

## **Anonym kritikk:**

### **Peer Review\* satt på prøve av vitenskaps-teori og - forskning<sup>1</sup>**

---

*Gerhard Fröhlich*

#### **0 Evaluering som slagord**

Evalueringer av vitenskapelige prestasjoner er for øyeblikket gjenstand for mange politiske debatter, lovforslag og oppsiktsvekkende nyheter. Men selv ikke de øverste beslutningstakere støtter evalueringenes synspunkt og reformforslag med bakgrunn i det fond av modeller og resultater som den teoretiske-

---

\*Det engelskspråklige uttrykket er velkjent, men på norsk snakker vi heller om *fagkyndige konsulenter* ved vitenskapelig publisering o.l. Dette er den første fotnoten til oversettelsen som er SÅ-redaksjonens tillegg, ikke forfatterens, alle merket med \*. Fotnoter merket med tall er forfatterens opprinnelige.

<sup>1</sup> Deler av denne artikkelen ble forkortet publisert i Fröhlich 1999b og 2002b. Takk for en kritisk gjennomlesning går til Simone Griesmayr og til Michael Strähle for verdifulle kommentarer.

empiriske vitenskapsforskning og informasjonsvitenskap<sup>2</sup> har frambrakt; de kommer åpenbart bare med private eller ideologiske meninger. De vitenskapelige rammebetingelsene blir dermed nedvurdert både i den vitenskapelige analysen og i den metavitenskapelige refleksjon. Skal vitenskapene bli drevet på et irrasjonelt grunnlag?

Sett fra et vitenskapsteoretisk ståsted har evaluering, i betydning kritikk og vurdering av teorier, metoder og resultater etter “vitenskapsinterne” kriterier, alltid hørt til den vitenskapelige hverdagspraksis. Evaluering hører til grunnlaget for alle vitenskapelige metoder. Men mange vitenskapsmenn har i det siste blitt opprørt over at man forsøker å fastsette evalueringsprinsipper utenfra, prinsipper som ligner på slike som brukes innen markedsforskning eller innen produksjonskontroll.

Det det gjelder er mot-evalueringen: Hvem evaluerer evalueringen? Hvem bedømmer bedømmeren? Hvem målermålestokkene? Ved siden av den evaluerende sciento-metrien (stikkord: *impact factor*; for kritikk se Fröhlich 1999a, 2002c) har i det siste også forskjellige *peer-review*-metoder blitt gjenstand for kritiske debatter, medieomtale og empiriske undersøkelser.

## **1. Kritikk som definerende kjennetegn ved vitenskap og kritikkens avmakt i hverdagspraksis**

---

<sup>2</sup> Vitenskapsforskningen er anvendelsen av vitenskapelige metoder på vitenskapsmenn, evt. vitenskapelige institusjoner og deres praksis (Fröhlich 2000). De områdene innen informasjonsvitenskapen som holder på med vitenskapelig kommunikasjon og evaluering av vitenskapelige ytelser regnes til vitenskapsforskningen.

“Hensynsløs” (Karl Popper) kritikk er et kjennetegn på vitenskap. Men dette kravet fra den normative vitenskapsteorien blir sjelden oppfylt innen den vitenskapelige offentlighet. Den “realeksisterende” vitenskapshverdag ser ofte annerledes ut.

I den offentlig tilgjengelige vitenskapelige kommunikasjon dominerer *de facto* fra mild til ærbødig høflighet. Det er ikke risikofylt å fordele komplimenter. Mer eksplosiv informasjon får man via uformelle kanaler, internett og massemedia. Når kritikk blir offentliggjort blir den nesten ikke tatt til etterretning av dem det berører. Selv om artikler trekkes tilbake pga. forfalskning har det nesten ingen konsekvenser, og man fortsetter å sitere dem uten kommentarer.

Den offentlige kritikken er bare toppen av isfjellet av den kritiske forskningskommunikasjon, som i hovedsak går for seg i to kontekster. Kritikken uttrykkes først og fremst:

- i uformelle dyader og smågrupper (f.eks. i form av manuslesning i en vennekrets)

- innen tidsskriftforlagenes rammer, hvor man følger et arkan- eller hemmelighetsprinsipp: Interaksjonen mellom utgiver, sakkyndige konsulenter, og manusforfatter fungerer etter modell av konspirative organisasjoner. Det vil si at sakkyndige konsulenter vanligvis er anonyme, at nesten ingen av deltakere vet om hverandre, og at manusforfattere bare sjelden får se sine sakkyndige uttalelser i utdrag eller i det hele tatt\*.

---

\* SÅ-redaksjonen er kjent med at flere større norske samfunnsvitenskapelige periodika hyller prinsippene om double blind (se fotnote 3), og om at alle konsulentuttalelser skal sendes forfatteren, hva enten han/hun blir avvist eller ikke. I hvor stor grad slike prinsipper faktisk etterleves i praksis er derimot ukjent utenfor redaksjonene.

Kritikere ser at åpenheten til den vitenskapelige informasjonen trues ved en slik hemmeligholds-praksis som er vanlig nå for tiden (Jasanoff 1985, Munthe og Welin 1996). Som en følge av at disse informasjoner holdes skjult (Fröhlich 1998) blir “context of discovery” til en forretningshemmelighet, og dermed blir feil og villspor ofte og unødvendig gjentatt.

## **2. Peer-review og avvisningsrater som prestisjemål for vitenskapelige tidsskrifter**

I vitenskapen hersker en offisiell sensur som kan sammenlignes med den katolske kirkens “nihil obstat”. Riktignok fungerer den relativt: ingen blir hindret i å forske. Forutsatt finansering av egne midler og utgivelse på et lite forlag eller lagt ut på web finnes det en prinsipiell mulighet for alle til å forske. Men fra evaluerings-hardlinere blir bare bidrag i “*peer reviewed*” tidsskrifter akseptert som vitenskapelige. Bøker eller detaljerte forskningsrapporter (“ubedømt litteratur”) er bare av liten verdi.

“Peer-review” er navnet på det kontrollsystemet som er vitenskapens “biltilsyn”. Sakkyndige konsulenter (eller rettere sagt *anonyme* sakkyndige konsulenter som ikke trenger frykte for hevn) garanterer kvalitet, hindrer overflod og kaos – det er den offisielle mening. Sakkyndige konsulenter råder ikke bare over hvilke manuskripter som kommer på trykk, men avgjør også fordeling av forskningsmidler, hvem som skal delta på kongresser, og velger ut prisvinnere på alle trinn i den vitenskapelige karriere.

Men peer-review-systemet som *ett* enhetlig evalueringssystem finnes egentlig ikke, til tross for at dette ofte blir hevdet. Jeg påpeker derfor forskjellene: at fra fag til fag er prosedyrer og bedriftskultur omkring “refereeing” altfor vilkårlig forskjellige og spesielt altfor forskjellig vilkårlige. Dette gjelder også for kongressorganisasatorer, for

forskningsentreprenører og for utgivere av tidsskrifter og bøker. Selve definisjonen av “peer” (egentlig: jevngammel eller av samme rang; i hovedsak med betydningen en vitenskapelig konsulent fra samme fag som er spesialist på emnet) er meget upresis og variabel.

