

Talität versus Qualität. 1. Teil: Grundlagen einer Kritik überholter Impact-Begriffe¹

Thomas Stephenson

Kurzzusammenfassung

Als 1927 die Idee der quantitativen Erfassung der „Citedness“ aufgebracht, durch Garfield 1955 zur Bestimmung eines entsprechenden Zitationsindex weiterentwickelt und dann als „Impact Faktor“ in den Wissenschaftsdiskurs eingebracht wurde, ahnte noch niemand, zu welchen Auswüchsen diese Ideen führen würden. 2013 unterschrieben auch österreichische WissenschaftlerInnen bzw. Institutionen eine internationale Deklaration, in der (wieder einmal) scharfe und essentielle Kritik an diesem „goldenen Kalb“ des Impact-Faktors formuliert wurde. In diesem ersten Teil des zweiteiligen Artikels werden die Grundlagen dieser Kritik dargelegt, diskutiert, an Beispielen illustriert und einige Alternativen skizziert.

Abstract

In 1927 the idea of a quantitative assessment of "Citedness" came up and was enhanced by Garfield in 1955. The aim was to determine an appropriate citation index. Further on the "impact factor" was introduced into the scientific discourse. At that time no one could expect the excessive role this idea would lead to. In 2013 Austrian scientists and institutions signed an international declaration in which (once again) this „golden calf“ Impact Factor was criticized essentially. This is the first part of a two-part article, In this first part basic principles of critical discussion are presented, illustrated by examples and alternatives outlined.

¹ Dieser Artikel ist eine überarbeitete und erweiterte Version des Vortrages „Projekt Zitationsforschung: Die „Impact“-Faktoren psychoanalytisch-pädagogischer Texte“, gehalten im Rahmen des Internationalen DissertantInnenkolloquiums Psychoanalyse am 13.9. 2008 in Wien

Einleitung

„Hier steh´ ich nun, ich armer Tor, und bin...“

Haben Sie´s erkannt? Und können Sie dieses Zitat aus dem Stand weiterzitieren? Goethes Werke, und speziell sein „Faust“ (I und II) gehören wohl zu jenen geistigen Produkten, die in einem fiktiven „Ranking“ aller jener Texte, aus denen in anderen Texten zitiert wird, in den „top ten“ (oder doch zumindest „top hundred“) anzusiedeln wäre. Wenn es nun aber konkret darum ginge, in wissenschaftlich exakter Weise zu erfassen, *inwiefern und inwieweit, also in welcher Weise und in welchem Ausmaß* Goethe auf die schriftlichen Produktionen anderer Menschen Einfluss genommen hat, würde sich sofort eine Flut von Fragen auftun, deren Beantwortung viele ForscherInnenleben füllen könnte. Viele dieser Fragen müssten sich auf etwas richten, was nur teilweise mit dem Begriff der „Qualität“ „dingfest“ gemacht werden könnte. Gleichwohl darf angenommen werden, dass meine Eingangsvermutung bei den meisten LeserInnen auf Zustimmung stoßen wird, nämlich die Vermutung, dass Goethes Werke zu jenen gehören, die insgesamt in der menschlichen Textlandschaft seither besonders häufig herangezogen werden, um als Teil anderer Texte ebendiesen einen Zuwachs an „Wert“, also an Reputation, an Glaubwürdigkeit, an Anerkennung zu verleihen.

Was in einem solchen fiktiven Ranking für ein „Gesamtergebnis“ gelten könnte, wird aber wohl kaum für jede einzelne Untergruppe von Textproduzentinnen gelten können. Im Gegenteil ist anzunehmen, dass die Zahl der Zitate aus diesen Werken in den verschiedenen Teilen der Gesamt-Textlandschaft bisweilen enorme Unterschiede zeigen wird. Teilen wir diese fiktive Gesamt-Textlandschaft zunächst

