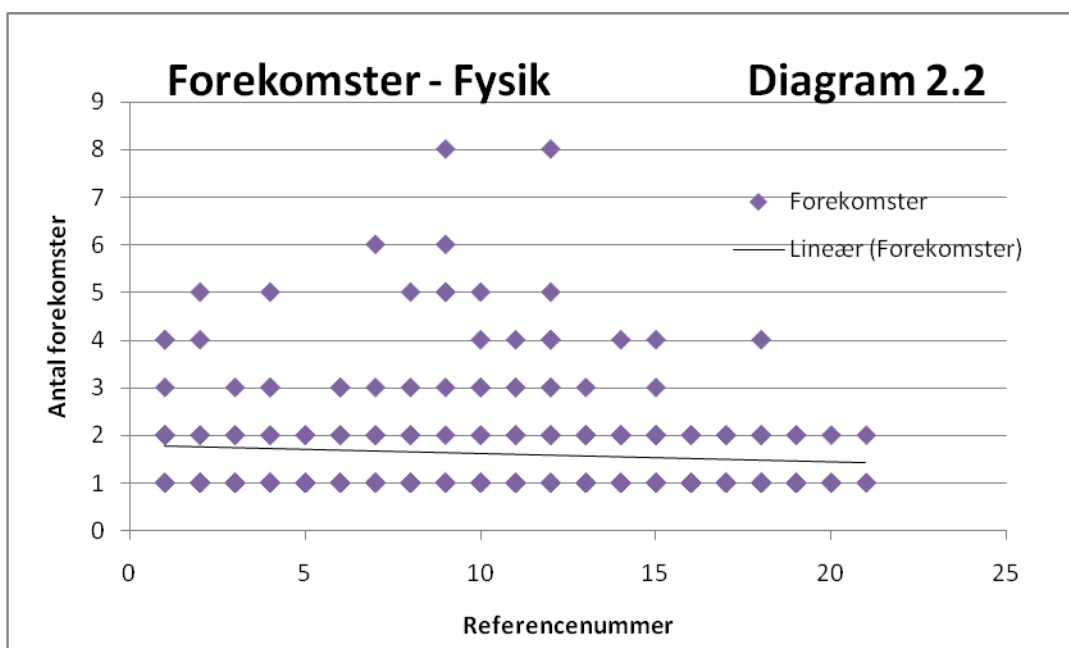
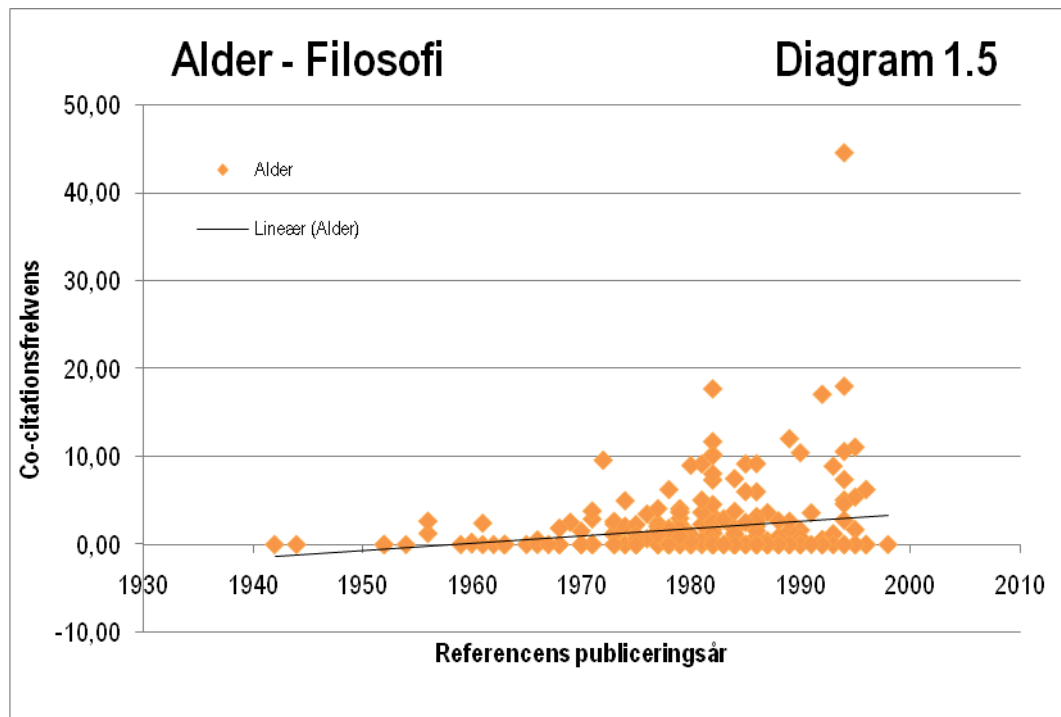


Diagram 1.5 viser referencernes udgivelsesår (y-aksen) holdt op imod co-citationsfrekvensen, der er præsenteret på x-aksen. Dette viser hvornår de mest og mindst co-citerede referencer er fra og hvordan de fordeler sig.



Tabellen på næste side, viser et udsnit af vores samlede data fra feltet filosofi. Første kolonne er artiklens referencenummer, der ses i listen vedlagt som bilag 4, her er det artikel 11 fra filosofi. Den næste kolonne henviser til det referencenummer referencen har fået i artiklen, altså den rækkefølge hvori de bliver nævnt første gang. Kolonnen 'citationer' viser, hvor citationer den enkelte reference har fået og i kolonnen 'Co-citationer' ses hvor mange co-citationer netop den reference har med kerneartiklen. Kolonnen Co-citationsfrekvensen er udregnet efter formelen præsenteret af Garfield (1980 i: Osareh, F., 1996) og viser den procentmæssige fordeling. Den gennemsnitlige co-citationsfrekvens for artiklerne inden for filosofi var 1,83 %, 3,22 % for fysik artiklerne og 1,74 % for artiklerne inden for sociologi. Alle tal over gennemsnittet er i denne kolonne markeret med farven grøn.

Vi ønskede yderligere at finde frem hvor mange gange de enkelte referencer blev anvendt i hver artikel og ved simpel tællemetode har vi noteret det i kolonnen 'Antal forekomster'. Dette gjorde vi for at kunne observere, hvorvidt de referencer

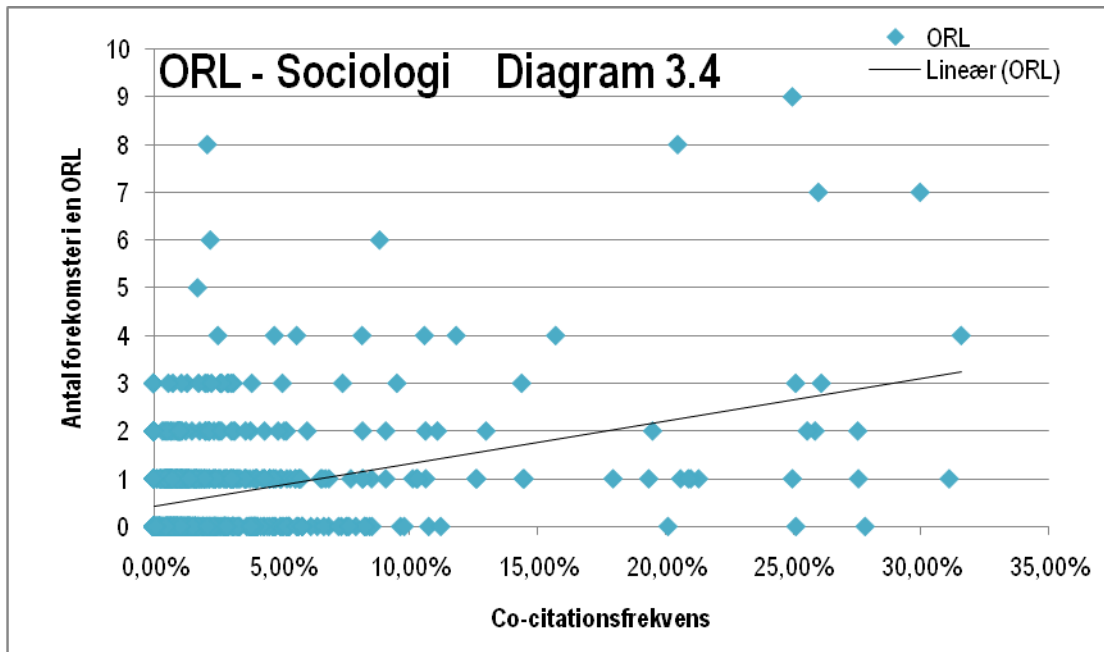
med flere forekomster var de samme som de med højere co-citationsfrekvens. Her er dem der forekom mere end en gang markeret med en lys lilla og dem der optrådte mere end 3 gange er markeret med en mørkere lilla.