Heterogeniteten til, ja usammenlignbarheten av de eksisterende peer-review-variantene (slik som enkelt-, dobbelt- og tredobbelt blindtest)<sup>3</sup> henger sammen med deres *forskjellige funksjoner*, noe som følger forholdet mellom tilbud og etterspørsel:

- I fag med en *lav avvisningsgrad (rejection rates)* dvs. fra 20 – 30%, f. eks. i fysikk, skal sakkyndige konsulenter foreslå gode artikler til trykking. Derfor blir ofte *den serielle metoden* praktisert, der bare én sakkyndig konsulent blir forespurt. Hvis han er enig, blir manuset trykt. Hvis han avslår, blir en annen sakkyndig konsulent forespurt.

- I fag med *høy avvisningsgrad (80-90%)* som i psykologi skal sakkyndige konsulenter derimot levere argumenter for at manuser kan bli avvist. Derfor foretrekker man den *parallele metoden* med to eller flere sakkyndige konsulenter (jo flere sakkyndige konsulenter desto flere negative sakkyndige uttalelser; ved bruk av fire sakkyndige konsulenter innen statsvitenskap ble for eksempel hvert eneste manuskript avvist av minst en konsulent. Ved uenighet om bedømmelsen blir et manus som regel avvist, eller sendt tilbake med en henstilling om videre bearbeiding (det siste er ofte er det bare en høflig form for avvisning, eller kanskje en mulighet til å bevare ansikt for en utgiver som har satset mye på det avviste manuskriptet).

---

<sup>3</sup> Ved (enkel) blindtest (*single blind*, den vanligste typen) kjenner manusforfattere ikke sine sakkyndige konsulenter. Ved dobbel blindtest (*double blind*) får heller ikke den sakkyndige konsulenten vite identiteten til forfatteren - ihvertfall prøver man dette. Ved tredobbelt blindtest (*triple blind*) får i tillegg utgiveren ikke vite noe om identiteten til forfatteren. Den siste metoden blir dessverre sjelden praktisert; men det forekommer, f. eks. av *Zeitschrift für Soziologie*, Bielefeld. Der skal det ha hendt at manus fra stjerneforskere har blitt avvist.

Avvisningsraten er stundom avhengig av helt banale faktorer. På den nevnte aksene mellom tilbud og etterspørsel finnes det forskjellige svar: tidsskrifter i fysikk er sterkt økende når det gjelder volum og undertidsskrifter: 10 000 sider per årgang er ikke noen sjeldenhet. Derimot virker 200-300 sider per årgang i psykologiske tidsskrifter ganske beskjedent, og nesten som en kunstig skapt tilstramning, ikke minst da med tanke på de prisgunstige digitale mulighetene.

Nå er den vanlige myten at jo høyere avvisningsrate desto strengere er tidsskriftet og desto høyere er den vitenskapelige kvaliteten både når det gjelder tidsskrift og enkeltartikler. Noen vitenskapelige tidsskrifter later som om kravet om å forkorte eller overflytte til andre avdelinger av tidsskriftet (f. eks. forskningsnotat i stedet for artikkel) er en avvisning. I konkurransen om forfattere, lesere og sponsorer fungerer både slike kunstig produserte avvisninger og "*peer-review*"-stemplet som *prestisjessymbol*:

- Noen ganger er det uklart hvilke tidsskrifter som bruker eksterne konsulenter, i hvilken grad, og hvilke overhode ikke. Forskjellige halvoffentlige tidsskriftskataloger gir bare delvis opplysninger (jfr. Eldredge 1997), Bare enkelte tidsskrifter rubriseres som peer reviewed.

- Ofte er det helt uklart hvordan peer review konkret blir praktisert, dvs. leseren får ingen opplysninger om dette.

- Mange *hevder* at de publiserer etter peer-review-metoder. Men de praktiserer dette ikke, eller bare i liten grad. Eller de gir generøse unntak som praktiseres enten uformelt eller formelt (f. eks. er nord-amerikanske akademimedlemmer unntatt fra vurdering i disse akademienes egne publikasjoner).

- Mange tidsskrifter forholder seg selvmotsigende. Peer-review-prosedyrer som har vært praktisert gjennom flere tiår har blitt stoppet etter oppsiktsvekkende feilavgjørelser.

- Selv blant velansette sakkyndig bedømte tidsskrifter blir også *bare en del*, ofte bare en brøkdel av tidsskriftet *faktisk kvalitativt eksternt evaluert*. Likevel lar de alt de publiserer smykke seg med dette rykte. Dette kan sammenlignes med den amerikanske passivrøyker-skandalen; bedriftsadvokater i et amerikansk tobakk-konsern som ble truet med høye klagekrav, fikk medisinske toppforskere til å signere leserinnlegg i meget anerkjente medisinske tidsskrifter – selvfølgelig mot godtgjørelse – slik at konsernets prosessfullmektig med suksess kunne påberope seg disse i retten. Fordi andelen av “ikke siterte” bidrag øker impact-faktoren på en kunstig måte, blir slike bidrag stadig mer populære.<sup>4</sup>

- Mange tidsskrifter smykker seg med en stor stall av sakkyndige konsulenter, men ofte blir faktisk bare en brøkdel av disse anvendt. Noen få “old boys” får nesten alle konsulentoppdragene, og behersker på denne måten et helt fag fordi de ofte arbeider samtidig for flere tidsskrifter.

Riktignok finnes det tallrike muligheter for å *unngå* peer-review-bedømmelse. Når det gjelder forskningsfinansiering kan følgende først og fremst gi resultater:

---

<sup>4</sup> Det markedsdominerende *Institute for Scientific Information (ISI)* definerer sin *impact-factor* på følgende måte: De teller alle siteringer som en tidsskriftårgang får i de første to år etter at publikasjonen foreligger - siteringer som finnes i tidsskriftene til ISI - databasene (som er en brøkdel til den verdensomspennende “output” av vitenskapelige publikasjoner). Summen blir delt på antallet “citable” artikler (som kan siteres) - og bare bare på disse. Slik blir allmenne tidsskrifter som “Science” med sine mange “non-citable” forord, leserinnlegg og journalistiske bidrag de ledende, pga. non-ciatable bidrag som dessuten ofte blir skrevet av kjente forfattere. Til tross for at disse bidrag klassifiseres som ikke-siterbare av dem som arbeider med ISIs impact factor-indeks, bidrar mange siteringer av dem til hele summen og øker dermed ISI-impact-faktoren (Fröhlich 2002c).