in zwei Gruppen, von denen die eine jene der „wissenschaftlichen“ Texte sei, und die andere „der Rest“, und betrachten nun die erste Gruppe genauer, so bieten sich mannigfaltigste Differenzierungsmöglichkeiten an. Würden wir nun jene der Unterteilung der Wissenschaftslandschaft in „Disziplinen“ heranziehen und dabei z.B. „Kernphysik“ und „Germanistik“ einander gegenüberstellen, so ist anzunehmen, dass eine quantitative Erfassung des „Einflusses“ der Goetheschen Werke in den beiden Disziplinen sehr unterschiedliche Ergebnisse zeitigen würde. In jedem Fall allerdings hätten wir zwei Seiten aller dieser „AutorInnen-spezifischen Zitationsphänomene“ zu berücksichtigen: die Seite des/der Zitierten und jene des/der Zitierenden!

Beide Seiten müssen bei den grundlegenden Fragen, die hier zu stellen sind, bedacht werden:

- Was macht die „Qualität des zitierten Textes“ und was die „Qualität des zitierenden Textes“ aus?
- Wie lässt sich das Ausmaß und die Art der *Wirkmächtigkeit*, die vom Zitierten auf die Zitierenden ausgeübt wird, erfassen?
- Und inwiefern hängen beide - *Qualität* und *Wirkmächtigkeit* - zusammen?

In diesem Artikel werden bestimmte Teilaspekte dieser Fragen näher betrachtet und der Autor wird den Versuch unternehmen eine spezielle Herangehensweise an die Lösung einiger substanzwissenschaftlicher und methodologischer Probleme zu skizzieren, die als Basis für ein Forschungsprojekt herangezogen wird, ein Forschungsprojekt, das bereits im Planungsstadium als interdisziplinäres Unternehmen angelegt ist und das einen Beitrag leisten möchte für eine differenzierte Erfassung jenes faszinierenden Phänomens, das im

Folgenden als „Fremdtext-Eigentext-Relation“ in den Be-Griff genommen werden soll.

Der Impact-Faktor – ein Maß wofür?

Das von Garfield beforschte Gebiet war wissenschaftspolitisch von Anfang an nicht nur mit höchster Brisanz² versehen, sondern auch dicht an dicht mit den Tretminen grundlegender Missverständnisse besetzt. Garfields Idee, die im Wesentlichen darin besteht, zu erfassen, wo und wie oft eine bestimmte wissenschaftliche Erkenntnis zitiert wird, war zunächst lediglich darauf ausgerichtet, eine Möglichkeit für die Bestimmung der Wirkmächtigkeit zu schaffen, die jene Texte, in denen die zitierten Erkenntnisse zu finden sind, bezüglich des einschlägigen Diskurses der scientific community erreichen. „Das bedeutet allerdings nicht a priori, dass viel zitierte Arbeiten objektiv eine größere Qualität aufweisen...“ (Jokic; Ball 2006, 147), wenn auch in der Öffentlichkeit ein solcher Trugschluss gang und gäbe ist (vgl. Nicolaisen 2002).

Diese Tatsache, dass die Zahl der Zitate, die auf ein Werk Bezug nehmen, nicht unmittelbar etwas über die Qualität resp. Bedeutung dieses Werks aussagen, hat Garfield anlässlich einer Rückschau auf 50 Jahre Erfahrung an sich selbst festgestellt:

„About 30 years ago we began the series called ‘This week’s Citation Classic’. Over the next 15 years, we asked thousands of authors to write commentaries on these highly cited papers (www.citationclassics.org). So I was surprised to be asked to do the same for my 1955 paper in Science. As a confirmed citationist, I must point out that it is not my most

cited work. It is my 1972 paper in Science, on using citation analysis to evaluate journals, which has attracted much more attention, although the 1955 paper is far more significant. In that sense, I am like many other authors who feel that their most-cited work is not necessarily their best. My most-cited work is in fact my 1979 book Citation Indexing.“ (Garfield 2006, 1)