	Reference nr.	Citationer	Co-citationer	Co-citationsfrekvens	Forekomster	ORL	Alder
11 - Filosofi	1	581	22	3,47	1	0	1976
	2	115	13	7,34	1	0	1982
	3	38	17	17,71	2	0	1982
	4	1	0	0,00	4	0	1984
	5	88	4	2,52	2	0	1982
	6	1	0	0,00	2	0	1985
	7	6	0	0,00	1	0	1973
	8	1	0	0,00	2	0	1987
	9	1	0	0,00	2	0	1980
	10	1	0	0,00	1	0	1963
	11	1	0	0,00	3	0	1984
	12	1	0	0,00	1	0	1971
	13	1	0	0,00	2	0	1984
	14	1	0	0,00	1	0	1979
	15	1	0	0,00	1	0	1971
	16	7	1	1,23	1	0	1982
	17	1	0	0,00	1	0	1980
	18	1	0	0,00	1	0	1984
	19	1	0	0,00	1	0	1984

Da vi ikke som Lowell Hargens (2000) havde mulighed for at gennemføre en omfattende referencenetværksanalyse og identificere forskningsfronten og de fundamentale artikler inden for hvert felt, sorterede vi referencernes alder ved at identificere de referencer der er udgivet over 10 år før den pågældende kerneartikel og definere dem som 'gamle'. Referencernes udgivelsesår er placeret i sidste kolonne 'Alder' og de 'gamle' i forhold til hver artikel er markeret med farven orange.

Dernæst gennemsøgte vi artiklerne og identificerede 'orienting reference lists' (Herefter ORL) (Hargens, LL. 2000) Vi ønskede at finde ud af om der var tendens til at artiklerne indeholdt ORL's, samt hvis de gjorde, om disse havde nogen forbindelse med de højt eller lavt co-citerede referencer. Vi definerede en reference liste som tidligere nævnt efter Lowell Hargens (2000). Disse resultater er præsenteret i kolonnen ORL og alle referencer der optrådte i en, er markeret med en lys blå og dem der optrådte i mere end én ORL i samme artikel er markeret med en mørkere blå. For at se om der var sammenhæng mellem en references

placering i en ORL og dennes co-citationsfrekvens med kerneartiklen, opstillede vi diagrammer der viste co-citationsfrekvensen som værdier på x-aksen og antallet af gange denne optræder i ORL på y-aksen. Her ses et eksempel fra sociologi.



Alle graferne i resultatet er blevet udstyret med en tendenslinje for at gøre resultaterne mere tydelige.

## *KARAKTERISTIK AF KERNEDOKUMENTERNE*

### FILOSOFI

Det var generelt svært at finde artikler inden for filosofi med samme antal referencer og stadig en forholdsvis høj mængde af citationer, hvilket resulterede i, at mange af kerneartiklernes referencer ikke har modtaget særligt mange citationer og derved ikke har mulighed for at have mange co-citationer med kerneartiklen. De 15 udvalgte artikler havde mellem 16 – 19 referencer med et sideantal på gennemsnitligt 23, 6 sider.

Forfatterne gør primært brug af integrerede referencer, men nogle anvender også enkelte non-integrerede i form af fodnoter. En enkelt artikel gjorde udelukkende brug af non-integrerede referencer.

Der blev identificeret ORL's i 9 ud af de 15 artikler. Som oftest forekom en enkelt ORL i de artikler, der indeholdte dem. Alle på nær 4 referencer, i de alt 15 artikler, kun blev nævnt i én ORL, hvis de overhovedet blev nævnt.

Desuden bliver referencerne gennemsnitlig nævnt 1,6 gange i løbet af artiklerne, hvilket også varierer en del fra artikel til artikel.

For at få den nødvendige mængde data vi havde brug for, for at kunne gennemføre en co-citationsanalyse var vi nødsaget til at udvide tidsperioden for de anvendte artikler fra filosofi. Artiklerne kommer fra tidsperioden 1985 – 1998 og er derved både ældre og inden for et bredere tidsrum end de andre to fag.

### FYSIK

Inden for feltet fysik skulle vi langt ned i søgeresultatet før det tyndede ud i antallet af citationer artiklerne havde modtaget, så her var det ikke noget problem at finde 15 højt citerede artikler, med samme antal referencer. De 15 udvalgte artikler har mellem 19-21 referencer og er på gennemsnitligt 5,4 sider.

I disse artikler gøres der primært brug af non-integrerede referencer, i form af en nummereret referenceliste. Meget få gange bliver forskerens navn integreret i teksten i forbindelse med en reference.

10 af de 15 artikler indeholdt ORL's, og deraf indeholdte størstedelen af de 10 mere end en ORL og forekom oftest i forbindelse med et emne eller et begreb. Referencerne bliver ligesom dem i filosofi nævnt gennemsnitlig 1,6 gange i løbet af artiklerne, og her er endnu større forskel artiklerne imellem.

Fysik artiklerne spænder over en tidsperiode fra 1995-2000, hvor 5 er fra 1998 og 5 af dem er fra 1996. Resten er fordelt på de øvrige år.

### SOCIOLOGI

Artiklerne fra sociologi var generelt længere og havde mange flere referencer, derfor var vi nødsaget til at vælge et højere antal referencer, nemlig mellem 76 og 79. Antallet af citationer som de udvalgte artikler fra dette felt har modtaget, er slet ikke lige så højt som i fysik, og derfor har nogle af artiklerne modtaget så få som 62 citationer. Derudover har de et sideantal på mellem 17 og 62 sider, hvilket også er en del mere end både fysik og filosofi.



Resultaterne ser da også anderledes ud for dette felt, da der er mange flere datapunkter i graferne, og der generelt var flere forekomster af hver enkelt reference.

I de sociologiartikler der har udgjort en tredjedel af vores empiri, kan det siges, at forfatterne i højere grad end i de andre felter, benytter sig af ORLs, og også gerne nævner de samme referencer flere gange.

Det er i øvrigt meget kendetegnende for de fleste af artiklerne, at de benytter sig af integrerede referencer, der bruges aktivt i teksten (undtaget ORLs, der bare listes op). Fodnoter er dog også hyppigt forekommende, men oftest med uddybende kommentarer og knap så hyppigt med tilføjelse af referencer.

## ANALYSE OG DISKUSSION

Følgende er en gennemgang er resultaterne analyseret og diskuteret i forhold til den foregående og gennemgåede teori.

Det første der falder i øjnene når man kigger ned over resultaterne er de store forskelle der er på de videnskabelige felter, men også fra artikel til artikel. Nogle af artiklerne har mange co-citationer med deres referencer og andre har stort set ingen, dette er specielt markant inden for filosofi (se bilag 4) Overordnet kan referencerne opdeles i fire grupper i forhold til antal citationer og co-citationer.

1. Den første gruppe indeholder de referencer, der selv har modtaget mange citationer, men kun har få eller ingen co-citationer med kerneartiklen.
2. Den anden gruppe indeholder de referencer som selv er højt citerede, og samtidig er ofte co-citerede med kerneartiklen.
3. Gruppen her repræsenterer de referencer, der ikke har modtaget særlig mange citationer og er lavt eller slet ikke co-citeret med artiklen.
4. Sidste grupper er de lavt citerede referencer, men hvoraf en betydelig del af disse alligevel er co-citeret med kerneartiklen.