- “*pork barrel*”-systemet (dvs. prosjekter som er politiske bestillingsverk til forskningsinstitutter i USA, ment å underbygge bestemte politiske standpunkter, (jfr. Savage 1999).

- forskning knyttet til rustningsindustrien (“classified”, dvs. unntatt offentlighet), og

- industrioppdrag eller bidrag fra private sponsorer. Særlig de sistnevnte mangler til dels universelle vitenskapelige kriterier. Der teller ofte bare sinnelaget.<sup>5</sup>

Private amerikanske universiteter som hele tiden fremstår som et forbilde og en elite, kan benytte penger både fra industri og stiftelser som erstatning for en peer-reviewed litteraturliste (ved ansettelse). Det er ikke bare Stanford som praktiserer prinsippet “*salary splitting*”, dvs. at professorer ikke bare må skaffe penger for å finansiere sine medarbeidere, men også til å finansiere egne stillinger.<sup>6</sup> Omfanget av eksterne midler gjelder nå også i Tyskland som et kjennetegn på vitenskapelig ytelse, i hvert fall gjelder er omfanget av eksterne midler som kjennetegn på instituttens kvalitet.

For vitenskapsmenn som ikke har en så stor finansiell “medgift” lønner det seg derfor i hvert fall innen naturvitenskapene å publisere artikler på en preprint-server.

---

<sup>5</sup> Et eksempel: Stanford-ledelsesfunksjonær Terman aksepterte uten problemer betingelsene for “funds that had been provided by William Coe, a generous benefactor (...) Coe, a right-wing businessman, had placed restrictions on his patronage; his money was to be used to meet “the threat of Communism, Socialism, collectivism, Totalitarianism and other ideologies opposed to the preservation of our System of free Enterprise” (Lowen 1977: 216, ftn67). Den daværende professor i statsvitenskap ble derfor sagt opp og det blir innstilt en beinhard antikommunist istedet.

<sup>6</sup> Sammenlign Stanford-casestudien til R.S. Lowen (1977). Også et annet amerikansk universitet praktiserer “to fire anyone who doesn’t attract research grants that provide between 50 and 100 per cent of his salary” (Lowen 1977: 237).



Innen fysikken har disse allerede oppnådd et så godt rykte at server-artikler som er sitert flere ganger ikke lenger kan bli avvist i tidsskrifter. Psykologiske tidsskrifter derimot avviser artikler som allerede er publisert på internett på grunn av den såkalte “Ingelfinger-regelen”, oppkalt etter en utgiver/redaktør som innførte forbud mot gjenpublisering innbefattet ren web-publisering.

### **3. Alvorlige prestasjonsfeil fra det sakkyndige konsulent- ”systemet” i forbindelse med bedrageri og forfalskning**

Tilliten til at peer-review-metoder fungerer burde også være svekket fordi den åpenbart har mislyktes i tilfeller med *plagiat, bedrageri og forfalskning*. Til tross for åpenbare og klossete manipuleringer avslørte anonyme sakkyndige konsulenter bare et fåtall av de tallrike kjente saker som bryter med vanlige vitenskapelige framgangsmåter. Når sakkyndige konsulenter unntaksvis oppdager en forfalskning, våger utgivere ofte ikke å offentliggjøre saken eller gå til anmeldelse fordi de er redde for å miste anseelse og for å kunne bli truet av klager fra forfalskerne. Alle saker som hittil er oppspørt ble som regel avslørt gjennom interaksjonsprosesser utenfor den formelle vitenskapskommunikasjonen, fortrinnsvis ved personlig selvavsløring eller ved tilståelser fra medarbeidere med dårlig samvittighet.

De som har æren av å avsløre vitenskapssvindler og av å forandre atferden til vitenskapelige institutter og profesjonssammenslutninger (jf. Fröhlich 2002a) i deres forhold til plagiat og forfalskning må først og fremst – ved siden av forskere med stort moralsk mot og som våger å gi eksplosiv informasjon videre – uten tvil tilskrives *utenomvitenskapelige* instanser som politikk og media. Uten disse ikkevitenskapelige informasjonskilder hadde det neppe blitt offentliggjort *en eneste* av de i “førstegenerasjons” kjente svindelaaffærer. Vitenskapelige institusjoner og vitenskapelige tidsskrifter var i

begynnelsen ikke villig til å nevne temaet i det hele tatt. Den kjente Cyril-Burt-skandalen (oppdikning av tvillingstudier av intelligens for å styrke arvelighetslæren) begynte først for fullt gjennom en presse-artikkel (Fröhlich 2001).

Ved de nyere tyske svindelskandaler innen medisin og biologi (kreft-forskningsskandalen Hermann, Brach og Mertelsmann; Max Planck Institutt for avlsforskning (*Züchtungsforschung*, Köln) klarte dusinvis av forfalskede publikasjoner å passere peer-review-prosedyrene til de mest anerkjente vitenskapstidsskrifter (inkludert *Science* og *Nature*) til tross for bedragerske manipulasjoner som delvis var helt åpenbare – for eksempel ble den samme illustrasjonen brukt som bevis for helt forskjellige målingsserier (Finetti og Himmelrath 1999).

Noen sakkyndige konsulenter var til og med selv innblandet i svindel-skandaler; det verserer saker for kontrollinstanser eller domstoler. De allerede nevnte kreftforskere Friedhelm Hermann og Marion Bach misbrukte også sine stillinger som referees: de avviste en prosjekt-søknad og sendte den selv inn uforandret til samme stiftelsen – med suksess: stiftelsen Thyssen i Köln innvilget dem 260 000 tyske mark (Finetti og Himmelrath 1999: 48). Generelt må man anta at sakkyndige konsulenter kan la seg inspirere av idéer som kommer fra innsendere og så bruker disse som 'egne idéer'. "Kryptamnesier" (kilde-hukommelsestap) dvs. ubevisste plagiater (det at man etter å ha glemt kilden husker det erfarte som ens egen nye idé) er derfor vanskelig å unngå helt blant de mest velplasserte forskere og forskningsentreprenører med mange verv.

#### **4. Peer-review-metoden bedømt ut fra empirisk forskning**

Tallrike negative resultater fra empiriske undersøkelser av peer-review-systemet bidrar til å ødelegge systemets rykte. Eksempler av dette skal nå kort presenteres.