Wenngleich die Idee der quantitativen Erfassung der „Citedness“ schon 1927 von Gross und Gross entwickelt wurde, stellte Garfield selbst erst 1955 seine Grundgedanken zur Bestimmung eines entsprechenden Zitationsindexes vor. Vier Jahre später hob er gemeinsam mit Sher den „Impact Faktor“ aus der Taufe. Nachdem Raising zunächst einen „Index of Research Potenzial Realized“ vorgeschlagen hatte, der lediglich das Verhältnis der Zitanzahl zu der Artikelzahl einer Zeitschrift wiedergeben sollte, errechnet sich der Impact Faktor (im weiteren als „IF“ bezeichnet), der als Grundlage für die Eintragung in die Datenbank „Science Citation Index“ (im Folgenden „SCI“) dient, „als Bruchzahl und gibt die Anzahl der Zitationen im laufenden Jahr von allen Publikationen der betreffenden Zeitschrift in den zwei vorangegangenen Jahrgängen dividiert durch die Anzahl der Artikel dieser Zeitschrift im gleichen Zeitraum an.“ (Jokic; Ball 2006, 80f.)

Aktive und passive Formen von „Impact“ und „Qualität“

Der Unterschied zwischen diesen zwei Zugängen, die zunächst von Raising, dann von Garfield eröffnet wurden, ist von größerer Tragweite, als ein erster Blick ahnen lässt. Raisings Idee, eine Zeitschrift könnte durch eine Art „Nutzungsfaktor“ gekennzeichnet werden, richtet das Augenmerk bei der „Beurteilung“ einer einzelnen Zeitschrift(enausgabe) darauf,

² Garfield selbst formuliert die Brisanz 1998 noch sehr vorsichtig: „There is considerable emotion when citation analysis is discussed in connection with research evaluation.“ (Garfield 2006)

in welchem Ausmaß die dort abgebildete Forschung *selbst bereits erarbeitete Erkenntnisse aufgreift* und das darin enthaltene Potential nutzt. Durch die Tatsache, dass sich in einem solchen Fall das zu beurteilende Medium via Zitation als von bestehendem Wissen „beeindruckt“ zeigt, gleichzeitig aber ebenfalls via Zitation ein wesentliches Gütekriterium von Wissenschaft (nämlich „Anschluss an bestehendes Wissen“) einlöst, werden im Folgenden solche und ähnliche Ansätze zur Erfassung von „Wirkmächtigkeit“ und „Qualität“ als *„aktive Qualität durch passiven Impact“* bezeichnet. Denn der Grundgedanke solcher Quantifizierungen richtet sich einerseits auf die *Eigenaktivität des zu beurteilenden Mediums*, das via Zitation die eigene Forschung mit schon bestehendem Wissen bzw. aktuellen der scientific community bereits zur Verfügung stehenden Forschungsergebnissen unterlegt, befördert, anregt, argumentiert u.ä.m. – und damit die *„Qualität“* seiner Arbeiten zu steigern versucht. Andererseits *lässt es sich „passiv“ von anderen Artikeln beeinflussen* (übt also dabei noch nicht „aktiv“ Einfluss auf andere aus).

Im Gegensatz dazu nehmen Garfields und ähnlich gelagerte Ansätze etwas ganz anderes in den Blick: Die oben genannte Formel stellt eine Beziehung her zwischen dem, was ein Medium an Forschung der scientific community zur Verfügung gestellt hat und dem Ausmaß, *in dem die restliche scientific community das zur Verfügung gestellte aufgegriffen hat!* Und bezüglich der Frage wissenschaftlicher Qualität des zitierten Artikels wird nicht der Artikel selbst untersucht, sondern lediglich jene, die ihn zitieren. Derartige Zugänge zur Lösung des Problems der Qualität wissenschaftlicher Forschung werden im Folgenden *„passive Qualität durch aktiven Impact“* genannt. Denn hier muss das Publikationsmedi-