	Lav eller ingen co-citationsfrekvens	Høj co-citationsfrekvens
Højt citerede artikler	<b>1</b>	<b>2</b>
Mindre citerede artikler	<b>3</b>	<b>4</b>

Gruppe 3 og 4 har som følge af deres mængde af citationer et dårligt udgangspunkt i forhold til at opnå co-citationer med artiklen. Det er alligevel lykkedes gruppe 4 at opnå en vis andel co-citationer i forhold til begrænsningerne, og disse kan derfor sagtens være høj-relevante for dokumentet. De lave citationer kan skyldes flere ting; enten at dokumentet er druknet i mængden af den producerede forskning, eller at kendskabet til den er utilstrækkeligt, og at den derfor ikke har modtaget anerkendelse i form af citationer. Derudover kan det citerede dokument være inddraget som følge af en negativ citering, eksempelvis som genstand for kritik. Sidst kan den forskning der bliver citeret, efterfølgende være blevet modbevist

eller voldsomt kritiseret, hvilket bidrager til det lave antal af citationer. Det kan også bare skyldes, at der henvises til materiale som endnu ikke er publiceret, og derfor ikke er tilgængeligt for andre at citere. Disse årsager kan være de samme i forhold til de lave citationer i punkt 3, dog antages det at referencerne i punkt 4 har større lighed med artiklen end punkt 3 som følge af grundideen bag citationsanalyser. Det kan eksempelvis være fordi referencerne i punkt 4 repræsenterer dokumenter fra et meget specialiseret felt, hvorfor det dog stadig kan være høj-relevant for kerneartiklen.

Referencerne i gruppe 1 har modtaget anerkendelse i forskningsfeltet i form af citationer, men deler ikke mange co-citationer med artiklen. Som udgangspunkt kan dette være et resultat af emnemæssig variation og andre forskere har derfor ikke belæg for at citere begge dokumenter i et tredje. Som eksempel på dette, kan forfatteren til artiklen have brugt referencen for at henvise til en metode, som normalt ikke figurerer i forbindelse med det emne artiklen behandler. Metoden kan derimod være en respekteret og meget anvendt metode inden for et andet forskningsfelt. Dette kan især være forekommende inden for de humanistiske videnskaber, hvor man ofte vil kunne finde eksempler på, at en metode eller teori fra et andet felt, bruges til at skabe nye sammenhænge eller nye vinkler (Hellqvist 2010). Hvis referencen i gruppe 1 stammer fra en oversigtsartikel, vil dette også kunne forklare de mange citationer og den lave andel af co-citationer. Desuden kan der i gruppe 1 være tale om tvivlsom citationsadfærd, blandt andet i form af smiger eller kollegial citering. De kan også være nævnt som historisk baggrund for et perifært område (Oppenheim & Renn, 1978).

Hvis man tager Susan E Cozzens (1985) undersøgelse i betragtning, i forbindelse med at forskeres årsager til at citere en reference, kan ændres over tid, kan der være tale om, at en reference af denne slags kan være blevet standardiseret inden for feltet og derved er blevet medtaget for dens overordnede idé og ikke har den store relevans for den citerende artikels egentlige emne/undersøgelsesområde. Standardisering af referencer kan i den forbindelse ses som det modsatte fænomen af obliteration (Garfield, 1975), som også kan sidestilles med henholdsvis Matthæus-effekten og punkt 1 på event-data listen over problemer ved citationsanalyse (MacRoberts & MacRoberts, 1989)

Gruppe 2 er dem vi antager for at være høj-relevante for artiklen i særlig grad, det kan enten være på et emnemæssigt plan eller i forbindelse med anvendte metoder. Men de kan også være inddragede for eksempelvis at dække et helt videnskabeligt felt (De Bellis 2009) eller fungere som 'concept symbol' (Small 1978). I den forbindelse kan gruppe 2 ligeledes indeholde standardiserede referencer.

Fordelen ved co-citationsanalysen i vores tilfælde er, at den tager udgangspunkt i referencer, der allerede er direkte forbundet til artiklen via en citation. Derved får vi ikke blot et indtryk af, hvilke dokumenter forfatteren selv vælger at koble sig med, men også at andre forskere har fundet begge artikler relevante i forhold til deres egen undersøgelse. Som følge af grundideen bag co-citationsanalyser, ville vi hurtigt kunne konkludere, at den ovennævnte gruppe 2 og gruppe 4, indeholder de referencer der burde tildeles en merværdi, men det er dog ikke så simpelt, da der i forbindelse med citationsadfærd og citationsanalyser eksisterer så mange usikkerheder og forbehold der skal tages højde for.

Inden for feltet filosofi tilhører en forholdsvis stor mængde af referencerne gruppe 3 og har i denne forbindelse mange referencer med en co-citationsfrekvens på 0,00 % og flere artikler indeholder kun en eller meget få referencer de selv er co-citerede med. Til trods for dette, er den gennemsnitlige co-citationsfrekvens højere end artiklerne inden for sociologi. Dette skyldes først og fremmest den langt lavere mængde af referencer og citationer, men også den deraf store mængde af referencer der tilhører gruppe 4, som gør at nogle af artiklerne har mange referencer over gennemsnittet. Artiklerne fra sociologi indeholder i lighed med filosofi en stor mængde referencer fra gruppe 4, denne er dog ikke ligeså markant, da der generelt bliver citeret mere inden for sociologi end filosofi, hvilket også afspejles i længden af deres referencelister. Af samme årsag har sociologi til forskel fra filosofi, en stor del referencer der tilhører gruppe 2. Desuden er der forholdsvis mange referencer både fra filosofi og sociologi, der befinder sig i gruppe 1. Dette kan skyldes at forskningsfelter som filosofi og sociologi ofte har tendens til at citere uden for deres eget felt, deres citationsradius er generelt stor, hvilket også kan forklare nogle af referencerne i gruppe 1 (Hyland K. 1999, Hellqvist B, 2010). Artiklerne fra fysik indeholder primært referencer fra gruppe 1 og 2, fordi referencerne generelt har modtaget langt flere citationer. Samtidig skyldes det

høje gennemsnit i co-citationsfrekvensen også det høje antal citationer, der er givet kerneartiklerne, da der skal flere citationer til for at opnå en høj co-citationsfrekvens. Det ses eksempelvis i reference 2 i artikel 1 (se bilag 6), den er ud af sine 89 citationer blevet co-citeret 10 gange med kerneartiklen, men i forhold til hvor mange citationer kerneartiklen har (se artikel 1, bilag 5), er dette ikke af større betydning. Afslutningsvis indeholder både artiklerne fra fysik og sociologi også referencer der tilhører gruppe 3, disse er der dog ekstremt få af inden for fysik.

Observationerne af referencerne viser meget tydelige forskelle i citationsadfærd, specielt er der afvigelser fra de mere 'bløde' felter sociologi og filosofi og til den 'hårde' fysik. Dette kan blandt andet skyldes, at især databasen Arts & Humanities Citations Index ikke har en særlig god dækningsgrad, når det kommer til monografier og antologier. Derfor vil co-citationsfrekvensen være meget lav, hvis referencen er givet til fx en bog, der typisk ikke vil være indexeret i en citationsdatabase.

### FILOSOFI

Selvom en stor del af referencerne har en co-citationsfrekvens på 0,00 % kan der stadig ses en tydelig tendens til, at de referencer der har en høj co-citationsfrekvens i forhold til kerneartiklen ligger blandt de første referencenumre og at de derfra er faldende. Der er dog en reference der skiller sig særligt ud, med en co-citationsfrekvens på 44,54%, hvilket er usædvanligt, specielt inden for dette felt, hvor der ikke bliver citeret nær så meget som i andre. Denne reference vil have mulighed for at gå ind og påvirke resultatet af grafen, men ikke desto mindre er tendensen stadig tydelig, som det ses på diagram 1.1 i bilag 1.

I forbindelse med de filosofiske artikler, blev referencerne ofte nævnt flere gange i løbet af artiklen, uden at der forekom en formel reference, selvom der ofte blev gjort brug af de to udtryk 'ibid.' og 'op.cit', hvilke her er betragtet som formelle referencer. Dette er der ikke taget højde for i optællingen af forekomster, men bør stadig nævnes, da det indikerer hvordan flere referencer aktivt bliver brugt og nævnt flere gange end resultaterne afspejler. Samtidig siger det noget om traditioner og vaner når der skrives videnskabelige artikler inden for forskellige

fagdiscipliner. Til trods for dette viser graferne en lille tendens til at de referencer der blev nævnt flere gange også havde for vane at befinde sig blandt de første i artiklerne. (se diagram 1.2 – bilag 1)

Desuden ses samtidig en sammenhæng mellem antallet af forekomster og en stigning i co-citationsfrekvensen, dette er igen hjulpet på vej af den førnævnte reference, der er nævnt 4 gange i den pågældende artikel. (se diagram 1.3 – bilag 1)

Gennemsnitligt forekom referencerne i en ORL 0,3 gange i alle 15 artikler. Dette er altså ikke en ofte set tendens i filosofi, hvilket også medfører at den repræsenterende graf er meget svag og ikke giver et særligt brugbart resultat (diagram 1.4 – bilag 1). Det kan dog nævnes at selvom ORL's fordelte sig vel over artiklernes referencenumre var der et flertal der befandt sig i blandt de første. Dette stemmer godt overens med teorien om en orienterende liste (Hargens, 2000) eller med formålet om at bruge disse til at etablere litteraturen på området i artiklens begyndelse (Bazerman, 2000).