#### 4.1. “Eksperimentelle” peer-review-studier

Ved “eksperimentelle” peer-review-studier blir det f. eks. bygd inn feil i manus, eller forskningsresultater blir variert (ved å holde konstant de øvrige manusedelene) for å avsløre referee-fordommer. Iblant blir det til og med sendt inn igjen allerede trykte manuskripter til de samme tidsskriftene:

Peter og Ceci (1982) tok ut en artikkel fra hver av tolv forskjellige psykologiske tidsskrifter. Disse allerede publiserte artiklene ble (etter endring på forfatteren navn og hans institusjonelle tilhørighet pluss noen små ‘kosmetiske’ forandringer) sendt inn for å bli publisert på nytt i de samme tidsskriftene. Bare tre av 12 utgivere eller redaksjonskolleger – mer eksakt *bare tre av 38 utgivere og referees – kjente igjen artiklene* som ble publisert fra ett og et halvt til tre år siden. De øvrige innsendte artikler ble på nytt utsatt for den vanlige vurderingsprosedyren. Resultat: *Bare en eneste artikkel ble godkjent.* 8 av 9 artikler, som allerede var publisert, ble avvist pga. negative sakkyndige uttalelser, først og fremst ble det henvist til “tungtveiende metodologiske feil”. (De resterende tre av 12 var de tre som ble gjenkjent som ‘plagiat’.)

Abramowitz m.fl. (1975) konstaterte en entydig *politisk bias* blant sakkyndig konsulenter innen psykologi (n=322). Overensstemmelse mellom konsulentens og den innsendte artiklens politiske holdning korrelerer negativt med konsulentens evne til å oppdage feil. Hvis en sakkyndig konsulent avviser det politiske standpunktet i en artikkel, finner han langt flere feil og kommer dermed til en langt mer negativ bedømmelse enn ved politisk overensstemmelse.

Baxt m. fl. (1998) laget et fiktivt manuskript (et placebokontrollert dobbel blindstudie av en type migrene-terapi), der de plasserte 10 alvorlige *feil* og 13 *småfeil*. Studien ble sendt til alle sakkyndige konsulenter av *Annals of Emergency medicine* (n=262) med anmodning om å sende en sakkyndig uttalelse, noe 78% av alle sakkyndige konsulenter gjorde. De 15 referees som anbefalte manuset til publika-

sjon, fant 17% av de alvorlige feil og 12% av de mindre alvorlige. De 117 sakkyndige konsulenter som avviste manuskriptet fant riktignok flere feil, men bare 39% av de alvorlige og 25% av småfeilene. De 67 sakkyndige konsulentene som anbefalte en grundig gjennomgang fant 30% av de alvorlige og 22% av småfeil:

“Sixty-eight percent of the reviewers did not realize that the conclusions of the work were not supported by the results (...) Peer reviewers in this study failed to identify two thirds of the major errors in such a manuscript” (Baxt m. fl.: 310).

#### *4.2 Rekonstruksjon eller replikasjon av vurderinger*

Ved rekonstruerte undersøkelser tas allerede avsluttende prosjekt- og manusbedømmelser opp på nytt. For eksempel vurderes aksepterte og avviste prosjektsøknader eller manuskripter en gang til av samme sakkyndige konsulenter fra den samme refereestallen.

50 avviste og 50 godkjente prosjektsøknader til National Science foundation (NSF) ble vurdert en gang til av NSFs sakkyndige konsulenter. Cole m.fl. (1981) kom til følgende resultat: utfallet av konsulentvurderingen og dermed godkjennelse eller avvisning av en søknad er “tilfeldig”.

En kvantitativ svensk undersøkelse (Wenneras og Wold 1997) påviste *nepotisme og sexism* ved sine “postdoctoral fellowship applications” når det gjaldt fordelingsrutiner til Swedish Medical Research Council (MRC). Tilgangen til forskningsdata, dvs. til bevismateriale om søkere og vurderingen av dem hadde de kvinnelige forskerne kjempet seg til rettslig ved hjelp av den svenske “Freedom of Press Act”. MRC nektet først å utlevere forskningsdata (i nesten alle andre land hadde det neppe vært mulig å gjennomføre en slik undersøkelse). De nøkterne resultatene: *Kvinner måtte være 2,5 ganger så produktive som menn for å få en like god vurdering som sine mannlige kolleger fra MRC-referees*. Det ble i tillegg påvist at bekjentskap har mye å si.

*Personlige relasjoner til et kommisjonsmedlem* skaffet vesentlige fordeler ved vurderingen, dvs. en enorm innsparing ved de dokumenterte publikasjonene:

“(A)n applicant lacking personal ties with the reviewers needed to have 67 more impact points than an applicant of the same sex who was associated with one of the reviewers, to be perceived as equally competent. So, although MRC policy does not allow ‘biased’ reviewers to participate in the scoring of applicants they are associated with, this rule was insufficient, as the ‘neutral’ committee members compensated by raising their score when judging applications affiliated with one of their peers.” (Wenneras og Wold 1997: 342)

Flere studier rapporterer også en “*language bias*” i peer-review-systemet (f.eks. Herrera 1999), dvs. at forfattere som ikke har engelsk som morsmål blir forbigått.

I følge den empiriske undersøkelsen til Nylenna (1994) er strengheten til referee- vurderinger avhengig av *alderen* til de sakkyndige konsulentene. Som tommelfingerregel gjelder: jo yngre desto strengere. Den strengere vurderingen til unge sakkyndige konsulenter skyldes deres ambisiøse og paniske redsel for å overse feil og et press i retning av å gjøre godt inntrykk på tidsskriftsutgivere.

Relasjoner mellom utgivere og sakkyndige konsulenter synes å være preget av *ikke-kommunikasjon og likegyldig uinteresse* for de bedømte artiklene. Ifølge Glogoff (1988) har halvparten av alle sakkyndige konsulenter (n=121 fra 31 vitenskapelige tidsskrifter innen biblioteksvitenskap) ingen formelle vurderingskriterier. De er i overveiende grad ikke informert om hvorvidt deres kommentarer blir videresendt fra utgiver til forfatter. Nesten trefjedredelen av de sakkyndige konsulentene ble heller ikke informert om den endelige “skjebnen” til manus som de har vurdert.

Kvaliteten av publiserte artikler, dvs. slike artikler som klarte seg gjennom alle konsulentvurderingshindre har blitt kritisk undersøkt i en rekke innholdsanalytiske studier på grunnlag av et større antall cases. I en undersøkelse av artikler i velrenommerte tyskspråklige sosiologiske tidsskrifter hevder Meinfeld (1985) at det ikke er mulig å gjennomføre en kritisk intersubjektiv kontroll av tidsskriftsartikler. *Relevante opplysninger mangler i nesten alle kontrollerte bidrag*, noe de sakkyndige konsulentene altså overhodet ikke kan ha lagt vekt på. Den hyppige unnskyldningen “plassmangel” er bare en bortforklaring fordi det trengs så lite plass til relevante opplysninger (f.eks. avslagskvoter) i småskrift.

#### *4.3. Rundspørring til sakkyndige konsulenter*

Undersøkelser av de involverte og aktive forskernes holdninger viser i første rekke negative vurderinger av peer-review; først og fremst med tanke på rasjonalitet, effektivitet, objektivitet og innovasjonsglede. Disse negative vurderinger gjelder både for avgjørelsene til utgivere og hele systemet med sakkyndige konsulenter.