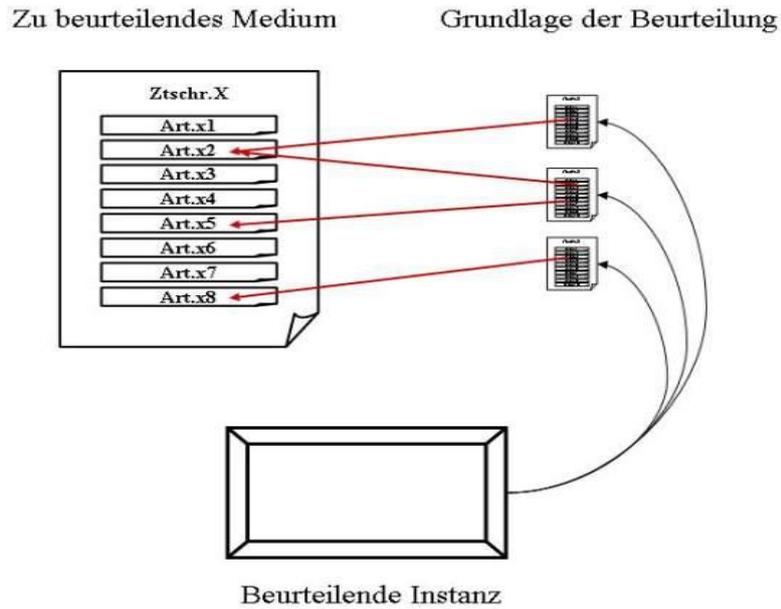
um „warten“, bis via Zitation eines anderen Mediums die eigene „Qualität“ „bestätigt“ wird.

Verzerrungen wissenschaftlicher Realitäten durch den IF

1. Verzerrungen durch die Kombination „passive Qualität - aktiver Impact“

Die oben eingeführte Unterteilung in „aktive“ und „passive“ Formen von „Wirkmächtigkeit“ und „Qualität“ impliziert teilweise bereits eine Kritik. Verdeutlichen wir uns zunächst noch einmal die Grundsituation³: Wie die folgenden Graphiken 1 und 2 veranschaulichen, ist die Grundlage der Beurteilung eines Publikationsorgans beim Garfield-IF eine „externe“: Die beurteilende Instanz zieht als Grundlage ihrer Beurteilung eines Publikationsorgans nicht Daten heran, die sich aus der Erfassung des Publikationsorgans selbst ergeben, sondern wendet sich an eine „externe“ Datenquelle, in diesem konkreten Fall an die Literaturverzeichnisse anderer Publikationsorgane.

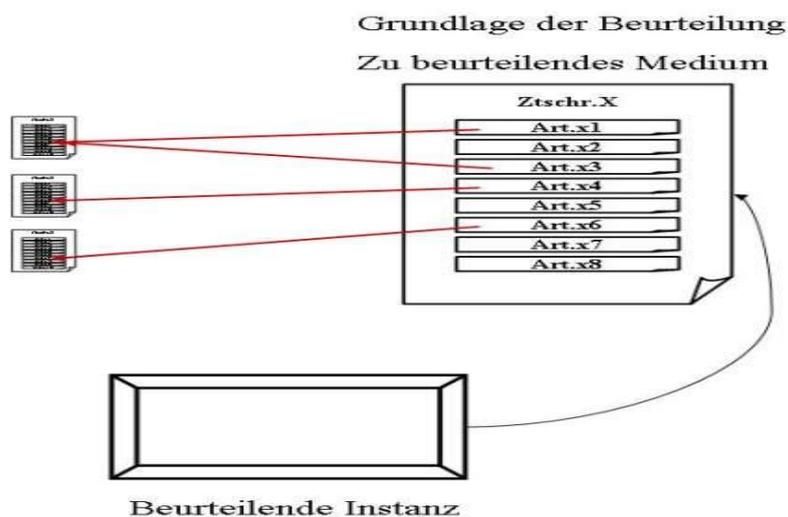
³ Um leichtere Erkennbarkeit und bessere Übersichtlichkeit (v.a. in den Graphiken) zu ermöglichen, werden im Folgenden einige Details der IF-Berechnung außer Acht gelassen, die am Prinzip der Erfassung nichts ändern, wie z.B. der Umstand, dass nicht ein sondern zwei Jahre für die Berechnung herangezogen werden.