I forhold til referencernes alder ses der i diagram 1.5 (bilag 1) en tendens til at jo nyere referencerne er, jo oftere var de højt co-citerede med artiklen. Det skal her siges at filosofi var den fagdisciplin der indeholdt flest ugyldige referencer i forhold til deres alder og de dermed er blanke. Eksempelvis havde flere en reference til Immanuel Kant ("Kant, I") uden hverken publikationsår eller kilde på. Dette kan tolkes som en standardiseret reference, hvor det refereres til skaberen af en generel idé, som muligvis er på vej til at være oblitereret. Hvilket yderligere begrundes ved, at ingen af disse ugyldige referencer var højt co-citerede med deres artikel, til trods for at flere af dem, selv havde modtaget en stor mængde citationer.

Desuden er der en lille tendens til at jo nyere artiklen er, jo ældre er de referencer der anvendes set i forhold til artiklens udgivelsesår.

### FYSIK

I den første graf (diagram 2.1 – bilag 1) ses en meget tydelig tendens til at de højt co-citerede referencer er tilbøjelige til at befinde sig blandt nogle af de første i artiklen. Denne tendens er dog faldende jo længere hen i referencenumrene man bevæger sig. De høje co-citationer er altså, trods den klare tendens, mere bredt

over referencenumrene. Den gennemsnitlige co-citationsfrekvens var her langt højere end i de andre felter oppe på 3,22 %, hvilket tydeligvis har sammenhæng med, at der generelt bliver citeret mere inden for fysikken som fagdisciplin, end der gør inden for de andre felter. Jo oftere dokumenter bliver citeret, jo større er chancen logisk for at blive co-citerede, dette betyder dog ikke noget for referencernes relevans for artiklen i forhold til de andre felter.

Diagram 2.2 viser en svag tendens til, at de referencer der forekommer mere end en gang, optræder første gang i artiklens begyndelse. Her skal dog påpeges, at antallet af forekomster bør ses i forhold til de få antal sider der generelt prægede artiklerne i fysik. Der ses desuden en meget tydelig tendens til, at de referencer der bliver nævnt flere gange, også har højere co-citationsfrekvens med kerneartiklen. Dette kan tolkes som at forskere inden for fysik, har tendens til at bruge de højt-citerede referencer mere, eftersom de bliver nævnt flere gange på meget få sider. Dette er en god indikator for relevans i forhold til vores tese om co-citationsanalysens resultater.

Der er, ligesom inden for filosofi, ikke særlig stor tilbøjelighed til at anvende ORL's i fysikkens artikler. 0,4 referencer forekommer i ORL's i alle 15 artikler. Men en forskel er, at der hvor der forekommer ORL's har de til gengæld tendens til at indeholde langt flere end 3 referencer. En af artiklerne indeholder en ORL på 11 referencer i artiklens introduktion. Et tilfælde som dette kan dog ikke ses i grafen, da referencerne er tildelt numre selv om de nævnes samme sted. Dette medfører at selv om reference nummer 11 i dette tilfælde nævnes allerede 3 linjer inde i artiklen, vil denne stadig syne som at forekomme i midten af artiklen, grundet det lave antal af referencer og artiklens sideantal. På denne måde kan ORLs gå ind og påvirke resultaterne i diagrammerne. Hvis vi havde tildelt alle referencer i en ORL det samme referencenummer, ville det have påvirket resultatet i en anden retning. Da ville der nemlig være langt flere referencer i begyndelsen af dokumentet og vores data fra de forskellige artikler havde ikke været sammenlignelige.

Referencerne i fysik var generelt hurtigere at identificere end i de to andre felter fordi de i stort set alle artikler er nummererede og fremkommer først i referencelisten.

Aldersmæssigt fordeler referencerne sig over en periode på næsten 100 år med den ældste fra 1901 og op til 1999. Det er dog meget få referencer, der er udgivet før 1950 og langt de fleste stammer fra 1990'erne, hvor artiklerne også er udgivet. Dette forekommer naturligt, eftersom fysik er et ældgammelt, men stadig hurtigt voksende felt, hvor der bliver publiceret rigtigt meget litteratur. Der ses også en tendens i diagram 2.5 (bilag 1) til at jo nyere referencerne er, jo oftere er de co-citerede med artiklen. Dette er specielt bemærkelsesværdigt da dokumenter logisk har mulighed for at opnå flere citationer og dermed også co-citationer jo længere de har eksisteret. Men ofte hænger det ikke således sammen, fordi emner ofte bliver behandlet i en periode hvor de er 'hot topics' indtil der, specielt inden for naturvidenskabelige felter, findes en løsning eller en afslutning på debatten. Hvorimod samfundsvidenskabelig og humaniora artikler kan analysere og fortolke i årtier på de samme tekster/metoder/idéer osv.

### SOCIOLOGI

Graferne som vi har lavet ud fra vores undersøgelse viser nogle forskellige tendenser.

Diagram 3.1, der viser co-citationsfrekvensen ud fra referencenumrene, indikerer, at der er en tendens til, at de referencer med den højeste frekvens befinder sig i begyndelsen af dokumentet. Dette giver meget god mening, også når man ser i sammenhæng med forfatterens brug af ORLs. De referencer der forekommer flere gange i en ORL er nemlig også dem, der er mest co-citerede med kerneartiklen (se diagram 3.4 – bilag 1).

Da forfatterne/forskerne inden for dette felt tilsyneladende gør stor brug af ORLs, især i begyndelsen af deres artikler, så er det naturligt at de referencer der har flest forekomster placerer sig dér. Vi har tildelt referencerne numre efter hvornår de optræder første gang i artiklen, og tendenslinjen i diagram 3.2 viser at forfatterne også rent faktisk *bruger* de referencer som de nævner i begyndelsen, ved at citere dem flere gange (=mange forekomster).

Tendensen er også tydelig, når man ser på de to grafer der viser forekomsterne. På den første kan man se, at antallet af forekomster er størst ved de laveste



referencenumre. Og på den anden kan man se at co-citationsfrekvensen stiger, jo flere forekomster referencen har (diagram 3.2 og 3.3 – bilag 1).

Den sidste graf viser en tendens til, at de artikler som kerneartiklen er mest co-citeret med, også er de nyeste. Man er altså ikke så ofte co-citeret med sine gamle referencer, måske fordi de blot er med til at klarlægge et felt eller fungerer som eksemplariske forbilleder.

Aldersmæssigt er den ældste reference i sociologi fra 1859, den forekommer i artikel 15 og er på trods af dens alder tilsyneladende ikke særlig højt citeret, hvilket indikerer at den *ikke* er en standardiseret reference. Den er nok bare gammel (Shadish et al., 1995).

Overordnet virker resultaterne fra sociologi mere markante, men det er sådan set klart nok, eftersom der er flere datapunkter i diagrammerne, grundet det større antal af referencer.

#### FEJLMARGIN

Det kan være visse usikkerheder ved at bedømme en references placering, udelukkende ved at opstille dem i den rækkefølge de optræder i. For det første er der stor forskel på sideantallet hos keredokumenterne, specielt fysik skiller sig ud, hvor en god mængde referencer bliver fordelt over meget få sider. Hvorimod i de sociologiske og filosofiske artikler breder referencerne sig over mange sider, og der er således ingen grænser for, hvor langt inde i artiklen den første reference kan forekomme eller hvornår den sidste gør. Dette aspekt er især usikkert inden for filosofi, hvor relativt få referencer, er fordelt på relativt mange sider. Her er det mere plausibelt inden for sociologi, der også har et højt sideantal, men derimod også langt flere referencer der skal fordeles.