Bradly (1981) spurte 672 universitetsprofessorer om deres erfaring med sakkyndige konsulenter og sakkyndige uttalelser. De klaget over *konformitetspress* fra de sakkyndige konsulentene basert på idiosynkrasier (dvs. meget egenrådige eller eksentriske ønsker). Videre klaget professorene over upassende kritikk, manglende faglig kompetanse, konsentrasjon om trivielle ting, over at de hadde opplevd underlødige behandling av seg selv som forfatter, og over en likegyldig, uoppmerksom lesning av manuskripter. Det virket snarere som at de sakkyndige uttalelsene ble skrevet for å imponere tidsskriftutgivere enn for å øve en saklig kritikk.

Sakkyndige konsulenter blir i hele litteraturen karakterisert som *motstandere både av teoretiske og metodiske innovasjoner*. Armstrong (1997) trekker følgende konklusjon av det empiriske resultat av en

omfattende undersøkelse (hans litteraturbasis: 68 artikler, publisert siden 1975) vedrørende fagkonsulenter i tidsskrifter:

”Current procedures (...) seem to discourage scientific advancement, especially important innovations, because findings that conflict with current beliefs are often judged to have defects.” (Armstrong 1997)

På samme måte konstaterer Yamazaki (1992, 1995) (Japan, farmasiforskning) faren for konservatisme.

Kritikere av den (ofte ensidige) anonymiteten til referee-systemet frykter at disse mekanismene fungerer som en usynlighetskappe for de oftest anonyme konsulentene. Sakkyndige konsulenter kjenner jo ofte identiteten til artikkelinnsendere, men ikke omvendt. Og manusforfatterne får ofte ikke en gang se de sakkyndige uttalelsene slik at de får mulighet til å rette på feil og feilslutninger som de sakkyndige konsulentene kan ha gjort.

I sum kan klagene til de involverte sammenfattes på følgende måte: *Det kjente* (dvs. det som en sakkyndig konsulent kjenner pga. egen kompetanse og kjennskap til den vitenskapelige utviklingen) *blir fremmet, det mindre kjente blir hemmet*. Noe som det ofte blir klaget over i litteraturen og i den uformelle kommunikasjonen er at kjente og velbrukte temaer, modeller og teorier har bedre sjanser ved en sakkyndig vurdering da sakkyndige konsulenter er fortrolige med disse, eller i det minste kjenner til dem. For innovative bidrag eller bidrag som står lengre fra “mainstream” er derimot faren særdeles høy for at innsendte søknader eller manuskripter blir avvist av de sakkyndige konsulenter – på grunn av deres egen uvitenhet. Som det vanlig (drepende) standardargumentet benyttes i denne sammenhengen ofte den stereotype formuleringen at det ikke er ‘henvist til aktuell diskusjon og til annen forskning (som bestandig blir likestilt med kunnskapsnivået til de sakkyndige konsulentene). Det innebærer at det ikke forventes at det bedømte bidraget kunne være et vesentlig

fremskritt for erkjennelsen. Her lar også den dårligere behandlingen, dvs. *høyere avvisningsrater* for prosjekter eller manuskripter som forfølger *interdisiplinære* idéer og framgangsmåter seg innordne. Det finnes empirisk belegg for denne påstanden<sup>7</sup>.

Beinharde peer-review-tilhengere vil som motargument til dette si at slike negative vurderinger kommer fra skuffede og mislykkede vitenskapsmenn. Men tyske forskere fra mellomsjiktet fører sine foreløpige publikasjonsuksesser tilbake til “flaks og sosiale forbindelser” i stede for til evne og anstrengelse (Borkenau og Kammer 1988). På motsatt side forteller nobelprisvinnere om en urettferdig overøsing av belønninger og begunstigelse, noe som Harriet Zuckermann (1977) har forsket på.

#### *4.4. Uoverensstemmelser mellom sakkyndige konsulentvurderinger og i forhold til sitathyppighet*

Ofte blir det kritisert eller påvist *manglende overensstemmelse* mellom de sakkyndige konsulentene. Cicchetti (1997) fokuserer f.eks. på den lave reliabiliteten i peer-review-prosessene. Utgivere må ofte treffe en avgjørelse på grunnlag av motstridende konsulentanbefalinger. Mot dette erklærer de få offensive forsvarere av det etablerte review-systemet at de korrelerer høyt – fra 0,2 til 0,3 (jf. Lindsey 1991). Eller de fortolker den manglende overensstemmelse mellom referee-vurderinger som et bevis på det paradigmatisk mangfold blant et hvert tidsskrifts sakkyndige konsulenter. Men enten de vil eller ikke impliserer de med dette deres egen mangel på paradigme-overgripende kvalitetskriterier.

---

<sup>7</sup> Et positivt unntak burde her være å fremheve interdisiplinariteten ved EU-prosjekter.



De som forsvarer status quo sammenligner vurderingene til referees med *sitathypighet*. Men dette er meget omstridt, og kan være gjenstand for nøyaktig samme bias: allerede etablerte teorier, metoder, vitenskapsmenn og institusjoner blir belønnet gjennom "impact"-verdier (f.eks. gjennom å bli tatt opp i en databank) og dermed blir deres virkning styrket. Innovative outsiders som foreløpig har en lav status, blir derimot straffet (jf. Fröhlich 1999a).

## 5. Motstanden mot objektivisering av en selv

Til tross for alle negative funn angående deres kompetanse, gjøres sakkyndige konsulenter delvis urettferdig til sydebukker. Det som er alt for lite i offentlighetens og kritikkens lys, er arkan- eller hemmeligholdspraksisen og vilkårligheten til utgivere som iblant leder sitt tidsskrift som en diktatorisk enmannsbedrift. Det som nesten ikke er undersøkt er: på hvilken måte blir man utgiver og hvordan fungerer avgjørelsesprosessene blant utgivere? I fag med et svakt paradigme blir skjebnen til et manuskript i praksis avgjort idet manuskriptet tildeles en sakkyndig konsulent som tilhører en kjent skoleretning. I litteraturen og på internett finnes det tallrike (delvis ganske negative) anekdotiske informasjoner om vurderingsmetodene og de avgjørende prosessene som tar for seg innsendelsen av prosjektforslag eller stipendsøknader til EU, til departementer, til akademier eller til private stiftelser. Men det finnes nesten ikke noen empiriske undersøkelser som er solide og kritiske.

Man må se det som negativt at utgivere, de som disponerer forskningsmidler og de sakkyndige konsulentene helt åpenbart ikke kjenner til eller ignorerer de tallrike, empiriske undersøkelsene til vitenskapsforskningen (som dessverre ofte viser negative resultater). De som selv skal være objektive viser som kjent ofte den største motstand mot *egen* objektivisering. Det gis følgelig heller ikke noe informasjon til eller skolering av sakkyndige konsulenter, eller i det hele tatt drevet noen systematisk kompetanseheving av dem. Det virker som at det sees som en slags naturlig gave: vitenskapsmenn som har kommet seg frem i et trang spesialdisiplin skal uten videre kunne vurdere alt objektivt.