Graphik 1: Aktiver Impact - Beurteilungsgrundlage 1

Im Gegensatz dazu gewinnen Verfahren, die sich zur Beurteilung wissenschaftlicher Publikationen wie Raisings Alternative mit „passi-

vem“ Impact, aber „aktiver“ Qualität beschäftigen, ihre Daten direkt aus dem zu beurteilenden Medium.



Graphik 2: Passiver Impact - Beurteilungsgrundlage 1

Würde man sich also - wie Raisings Idee nahe legt - darauf konzentrieren, inwiefern ein Artikel (und damit indirekt auch das jeweilige Publikationsorgan) sich von anderen Artikeln („passiv“) beeinflussen lässt (und dadurch („aktiv“) Anschluss an das gemeinsam geteilte Wissen der scientific community sucht und findet), und damit erheben, *ob und in welchem Ausmaß dieser Artikel (bzw. dieses Publikationsorgan) an aktuelle Forschungsergebnisse anschließt*, hätte man tatsächlich einen Aspekt wissenschaftliche Qualität miterfasst - einen, nicht mehr, aber auch nicht weniger - und jedenfalls einen, der mittlerweile in seiner Prominenz außer Streit stehen dürfte: „Die sich aus dem Funktionieren moderner Wissenschaft ableitende Notwendigkeit, Wissen in Texten zu vernetzen, ist als gemeinschaftlicher Konsens in den Katalog von Anforderungen an wissenschaftliche Texte eingegangen.“ (Jacobs 1999, 63)

2. Verzerrungen durch die Exklusivität der Datenquelle „Literaturverzeichnis“

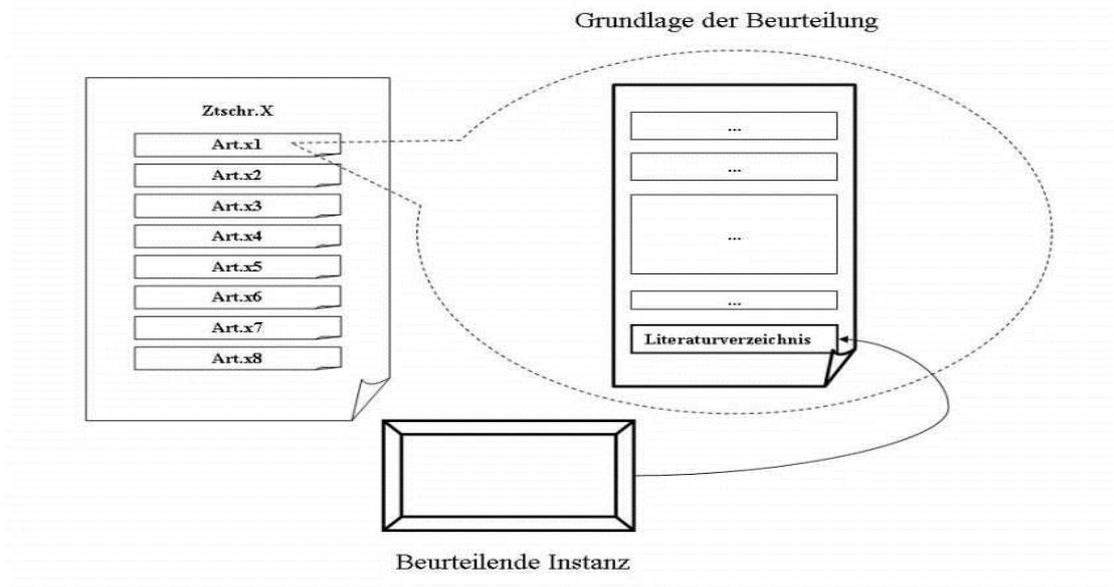
Nehmen wir nun, nach der Klärung der Beurteilungsgrundlage 1 des IF („extern“ via rezipierende Publikationsorgane) die zweite Spezifizierung der Datenquelle des IF näher in Augenschein: Aus der anfangs genannten Formel geht nicht unmittelbar hervor, was Berechnungspraxis ist: die „Anzahl“ der „Zitationen“, von denen in der Formel die Rede ist, wird *aus dem Literaturverzeichnis der zitierenden Artikel* errechnet! Ganz abgesehen davon, dass dadurch dem (tatsächlich stattfindenden) Missbrauch Tür und Tor geöffnet ist⁴, ergibt