I og med at artiklerne er valgt ud fra et kriterium om samme antal referencer, så kan datasættet siges at være repræsentativt, for både højt citerede og mindre højt citerede artikler. Denne gennemgang af resultaterne viser, at selvom der analyseres på tværs af forskningsfelterne, er der en gennemgående tendens til at de relevante referencer optræder første gang i artiklernes første referencer. På

baggrund af vores viden om den videnskabelige artikel som genre, havde vi nemlig en formodning om at det ville forholde sig sådan.

### *HVAD CITERES HVOR?*

Følgende vil uddybe årsagen til, at denne tese er opstået, med udgangspunkt i den teoretiske gennemgang af Swales' teorier, med perspektiv på betydningen af ORLs (Hargens LL. 2000)

Hvorfor der i introduktionen skulle forekomme flere co-citationsreferencer, begrundes vi med, at der her fremhæves alle de vigtige elementer som følger i artiklen. Det inkluderer reference af forskellig art, nemlig til:

- tidligere forskning (sin egen eller andres) som man bygger videre på
- ny og aktuel forskning, der beskæftiger sig med noget inden for samme emne
- oversigtsartikler, der giver et bredere billede af området
- andre artikler som har undersøgt, noget lignende det, som man selv undersøger
- metoder, som man vil anvende i sin undersøgelse
- mangler i tidligere artikler, som man vil rette op på (egne eller andres)

Der er hermed grundlag for en stor mængde af referencer i begyndelsen af artiklen, hvor emne og problem præsenteres.

Desuden vil citationer brugt retorisk som overtalelsesredskab (også nævnt som 'persuasiveness' ifølge Brooks 1986), også sandsynligvis fremkomme i introduktionen, for at inddrage læseren og forklare hvorfor netop denne artikel er relevant.

Disse referencer i introduktionen vil sandsynligvis senere hen blive nævnt i sammenhæng med denne kerneartikel og vil derved opnå en høj co-citationsfrekvens. Ikke fordi de nødvendigvis er emnemæssigt beslægtede, men fordi de sammen er med til at klarlægge et felt.

I forhold til Garfields (1965) 15 punkter (G) og Case & Millers (2011) 7 faktorer (CM), der opstilles som grunde til at citere, stemmer nogle af punkterne direkte overens med de ovenstående. Blandt andet punkt 4 på både Garfields og Case & Millers lister, der handler om at henvise til baggrundsstof og punkt 12 (G)

som henviser direkte til tidligere forskning. Desuden er punkt 2 (CM) overensstemmende med flere af de ovenstående referencer, som typisk vil blive givet i introduktionen. Yderligere vil de 'moves' som er beskrevet (og vedlagt som bilag 9) også på forskellig måde kræve en reference.

I et metodeafsnit vil de referencer der gives, oftest være til en eller flere benyttede metoder, som det nævnes i punkt 3 på Garfields liste, og co-citationerne vil ikke nødvendigvis indikere en emnemæssig relation, da den samme metode godt ville kunne benyttes på flere discipliner. Det vil dog også kunne forekomme her, at referencer undlades, da de måske opfattes som almen viden af forfatteren jævnfør Garfields punkt 13, Ikke at forglemme MacRoberts&MacRoberts, der kritiserer at 'formal influences' ikke bliver citeret, hvilket de mener, har en stor betydning for citationsanalysers troværdighed.

Det gælder for videnskabelige metoder, at de skal være beskrevet og dokumenteret således at man vil kunne efterprøve dem. I praksis vil det dog i de fleste tilfælde være umuligt, da der ofte ligger mange tanker og overvejelser forud for selve nedskrivningen af metoden, og fravalg foretaget undervejs i processen bliver ikke altid beskrevet. Disse fravalg mister derved citationer, som de kunne have fået, hvis de var blevet nævnt. Vi mener ikke, at dette nødvendigvis er negativt, for ved at medtage denne form for negative citationer, ville der skabes støj i søgningerne, og der ville være endnu flere usikkerheder forbundet med citationsanalyser.

Det er på baggrund af de overvejelser, der er nævnt i dette afsnit, at vi fremsatte vores hypotese. Vi afviser dog ikke, at referencer som først optræder i et andet afsnit også kan være kan være relevante. Men skulle man vurdere citationers relevans efter deres placering, vil vi mene at det er hensigtsmæssigt at vurdere de mest co-citerede, først forekommende citationer højest. Det at tendensen viser sig at være gældende for alle de tre fagdiscipliner, gør blot denne formodning mere troværdig.

At sammenstille referencernes placering med deres antal af forekomster og sige noget om deres relevans, bringer os til undersøgelsen af Voos og Dagaev (1976), hvor de konkluderer, at denne tendens er så klar, at der kan tildeles en værdi til de

citationer der nævnes flere gange. Da der på nuværende tidspunkt kun tildeles én citation til en anvendt reference, mener vi at det forekommer nyttesløst at nævne samme reference mange gange, hvis ikke det var fordi den var relevant. I forbindelse med citationsadfærd kan årsager som 'persuasivenss' spille en rolle i denne sammenhæng, ved flere gange at gøre læseren opmærksom på forbindelsen til samme reference, men til trods for dette, finder vi alligevel resultatet troværdigt. Sociologi og Fysik præsenterer meget overbevisende resultater i forbindelse med, at de mest co-citerede referencer forekommer flere gange i løbet af artiklerne. Dette kan specielt diskuteres i forhold til sociologi, da rigtig mange af referencer bliver nævnt mere end en gang. Vi er dog skeptiske overfor en eventuel værdibasing på baggrund af antallet af forekomster, da flere af de meget højt co-citerede referencer i filosofi og fysik kun havde en enkelt forekomst.

### *BETYDNINGEN AF ORL*

Inspireret af Hargens (2000), har vi også i empirien undersøgt, hvilke referencer der optrådte i en ORL. Med hvilket formål en ORL anvendes, er generelt op til den enkelte forfatter, og med udgangspunkt i den anvendte teori, er dette både positivt og negativt. Den kan i et positivt lys betragtes som en hjælp for både læseren og forfatteren, som et redskab til læseren om relevant supplerende læsning og som et redskab for forfatteren til at tydeliggøre perspektivet, hvori den pågældende forskning bør ansues af læseren.(Bazerman, 2000);(Oppenheim & Renn, 1978) Denne ORL kan dog 'misbruges' og anvendes som en form for 'namedropping' i forsøg på at overtale læseren om vigtigheden af artiklen og opnå anerkendelse af forskningen.(Brooks, 1985) Dette mener vi dog ikke nødvendigvis bør betragtes som noget negativt. Med den mængde af forskning der i dag bliver produceret og med de mange videnskabelige perspektiver der tilbydes, kan det virke forståeligt, at forskning der i forvejen er anerkendt, bruges til at vise læseren og ikke mindst eventuelle peers, hvorfor ens forskning er relevant i lyset af referencerne i disse ORL's.

Hargens (2000) argumenterer for, at ORL's ofte vil bestå af fundamentale referencer eller liste 'hot topics' At dømme på alderen af de referencer der forekom, fandt vi ikke en sammenhæng i den retning. Men vi mener, at en ORL, der ikke præsenterer standardiserede referencer, kan siges at afspejle en vis distance

fra forfatter til reference. Dette begrundes vi med fremstillingsformen, der ikke viser det store engagement til hver enkelt, men blot er en opstillet liste.

At det lige netop er sociologien, der i langt højere grad end i de andre felter gør brug af ORLs er ikke overraskende. Betragter man det i forhold til den brede målgruppe, der kan anvende sociologiske artikler, er det nødvendigt for forskerne at tage hensyn til, at ikke alle er lige bekendte med feltets tidligere forskning. Der ses ingen sammenhæng mellem alderen på referencerne og deres forekomster i en ORL. Det kunne tænkes, i forbindelse med name-dropping, at en del fundamentale tekster kunne forekomme i ORL for at retfærdiggøre den som overtalelsesredskab, dette ser dog ikke ud til at være tilfældet. Dette bør selvfølgelig undersøges yderligere før en eventuel konklusion og kan ikke vurderes udelukkende på referencernes alder.

Det præsenteres yderligere i graferne, hvorvidt der er sammenhæng mellem referencernes co-citationsfrekvens med kerneartiklen og hvorvidt disse forekommer i en ORL. Der er i den forbindelse en ikke nogen tendens til at de referencer der forekommer i ORL's også er højt co-citerede med artiklen inden for fysik, da tendenslinjen holder sig under 0. Der er dog en forholdsvis klar tendens i de sociologiske artikler, der på sin vis retfærdiggør disse ORL's som generelt relevante for artiklerne og ikke en overflødig liste af referencer. Filosofi viser også en stigende tendens op til 1, denne er dog primært hjulpet på vej ved at den ene reference der har en yderst høj co-citationsfrekvens også er nævnt i en ORL. Det skal dog siges at disse grafer er noget usikre med meget varierende og lave mængder af ORL's i filosofi og delvist fysik.