Til sist skal det enda en gang fremheves at det finnes *meget forskjellige "bedriftskulturer"* fra en vitenskapelig disiplin til en annen. Innenfor hver av disiplinene er det slik at vanene til tidsskrifter, forlag, organisasjoner som fremmer forskning osv. skiller seg betydelig fra hverandre. Ta to eksempler: I noen fag ( f.eks.

sosialøkonomi) er det vanlig at forfattere får innsyn i de sakkyndige konsulentenes uttalelser, i andre fag er det derimot ikke slik. En anonymisering av forfattere, som er ett reformforslag, er nesten umulig innenfor store deler av fysikk eller “Big science”. Dette har sammenheng med at det i visse spesialområder av disse fagene på verdensbasis finnes så få forskningsgrupper hva enten de konkurrerer eller samarbeider med hverandre. Det betyr at forskergrupper kan bli identifisert på grunn av det tunge og svært kostbare utstyr, eller av metodene de bruker. Slike begrensninger burde det tas hensyn til ved følgende reformforslag.

## 6. Reformforslag

Reformforslagene i litteratur og på internett anbefaler *mer transparens*. De vil ha innsikt i de fullstendige sakkyndige uttalelsene, og de krever forpliktende og offentlig tilgjengelige kriterier for sakkyndige konsulenter, men de krever sjelden en regelmessig vurdering av de sakkyndige konsulentene selv.

Mer radikale kritikere krever et *åpent system* istedenfor den hittil vanlige praksisen med hemmelighold. Dette inkluderer å lagre manus, sakkyndige uttalelser og kritikker på offentlig tilgjengelige internett.<sup>8</sup> I hvert fall vil man gjøre slutt på anonymiteten til de sakkyndige konsulentene. Anonymiteten innskrenker åpen diskusjon og kritikk pga. redsel for hevn fra dem som blir kritisert, og som hittil blir vernet av anonymiteten.

Andre forslår derimot å gjennomføre en allmenn *dobbelblind-vurdering*. Dette vil si. at også identiteten til de som blir vurdert blir

---

<sup>8</sup> For drøftelse av “electronic peer review conferences”, se Peters (1996), for praksiserfaringer av “interaktive publikasjoner” på internett, se Harnad (1996).

hemmeligholdt under vurderingsprosessen og kommentarer i samme hefte. (Se f. eks. Peters og Ceci 1982) Sammenlignet med enkelblind-metoder klarer dobbelblind-metoder seg vanligvis bedre og bidrar dermed til å begrense det at vitenskapsmenn fra bedre stilte (“near-top-ranked”) institusjoner blir foretrukket eller at vitenskapsmenn fra ikke-akademiske institusjoner blir forbigått (se f. eks. Blank 1967).

Det kan være verdifullt å forestille seg et vitenskapelig publikasjonssystem *uten* peer review, dvs. uten offisiell sensur i begynnelsen av publikasjonsprosessen, som et *tankeeksperiment*:

Det er nettopp tilliten til de “kvalitetskontrollerte” publikasjonene som leder til hemningsløs sitering av og tiltro til disse. Slik oppmuntres det til å publisere artikler som er laget for raskt og for overfladisk. For øyeblikket dominerer åpenbart denne avlastende tillitsstrategien, fordi den reduserer kompleksitet (Luhmann). Vi vet fra mange undersøkelser at forfattere ofte ikke har lest originallitteraturen som de siterer. De stoler på at litteraturen allerede har blitt kontrollert av noen kompetente “gate keepers” – konsulentene.

Uten peer-review-sikkerhet ville det på nytt kreve større egenansvar av forfattere hva angår kritisk kontroll av tekster de siterer. Uten peer-review ville større ressurser frigjøres til kontroll og vurdering av slik sitering, og det fullstendig ‘overopphetete’ publikasjons-maskineri kunne ‘avkjøle’ seg litt. Det er feil å vente at tidsskriftssektoren, og heller ikke den vitenskapelige sektoren av internett (dvs. internett-aktiviteter som vitenskapelige selskaper står for) vil utvikle seg til en tumbleplass for dilettanter. Den konsekvente digitale vitenskapelige kommunikasjonen ville gjøre det lettere å finne tekster i søkemaskiner eller datalinks, og gjøre det lettere for leseren å skrive kritiske kommentarer. Faren for å blamere seg offentlig ville i så fall være vesentlig høyere enn hittil – pinligheter blir i dag skjult bak kulissene til institusjonene og til tidsskriftene.

Riktignok løser man ikke det store problemet med fordelingen av forskningsmidler og stipend skal foregå. Funksjonsdyktigheten og prestisjen til fysikkens preprint-servere henger nok sammen med at man innen disse “big-science-disipliner” som nevnt allerede har et høyt hinder i og med at man må tilhøre en forskningsgruppe som har tilgang til svært kostbart utstyr, eller på annen måte å få finansiert forskningen. I tekstvitenskapene derimot er det fremdeles ofte tilstrekkelig med papir og blyant som forskningsmiddel (ved siden av tilgangen til et vitenskapelig bibliotek).

Hvis man nå vil beholde peer-review-systemet, ville følgende reformer være å anbefale:

- I Tyskland blir sivile rettssaker tilfeldig tildelt dommere ved hjelp av en datamaskin for å unngå den skjevhet som kunne oppstå ved en personlig fordeling. Dette kunne man også overføre til artikkelvurderingssystemet. Tildelingen av sakkyndige konsulenter bør foregå *tilfeldig*, dvs. at alle som arbeider som sakkyndige konsulenter systematisk blir benyttet – på den måten utelukkes vilkårligheten og makt-konsentrasjonen hos noen få “old boys”. I fag med svakt paradigme kunne dette foregå sortert etter paradigme, idet man som et minimum alltid finner en paradigmeintern og en paradigme fremmed sakkyndig konsulent.
- Det burde være en selvfølge å påskynde kommunikasjonen via internett mellom forfattere, redaksjoner, sakkyndige konsulenter o.l.
- Det beste ville være en vurdering som er ’tre dobbelt blind’, altså at identiteten til manusforfattere holdes skjult ikke bare for de sakkyndige konsulentene, men også for utgivere eller for medlemmer av forskningsråd o.l.
- Alle som blir vurdert bør få vite hele innholdet av samtlige sakkyndige uttalelser.

- Ankeretten for innsendere bør være standard. Først og fremst må manusforfattere få muligheten til å korrigere sakkyndige konsulenter når de gjør klare feil, eller når det oppstår misforståelser.

- Alle ytringer, dokumenter og prosedyrer bør bli *offentliggjort* etter bedømmelsen når den avsluttes med suksess – og hvis den avsluttes negativt, når søkeren er innforstått med det. Kritikken til de sakkyndige konsulentene og til utgivere bør også bli offentliggjort slik at hele det vitenskapelige samfunn kan ha nytte av kritikken. Det potensielt uendelige digitale rommet gjør dette mulig.