### *VALG AF REFERENCER*

I forhold til Ken Hylands (1999) undersøgelse viste vores artikler, i overensstemmelse med hans resultater, at der inden for filosofi er en præference for integrerede referencer og fysik ligeledes for ikke-integrerede. Der var dog forskel i vores og hans undersøgelse, da kerneartiklerne fra sociologi primært anvendte integrerede referencer. Men som han konkluderede, ser vi også en klar tendens inden for de filosofiske tekster til at forfatterne forholder sig meget subjektivt til referencerne og teksten foregår nærmest i en dialog med lange citater fra de citerede artikler, samt både negativ og positiv respons på disse fra

forfatterens side. Den analyserende og fortolkende tone går igen i alle de filosofiske artikler, og viser en tydelig uformel tone. Til forskel tager artiklerne fra fysik i langt større grad afstand til deres referencer og forholder sig neutralt til dem. Der er en gennemgående forklarende og konstaterende tone, som afspejler den jargon, der eksisterer i denne form for artikler. Artiklerne fra sociologi kan siges at være en god blanding, da de både forholder sig fortolkende til nogle af deres kilder og nøgternt til andre. Dette, mener vi, viser meget tydeligt, hvor store forskelle der er på felterne i henhold til forskningstraditioner, og samtidig hvordan forskere socialt interagerer med hinanden og deres kilder.

Desuden er det karakteristisk, at de filosofiske artikler refererer til flere værker af samme forfatter, se bilag 3 for eksempel, hvor der refereres til otte forskellige værker af samme forfatter. Dette kan have noget at gøre med det høje niveau af personligt engagement forskerne afspejler til hinanden, samtidig kan det vidne om en dybdegående analyse af netop denne forskers filosofiske synspunkter. Desuden har flere af disse referencer ikke sidetal eller publikationsnavn hvilket tyder på at flere af disse er repræsentationer for monografier. Til forskel fra eksemplerne fra sociologi og fysik der primært består af artikler, hvilket er i overensstemmelse tidligere teorier (Hellqvist B, 2010; Hylland K, 1999)

Vores undersøgelse viser hvilke referencer, der i højere grad er en del af de enkelte forskeres citationsimage. I forhold til netop denne artikel (bilag 3) kan man antage, at 'Quine, W' er en forsker, som forfatteren af artiklen ser som en del af sin citationsidentitet, men i forhold til co-citationsanalysen er det udelukkende hans egne referencer, der er en del af hans citationsimage. Der kan dog ikke konkluderes noget om de enkelte forskeres citationsidentitet uden, at vi betragter dem over en tidsperiode.

Inden for fysik forekommer der sjældent mange referencer til samme forfatter og ser man på den foregående analyse, vil et sådan tilfælde sandsynligvis komme af at forfatteren anvender metoder eller resultater fra foregående arbejder til sin egen forskning eller sammenligner dem. Forskerne i sociologi har det også med at nævne flere værker af samme forfatter, dog ikke i samme grad som filosofi. Denne tendens forekommer da også logisk, da forskere normalt gennem tiden vil bygge videre på deres egen forskning og derved må citere tidligere værker for at skabe

den rette kontekst. Dette er ikke umiddelbart noget, der ofte optræder inden for fysik, de har derimod en større tendens til at citere artikler produceret inden for samme tidsskrift, som det de selv er publiceret i. Dette kan ses i lyset af et perspektiv af White (2001) om kollegial citering og sociale netværk, men kan også forklares ved et langt smallere emne og igen det mere specificerede publikum.

Når man taler om at vurdere en references relevans for det citerende dokument, kan der sjældent opstilles enslydende kriterier. Der er for hvert tilfælde mange pragmatiske, hermeneutiske og paradigmemæssige aspekter, der bør tages i betragtning.(Hjørland 2010)

Foskett(1972: Ifølge Hjørland 2010) adskiller begrebet *relevans*, der bruges for et givent paradigme og konsensus inden for et felt, fra begrebet *pertinent*, der ses i forhold til den individuelle brugers behov. En reference kan altså være relevant, både for feltets sammenfattede tankegang, men også i forhold til den enkelte brugers opfattelse. En reference, der er relevant for én læser, kan være irrelevant for en anden og omvendt. Dette bør iagttages med fokus på, hvordan et eventuelt publikum læser den citerende artikel. Forfatterens valg må således afspejle de referencer, denne finder relevant for målet/målene med artiklen. Disse kan være pluralistiske og relateres til de forskellige begreber inden for citationsadfærd og kan ikke defineres bredt.

Co-citationsanalysen afspejler, hvilke referencer andre har anset som værende relevant i sammenhænge, der så skal ses i kontekst af deres individuelle mål med et helt 3. dokument.

Altså kan en eventuel relevansvurdering variere væsentligt, alt efter hvilken kontekst man ser den i. Vi argumenterer her for at bruge begrebet relevans, mere som en overvejende faktor. Vores co-citationsanalyse afspejler en konsensus af referencernes tilhørsforhold på forskningsområdet (niveau), da de er dannet ud fra flere forskeres mål, med at anvende både vores kerneartikler og dennes referencer i samme dokument. Hvad disse så end måtte være, ville det kræve en dybdegående kvalitativ undersøgelse at finde ud af.

### *STANDARDISERING VS. OBLITERATION*

Generelt er der bred enighed mellem de forskere, vi har gennemgået i teorien, omkring forskelle i citationsadfærd imellem videnskabelige felter. Det er den gængse opfattelse, at visse referencer ofte bliver højt citerede grundet en formatering til standardreference, et 'concept symbol', eller de bliver anerkendt i sådan grad, at der bliver skabt enighed om dennes eksistens i forbindelse med særlige ideer eller teorier. (jf. yderligere 'biased citing' og 'persuasiveness') Dette gælder også den modsatte vej, hvor en reference bliver betragtet som almen viden inden for feltet og derfor ikke citeres, dét som af Garfield bliver betegnet som obliteration. Man kan her diskutere, hvorvidt en standardisering af en reference kan være et skridt på vejen til obliteration. Det kan i hvert fald være et problem når man sammenligner videnskabelige felter, da obliteration er et fænomen der oftere forekommer inden for naturvidenskabelige felter og standardisering oftere forekommer inden for de humanistiske felter. (Hyland, K 1999) Dette kan have noget at gøre med, hvordan forskere betragter deres målgrupper; en artikel inden for fysik vil ofte henvende sig til akademikere inden for feltet og indeholde en høj mængde af fagtermer og formler, der kræver en vis viden på området, derfor kan det virke nedgørende at citere det åbenlyse. Humaniora og specielt samfundsvidenskab på den anden side kan henvende sig til - og anvendes inden for - meget brede målgrupper og det vil derfor være nødvendigt at dokumentere den brugte baggrundsviden. (Hellqvist 2010)

Som følge af dette, ses også en tendens til at artiklerne inden for fysik generelt bygger på nyere litteratur end filosofi, hvor filosoffer som Platon og Kant stadig bliver citeret flittigt. De sociologiske tekster viser den største tendens til at bruge 'gamle' referencer. I den forbindelse er det samtidig vigtigt at tage højde for, hvilken artikel der er tale om, en historisk artikel eller en oversigtsartikel vil logisk nok indeholde flere ældre referencer. Man kan antage at sociologi ikke har tilbøjelighed til, at referencer bliver oblitereret i samme grad som de andre felter. Men at de derimod har større tendens til at referencer bliver standardiseret.

De mange ældre referencer inden for filosofi kan også være et resultat af at referencen enten er genstand for kritik eller at den skal fungere eksemplarisk (Shadish et al. 1995)



### *I VIDENSKABSTEORETISK PERSPEKTIV*

I teoriafsnittet om citationsadfærd er nævnt forskellige undersøgelser omkring citationsadfærd, samt forskellige teorier omkring dette emne. En pionér inden for området er Garfield, der allerede i 1965 opstillede den tidligere nævnte liste over de forskellige bagvedliggende årsager der kan være for at citere. Det første punkt på listen er 'Paying homage to pioneers' – altså at grunden til at citere er at ære og anerkende en tidligere forskers værk. Følger man den normative citationsteori, ser man dette som den hovedsagelige grund til at citere andre forskere. På Brooks' liste fra 1985 svarer nogle af punkterne til dem på Garfields liste. Det interessante ved undersøgelsen er, at Brooks punkt omkring 'persuasiveness' viste sig at være den primære grund til at citere. Dette punkt figurerer slet ikke på Garfields liste, hvilket sandsynligvis skyldes at Garfield er tilhænger af den normative citationsteori, der er udsprunget af Mertons videnskabelige norm. I modsætning hertil står den socialkonstruktivistiske videnskabsteori, som også beskæftiger sig med selve formidlingen af viden. Følger man denne teori, ser man udover resultaterne af videnskabelige undersøgelser, også på de sociale og politiske kontekster produktionen af viden opstår i. Socialkonstruktivisme i videnskabssociologisk **bibliometrisk** sammenhæng, tolker referencer som udtryk for en overtalelsesstrategi, altså den 'persuasiveness' som Brooks opstiller, som en bevæggrund for at citere. Den normative citationsteori vil fordre at alt skal citeres efter CUDOS normerne, mens den socialkonstruktivistiske vil mene, at man skal bruge de referencer, der er nødvendige for at overbevise læseren om sine påstande.

Case & Miller (2011) stiller også spørgsmålstegn ved, hvor stor en rolle 'persuasion' egentlig spiller i den videnskabelige skrivetradition og inden for videnskabssociologien. De anerkender at situationen ikke nødvendigvis er så sort/hvid, som nogle forskere gerne vil have det til at se ud. Dette understøtter Nicolaisens (2003) påstand om, at man kan se på citationer som en fælles diskursiv handling inden for de forskellige felter, der muligvis kan skabe bro mellem de to videnskabsteoretiske positioner.

Til trods for mange års forskning på området findes der stadig ikke nogen specifik citationsteori, men forskellige videnskabelige tilgange, hvorpå man kan anskue

valget af referencer. I et nyere paper fra 2009 genoptages diskussionen om citationsteori med et nyt perspektiv (Camacho-Minano. & Nüñez-Nickel, 2009). Her opdeles valget af en reference i tre lag, hvor det første er givet og består af at sortere de dokumenter fra, der egentlig ikke er en mulighed på grund af eksempelvis en sprogbarriere eller utilgængelighed. Dernæst er andet valg et funktionelt valg, der argumenteres for at være objektivt, da der fravælges alle dokumenter, som ikke bidrager til indholdet af forskningen. Der er først i det sidste og tredje lag, at der vælges subjektivt mellem de resterende referencer. Og her bliver valget truffet ud fra eventuelle forudindtagede holdninger der eksisterer hos forfatteren.

De mener dermed heller ikke, at den socialkonstruktivistiske teori er et modstykke til den normative teori, men en anderledes måde at anskue den tredje fase af referencevalget på. I den tredje fase vælges nemlig subjektivt mellem en række sammenlignelige dokumenter, ud fra kriterier som kun forfatteren kender. Disse kriterier kan bunde i forudindtagede holdninger om, hvilket tidsskrift det pågældende dokument er udgivet i, hvem har skrevet artiklen, osv.

Så det er altså ikke et spørgsmål om, hvorvidt man skal tro på den socialkonstruktivistiske eller den normative model når det kommer til citationsadfærd, men derimod hvor stor indflydelse de hver især har i forhold til forfatterens citationsadfærd.

I forhold til de forskellige undersøgelser foretaget vedrørende forfatteres citationsadfærd, kan man sætte spørgsmålstegn ved de årsager, der angives til at vælge forskellige referencer og hvad der i virkeligheden ligger bag. Disse valg og fravalg i forhold til hvornår der gives en reference, indvirker på, hvordan selve referencelisten i artiklen kommer til at se ud, og har dermed også direkte betydning for, hvilke resultater en citationsanalyse vil give. Man kan hævde, at de næppe vil indrømme, hvis deres arbejde primært fokuserer på at beskæftige sig med andres "dårlige" forskning i forhold til deres egne vigtige og relevante forskning. Ligeledes vil de næppe sige, at den vigtigste grund til at citere er at citere, hvad andre forventer af dem, da dette vil vidne om en uselvstændighed i forhold til dem som forskere. (Brooks 1985)

Vi mener, at den menneskelige faktor gør, at det under alle omstændigheder er forskellige valg man foretager. To forskere vil næppe angive nøjagtig de samme referencer. Det er netop denne menneskelige faktor, der gør det problematisk at vurdere, om citationer har den samme værdi, da valget af referencer og årsagen til at benytte netop disse, varierer helt ned på individplan. De forskelle i resultaterne vi bliver præsenteret for udspringer, givetvis af den enkelte forskers måde at håndtere sin viden på, i sammenhæng med normer og traditioner inden for dennes forskningsmiljø.

Sammenfattende kan man sige, at der i forhold til spørgsmålet om, hvorvidt nogle citationer overordnet kan siges at være af større relevans for artiklen, er en masse bagvedliggende grunde man skal være opmærksom på. Alligevel er det i sidste ende kun den enkelte forsker eller forskergruppe, der kender den reelle årsag bag hver enkelt reference.

## KONKLUSION

Vi har nu fremlagt en undersøgelse, der viser, hvordan udvalgte artikler inden for fagdisciplinerne fysik, sociologi og filosofi er co-citerede med deres egne referencer. Undersøgelsen viser en stor variation i mellem felterne, hvilket både ses i antallet af citationer og dermed også co-citationer. Til trods for dette, danner der sig en gennemgående tendens til, at de referencer der har højest co-citationsfrekvens med den citerende kerneartikel, optræder første gang i begyndelsen af artiklen. Dette medfører, at vi kan konkludere, at vores hypotese 1 holder stik inden for alle 3 fagdiscipliner. Det kom desuden frem i vores undersøgelse, at brugen af ORL's kan have betydning for resultaterne, i form af de tildelte referencenumre og derfor bør tages i betragtning, når diagrammerne vurderes. Når vi tager de vigtigste forbehold for videnskabelige artiklers opbygning, relevante forskelle indenfor citationsadfærd og forskningstraditioner, samt overvejer metodens begrænsninger, vil vi dog ikke konkludere at en references placering er en sikker indikator for dennes relevans. Vi vil alligevel hævde, at tages der højde for problemstillinger, der omhandler referencernes placering, alder og udformning, vil en analyse som denne, kunne fungere som en god indikator for relevans. Vi vil dog ikke udelukke at dette er en risikabel metode, der kan medføre at relevante referencer bliver ignoreret eller overset.

For at minimere disse risici er det nødvendigt at have et godt kendskab til, hvad der karakteriserer de forskellige forskningsområders citationsadfærd. Undersøgelsen viste store forskelle, både i antallet af referencer og det gennemsnitlige antal sider, hvorpå disse var repræsenteret. Desuden blev der identificeret store forskelle i måden referencerne blev introduceret i teksten; dette var specielt markant i sammenligningen af sociologi og filosofi vs. fysik. Der bør altså tages forskellige forbehold, afhængigt af om man gør brug af naturvidenskabelige artikler eller samfundsvidenskabelige og humanistiske artikler. Det samme gør sig gældende, når det kommer til alderen på de anvendte referencer, hvor der med visse undtagelser generelt blev identificeret nyere referencer i fysik end i de andre felter. Dette verificerer Hypotese 2.

Det er i en undersøgelse som denne, ikke muligt at konkludere noget på citationsadfærd i forbindelse med hver enkelt artikel, men der ses generelt en tendens i alle disciplinerne til at forskere nævner deres højt co-citrerede referencer flere gange i artiklen, hvilket vi yderligere betragter som en indikator for relevans. Dette ses tydeligst i artiklerne fra fysik og sociologi.

Til trods for de relative ensartede tendenser, der prægede alle tre felter, var der stor forskel i styrken af disse, og mængden af afvigelser fra tendensen er stadig markant.

Vi konkluderer, at de meget store forskelle, ikke bare fra artikel til artikel, men også i særlig grad mellem felterne, primært bunder i forskelle i citationsadfærd. Traditioner og sociale normer når det kommer til at citere, er gennem teorien blevet identificeret og i mange tilfælde kan disse bruges som argumenter for de fremfundne resultater. Fænomener som obliteration og standardisering er specielt noget, vi anser for værende en afgørende feltforskelle.