- Hittil har arbeidet til sakkyndige konsulenter som regel vært et ulønnet tillitsverv. Dette er sikkert en årsak, noe det ofte klages over, til at det er så vanskelig å finne sakkyndige konsulenter og til at de legger så lite tid og arbeid i å skrive vurderingene. Det finnes empiriske belegg for dette. Arbeidet til sakkyndige konsulenter bør derfor bli passende avlønnet.

- *Kompetansen* til utgivere og sakkyndige konsulenter bør bedres systematisk. Dette bør senest begynne etter hovedfag eller på doktorkursnivå. De kritiske resultatene til vitenskapsforskningen bør tas til etterretning, og ses som en anledning til å gjennomføre reformer. Generelt bør en kritisk evne og en kritisk praksis systematisk fremmes allerede i undervisningen, men også gjennom det vitenskapelige belønningssystem. Slik kunne man også øke den vitenskapsteoretiske og forskningsetiske sensibilitet og refleksjons-evne.

- Praksis blant utgivere og til sakkyndige konsulenter har hittil virket *ganske gammeldags og håndverksmessig* – delvis med fatale følger. Ved å lete fram “abstracts” fra databaser, noen minutters arbeid, hadde man f.eks. lett funnet ut at tvillingforskeren Burt brukte akkurat den samme korrelasjonskoeffisienten – med tre desimalers nøyaktighet – i tre publikasjoner som omhandlet angivelig forskjellige stikkprøver for en- og toeggete tvillinger (Fröhlich 2001) Lik visualisering i forskjellige artikler, som angivelig var relatert til

forskjellige undersøkelser, ble først oppdaget etter flere års etterforskning av nye svindelskandaler. Disse ville blitt oppdaget raskt og uten store kostnader ved å finlese alle artikler på “referee”ens nettsteder. Plagiatovervåknings-programmer eller enkle “related documents” funksjoner (dvs. slike som finner frem dokumenter med identisk eller liknende siteringer) kunne vært satt i gang for lenge siden. “Moderniseringen” av utgiveres og sakkyndige konsulenters gammeldagse kontrollpraksis bør derfor være en meget viktig reformsak. Kompetansen til utgivere, redaksjoner og sakkyndige konsulenter bør utvikles når det gjelder deres omgang med databaser og med metoder fra informasjonsvitenskapen. I det minste burde kvalifiserte forskere på informasjon og vitenskap bli trukket inn.

- De rettslige forutsetninger for å offentliggjøre dokumenter fra stiftelser og andre som finansierer empirisk forskning burde forbedres eller skaffes til veie.

Utover dette foreslår jeg i en viss forstand å gi *tilfeldigheten* større rom:

- Ligningskontoret krever ikke alle vedlegg til selvangivelsen framlagt på forhånd når de beregner skatten. Men de skattepliktige må regne med at deres sak blir undersøkt nærmere og at de da må legge frem samtlige bilag. På lignende måte ville det forbedre kontrollen med slurv og bedrageri hvis de som innleverer sine prosjekter og manuskripter måtte regne med en risiko på 5-10% for å bli trukket tilfeldig ut og måtte legge frem sine rådata og sine laboratorie- eller felt-dagbøker (ifølge undersøkelser har hittil mesteparten av disse blitt “forlagt, brent eller mistet”, jfr. f.eks. Wolins 1962). De sakkyndige konsulentenes arbeid kunne få økt nøyaktighet og kvalitet hvis alle visste at det var 5-10% sannsynlighet for at en annen sakkyndig konsulent ville vurdere deres sakkyndige uttalelser, og at dette kunne føre til negative konsekvenser. *Systematiske stikkprøvekontroller* ville først og fremst roe ned de forsvarerne av

det nåværende systemet som avviser reformer ved å henvise til at det ville øke tidsbruk og kostnader sterkt.

- Man kunne gjøre en dyd av klagen om at peer-review-systemet er et "lottospill": Hvorfor kunne man da ikke bruke en liten del av forskningsmidlene til et *innovasjons-lotteri*, altså fordele noen midler ved loddtrekning?

- For dyrt? Forfattere og institusjoner som skyver sine evalueringsproblemer over på publikasjonsarenaen (noe som fører med seg at kritikk- og kommunikasjonsfunksjonen lider), bør overta kostnadene for dette. Ved omstilling til digital vitenskapelige kommunikasjon ville dessuten øvrige kostnader bli sterkt redusert. Fordi publisering hjelper forfattere og deres institusjoner mer enn leserne (interesserte vitenskapsmenn får likevel tak i relevante informasjoner direkte i form av uformelt bytte), burde de som har den egentlige nytten av dette betale. Ikke minst belønner det aktuelle systemet nettopp skribenter som om og om igjen sender inn gratis (til andre tidsskrifter): Jo "dårligere" en artikkel er, desto flere lesere (dvs. utgivere og sakkyndige konsulenter) finner den.

Penger til "evaluerings" blir likevel brukt. Private markedsforskningsinstitutter som arbeider med oppdrag fra departementer åpner f.eks. auditoriedører og vurderer med et blikk hvor mange studenter (som ofte tilhører en helt annen forelesning enn den offisielt registrerte på grunn av romforandringer) som er til stede. Eller de spør angivelige "opinion leaders" om hvilke vitenskaper som skal få innvilget rett til å eksistere videre ved universitetet. Den såkalte evaluerende scientometriens flittige virksomhet (som først og fremst tar for seg de enkelte skoleretninger og disipliner, og tidsskrifttyper som systematisk skorer høyt på ISIs to-års-impact-faktor) fører til kunstprodukter og til en kortpustet og kortsiktig salami-publikasjonstaktikk (Fröhlich 1999a).



Sammenlignet med slike fremgangsmåter er faglig kritikk nødvendig og meningsfylt, men da i form av den mest mulig kompetente *vurderingen av intellektuelle prestasjoner*. Egentlig skulle dette være hverdagspraksis for *alle* forskere<sup>9</sup>, og ikke en delegering til noen få. Ved all evaluerings- og reformiver bør vi ikke glemme det som så ofte er påvist: det vitenskapelige “establishments” *motstand mot innovasjon*, noe også vitenskapshistorien viser. Det ville sikkert få fatale følger dersom vi fikk et fullkomment teknokratisk vurderingssystem, med andre ord et hierarki: For ikke bare innen jordbruk viser det seg at monokulturer ikke er motstandsdyktige. *Åpenhet og pluralisme* i vitenskapelige arenaer bør derfor være det høyeste bud.

---

<sup>9</sup> Alt etter problemstillingen skulle det være en selvfølge at øvrige vedkommende og kompetente bli tatt med i forhold til en deltakende utforming.