Sammenfattende vil vi yderligere konkludere, at de store forskelle i citationsadfærd er en problematik, der forstærkes af, at det i sidste ende altid vil være ubevidst for udefrakommende, med hvilket motiv en citation gives. Dette vil altid være en beslutning, der tages af den forsker eller forskergruppe der udgiver artiklen og bør ses i konteksten af deres forskningsmiljø.

Vi vil afslutningsvis argumentere for co-citationsanalysen som metode i denne sammenhæng. Her er det nemlig ikke den enkelte forsker, men hele det videnskabelige felt, der mere eller mindre uafhængigt af hinanden, vælger at citere de samme dokumenter sammen.

# METODEKRITIK

I forhold til MacRoberts & MacRoberts liste over problemer ved citationsanalyser, er vi klar over, at vi ved en yderligere granskning af vores søgninger, kunne have udelukket nogle af de tekniske fejl som nævnes. Vi kunne i højere grad have gennemset søgningerne, eller endda have udvidet hver enkelt søgning, ved hjælp af 'expand' kommandoen i dialog, og derefter udvalgt de rette referencer. Eller – vi kunne i hvert fald have frataget dem med åbenlyse fejl, så som stavefejl, som vi havde mulighed for at identificere, da vi sad med kerneartiklernes referencelister i fysisk form. Vi har også valgt at se bort fra selvcitationer som en mulig fejlmargen. Og helt grundlæggende kunne vi have valgt tre helt andre felter, hvorved undersøgelsen nok havde set helt anderledes ud.

Desuden var der stor variation i antallet af citationer, som kerneartiklerne havde modtaget meget fra databasen Web of Science og fra resultaterne i Dialog. Dette kan være et eksempel på fejlindekseringer, hvor en forkert repræsentation af referencen modtager der var tiltænkt den egentlige repræsentation. Denne fejlmargen bør der tages højde for når resultaterne vurderes, da det i princippet er en fejl der kan forekomme i forbindelse med alle referencerepræsentationer.

Vi er altså klar over, at det ikke på nogen måde er en fyldestgørende undersøgelse, og at man kunne have valgt mange andre metoder til at undersøge nogle af de områder, som vi beskæftiger os med. Blandt andet kunne vi lige som Hyland (1999) have undersøgt, hvordan felterne differentierede sig fra hinanden i brugen af integrale og non-integrale referencer. Vi kunne også have undersøgt relevans i forhold til kerneartiklerne og de citerede artiklers keywords. Eller vi kunne have udført en kvalitativ spørgeskemaundersøgelse af forskere på IVAs citationsadfærd.

I det hele taget er der utrolig mange vinkler på og muligheder for at udføre citationsanalyser – og denne opgave er bare en af dem.

# KILDEFORTEGNELSE

## WWW

- **DIALOG POCKET GUIDE**, downloadet herfra:  
<http://support.dialog.com/searchaids/dialog/pocketguide/> (maj 2011)

## ARTIKLER

- **BAZEMAN, C.(2000)***Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. WAC Clearinghouse Landmark Publications in Writing Studies:  
[http://wac.colostate.edu/books/bazerman\\_shaping/](http://wac.colostate.edu/books/bazerman_shaping/) (lokaliseret 18. maj 2001).Originally **Published** in Print, **1988**, by University of Wisconsin Press, Madison, Wisconsin. Herikapitel 10.
- **BROOKS, T.(1986)**, *Complex Citer Motivations*. Journal of the American Society for Information Science, 37 (1): 34-36
- **CAMACHO-MINANO, M. &NÚÑEZ-NICKEL, M.(2009)** *The Multilayered Nature of Reference Selection*. Journal of the American Society for Information Science and Technology, 60 (4): 754-777
- **CASE, DO. & MILLER, JB.(2011)**,*Do Bibliometricians Cite Differently From Other Scholars?*. Journal of the American Society for Information Science and Technology, 62 (3): 421-432
- **COZZENS, SE. (1985)***Comparing the Sciences: Citation Context Analysis of Papers from Neuropharmacology and Sociology of Science*. Social Stuides of Science , 15: 127-153
- **GARFIELD, E., (1965)**, *Can Citation Indexing Be Automated?* Statistical Association Methods for Mechanized Documentation: Symposium Proceedings, National Bureau of Standards (NBS) Miscellaneous Publication 269, pp. 189-92.
- **GARFIELD, E. (1975)** *The 'obliteration phenomenon' in science and the advantage of being obliterated*. Essays of an Information Scientist, 2: 396-398, 1974-76 & Current Contents, 51/52 : 5-7, December 22, 1975
- **GARFIELD, E.,(1996)***When to Cite*. Library Quarterly, 66 (4): 449-458

- **GLÄNZEL, W. & SCHOEPFLIN, U.(1999)**, *Abibliometric study of reference literature in the sciences and social sciences*. Information Processing and Management, 35: 31-44
- **HARGENS, LL.(2000)**, *Using the Literature: Reference Networks, Reference Contexts, and the Social Structure of Scholarship*. American Sociological Review, 65 (6): 846-865
- **HELLQVIST, B.(2010)**, *Referencing in the humanities and its implications for citation analysis*. Journal of the American Society for Information Science and Technology, 61 (2): 310-318
- **HJØRLAND, B (2010)**, *The Foundation of the Concept of Relevance*. Journal of the American Society for Information Science and Technology, 61 (2); 217-237
- **HYLAND, K. (1999)**, *Academic Attribution: Citation and the Construction of Disciplinary Knowledge*. Applied Linguistics, 20 (3): 341-367
- **LARIVIÈRE, V. ET AL. (2006)**, *The Place of Serials in Referencing Practices: Comparing Natural Sciences and Engineering With Social Sciences and Humanities*. Journal of the American Society for Information Science and Technology, 57(8): 997-1004
- **MACROBERTS, M. AND MACROBERTS B., (1989)** *Problems of citation analysis: a critical review*, Journal of the American society for information science, 40(5):342-249
- **MERTON R K. (1965)** *On the shoulders of giants: a Shandean postscript*. New York: Harcourt Brace & World, p. 218-19.
- **MERTON, R., (1968)**, *The Matthew Effect in Science*. Science 159 (3810): 56-63
- **NICOLAISEN, J.(2003)**, *The Social Act of Citing: Towards new Horizons in Citation Theory*. Proceedings of the American Society for Information Science and Technology, 40 (1): 12-20
- **OPPENHEIM, C. & RENN, S. (1978)**, *Highly cited old papers and the reasons why they continue to be cited*. Journal of the American Society for Information Science, 29 (5): 225- 231
- **OSAREH, F. (1996)** *Bibliometrics, Citation Analysis and Co-citation Analysis: A Review of Litterature I, Libri*, 46: 149-158



- **OSAREH, F. (1996)** *Bibliometrics, Citation Analysis and Co-citation Analysis: A Review of Litterature II*, Libri, 46: 217-225
- **SEGLEN, P., (1996)** *Bruk av siteringer og tidsskriftimpaktfaktor til forskningsevaluering*, Biblioteksarbejde, 17 (48): 27-34
- **SHADISH, W., TOLLIVER, D., GRAY, M. & GUPTA, SKS. (1995)**, *Author Judgements about Works they Cite: Three Studies from Psychology Journal*. Social Studies of Science, 25 (3): 477-498
- **SMALL, H. (1973)**, *Co-citation in the scientific literature: A new measure of relationship between two documents*, Journal of the American society for information science, 24: 265-269
- **SMALL, H. (1978)**, *Cited Documents as Concept Symbols*. Social Studies of Science 8(3): 327-340
- **SMITH, L. (1981)**, *Citation Analysis*, Library Trends, Summer 1981
- **VOOS, H & DAGAEV, K. (1976)**, *Are All Citations Equal? – Or Did We Op. Cit. Your Idem?* The Journal of Academic Librarianship 1(6): 19-21
- **WHITE, HD. (2001)**, *Authors as citers over time*. Journal of the American Society for Information Science and Technology, 52 (2): 87-102

#### MONOGRAFIER

- **DE BELLIS, N. (2009)**, *Bibliometrics and Citation Analysis – From the Science Citation Index to Cybermetrics*, The Scarecrow Press, Inc.
- **SWALES, JM. (1990)** *Genre Analysis – English in Academic and Research Settings*, Cambridge University Press