## ***Litteratur***

- Abrmowitz, S. I m.fl. (1975): "Publish or politic: Referee bias in manuscript review" I *Journal of Applied Social Psychology* 5 (3), 187-200
- Ammstrong, J. S. (1997): "Peer Review for Journals: Evidence on Quality Control, Fairness, and Innovation" I *Science and Engineering Ethics* 3 (1), 63-84
- Baxt, W. G. m. fl. (1998): "Who reviews the Reviewers? Feasibility of using a fiction manuscript to evaluate peer reviewer performance" I *Annals of emergency medicine* 32 (3,1), 310-317
- Blank, R. M. (1967): "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence" *American Economic Review* 8, 1041-1067
- Borkenau, P. og Kammer, D. (1988): "Publizieren oder resignieren? Subjektive berufliche Perspektiven und Bewältigungsstrategien des Mittelbaus der Universität Bielefeld" I *Zeitschrift für Soziologi* 17 (1), 72-79
- Bradley, J. V. (1981): "Pernicious publication practices" I *Bulletin of the Psychonomics Society* 18 (1), 31-34
- Cicchetti, D. V. (1997): "Referees; Editors, and Publication Practices: Improving the Reliability – and Usefulness of the Peer review system" *Science and Engineering Ethics* 3 (1), 51-62
- Cole, S. m.fl. (1981): "Chance and Consensus in Peer Review" *Science* 214,881-886
- Eldredge, J. D. (1997): "Identifying peer-reviewed journals in clinical medicine" *Bulletin of the Medical Library Association* 85 (4); 418-422
- Finetti, M. og Himmelrath, A. (1999): *Der Sündenfall. Betrug und Fälschung in der deutschen Wissenschaft* Stgt. etc.

- Fröhlich, G. (1998): "Optimale Informationsvorenthaltung als Strategem wissenschaftlicher Kommunikation" I Harald H. Zimmermann og Volker Schramm (red.), *Knowledge Management und Kommunikationssysteme*. Konstanz, 535-549. Online: <<http://www.agmb.de/mbi/8/mb8.pdf>>
- Fröhlich, G. (1999a): "Das Messen des leicht Messbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie?" I Becker, J. og Goehring, W. (red.): *Kommunikation statt Markt. Gesellschaft für Mathematik und Datenverarbeitung Sankt Augustin, GMD Report 61*. Online s.a.: <<http://info.uibk.ac.at/sci-org/voeb/om65.html#gf>>
- Fröhlich, G. (1999b): *Von den Tempelwächtern der Wissenschaft. der Standard – Album Wissenschaft (bm:wv; FWF) 23. 10. 1999, W1*
- Fröhlich, G. (2000): "Wissenschaftsforschung: Die theoretisch-empirische Erforschung der `Medien und Institutionen des Wissens" I AG Kulturwissenschaften Graz (red.): *Kulturwissenschaften in Österreich*. Graz, 28-31
- Fröhlich, G. (2001): "Betrug und Täuschung in den Sozial- und Kulturwissenschaften" I Hug, T. (red.), *Wie kommt die Wissenschaft zu ihrem Wissen?* Hohengehren og Baltmannsweiler: bind 4/CD-rom2: *Einführung in die Wissenschaftstheorie und Wissenschaftsforschung der Sozial- und Kulturwissenschaften* 261-273
- Fröhlich, G. (2002): "Anonyme Krikik. Peer review auf dem Prüfstand der empirisch-theoretischen Wissenschaftsforschung" I Eveline Pipp (2002): *Drehscheibe E-Mittleuropa. Information: Produzenten, Vermittler, Nutzer. Die gemeinsame Zukunft* Wien: Phoibos Verlag, 2002, Biblos-Schriften 173,129-146
- Fröhlich, G. (2002a): "Verein.wissenschaft: Entstehung und Funktion wissenschaftlicher Gesellschaften" I Kammerhofer\_Aggermann U. (red.): *Ehrenamt und Leidenschaft* Salzburg, 255-278
- Fröhlich, G. (2002b): Peer Review: Contra. Forschung und Lehre 9 (2), 313

- Fröhlich, G. (2002c): "Einschlagende Neuigkeiten. Der 'Impact Factor' entscheidet über das gewicht von Artikeln und Forscherinnen" *Der Standard, Wissenschaft*, 9. 4. 2002, online: <<http://DerStandard.at/?id=917598>>
- Glogoff, S. (1988): "Reviewing the Gatekeepers: A Survey of Referees of Library Journals. American Society for Information" I *Science Journal* 39 (6), 400-407
- Harnad, S. (1996) "Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals" I Peek, R. og Newby, G. (red) *Scholarly Publishing: The Electronic Frontier* Cambridge/Mass., 103-118
- Herrera, A. J. (1999): "Language bias discredits the peer-review system" I *Nature* 397 (6719), 467
- Jasanoff, S. (1985): "Peer Review In The Regulatory Process. Science" I *Technology and Human Values* 10, 20-32
- Lindsey, D. (1991): "Precision in the Manuscript Review Process Hagens and Herting Revisited" I *Scientometrics* 22 (2), 313-325
- Lowen, R. S. (1977): *Creating the Cold War University. The Transformation of Stanford Berkley etc.*
- Meinefeld, W. (1985): "Die Rezeption empirischer Forschungsergebnisse – eine Frage von Treu und Glaube? Resultate einer Analyse von Zeitschriftenartikeln" I *Zeitschrift für Soziologie* 14 (4), 297-314
- Munthe, C og Welin, S. (1996): "The Morality of Scientific Openness" I *Science and Engineering Ethics* 2 (4), 411-428
- Nylenna, M. (1994): "Multiple blinded review of the same two manuscripts: Effects of referee characteristics and publication Language" I *JAMA – Journal of the American Medical Association* 272 (2), 149-151
- Peters, D. P. og Ceci, S. J. (1982): "Peer review practices of psychological journals: The fate of published articles, submitted again" *Behavioral and Brain Science* 5, 187-195

- Peters, J. (1996): "The Hundred Years War started today: an exploration of electronic peer review" I *Management Decision* 34 (1), 54-59
- Savage, J. D. (1999): *Funding science in America: Congress, universities, and the politics of the academic pork barrel* Cambridge etc.
- Wenneras, C. og Wold, A. (1997): "Nepotism and sexism in peer-review" I *Nature* 387. 341-343
- Wolins, L. (1962): "Responsibility for Raw Data" I *American Psychologist* 17, 657-658
- Yamazaki, S. (1992): "Production and transfer of information: the increasing volume of information and the referee system" I *Pharmaceutical Library Bulletin* 37 (3), 187-190
- Yamazaki, S. (1995): "Refereeing system of 29 life science journals preferred by Japanese scientists" I *Scientometrics* 33 (1), 123- 129
- Zuckermann, H. (1977): *Scientific elite. Nobel Laureates in the United States* N. Y. / London