⁴ Für diesbezügliche und andere Detailergebnisse unserer Forschung möchte ich LeserIn auf die verdienstvolle Arbeit von Julia Raphaela Strohmeyer

sich eine weitere fatale Verzerrung ebenjener wissenschaftlichen Realität, die hier beurteilt werden soll.

Wie Graphik 3 andeutet, ist ein wissenschaftlicher Artikel ein Wissenschaftsprodukt, das eine Struktur aufweist, also aus einem geordneten Ganzen unterschiedlicher Teile besteht. Eine genauere Differenzierung dieser Struktur, wie sie weiter unten entfaltet wird, sei einstweilen noch hintangestellt, aber folgende Grobeinteilung wird wohl als allgemein bekannt vorausgesetzt werden dürfen: Jeder Artikel weist so etwas wie einen „Haupt-“ und einen „Neben-Text“ auf. Im Haupttext finden wir die eigentliche Substanz dieses wissenschaftlichen Forschungsproduktes mit allen bekannten Anteilen wie Forschungsfragen, Hypothesen, Methodendiskussion, Ergebnisdarstellung u.v.m. (s. Anhang 2.2).

verweisen, die etliche Aspekte jener hier angesprochenen Alternativen zur Berechnung eines „neuen“ Impact-Faktors an einigen Musterbeispielen aufzeigt und dabei auch im Zuge Ihrer empirischen Untersuchung auf einen besonders eklatanten Fall von Missbrauch gestoßen ist (Strohmeyer 2008).



Graphik 3: Beurteilungsgrundlage 2: Textteile

Der „Nebentext“ beinhaltet auch unterschiedliches, vor allem aber „Anhänge“ verschiedener Art. Eine dieser Anhangskategorien ist für Verzeichnisse reserviert. Und unter diesen Verzeichnissen im Anhang findet sich ein prominentes, nämlich das Literaturverzeichnis, das bekanntlich die genauen Angaben zu allen jenen Werken auflistet, auf die innerhalb des Haupttextes in jeweils anderer Absicht, an verschiedenen Stellen, in unterschiedlichem Umfang und wechselnder Frequenz Bezug genommen wurde. Die Tatsache, dass es sich um einen prominenten Bestandteil des „Neben-Textes“ handelt, ändert allerdings nichts an seinem Marginalitätscharakter, wenn es um die Beurteilung der wissenschaftlichen Qualität des Haupt-Textes geht. Denn alle Informationen, die notwendig sind für die Beurteilung der wissenschaftliche Dignität des Gedankenganges, der Beweisführung etc., *also alle jene Aspekte des wissenschaftlichen Textes, die sinnvoll für so etwas wie „Quali-*

tätsbeurteilung“ dieses Produktes herangezogen werden könnten, finden sich ohnehin auch im Haupttext - und darüber hinaus noch viel mehr, nämlich alles, was wir für die Erfassung der „Talität⁵“ benötigen und damit alles, was wir im Literaturverzeichnis niemals finden können!

Berücksichtigt man zunächst nur die allerwichtigsten Punkte, die sich allein bezüglich dieser Exklusivität der Datenquelle „Literaturverzeichnis“ für eine Kritik des IF formulieren lassen, ergibt sich bereits eine besorgniserregende Liste:

„IF-Forschungsqualitätsmessung“ verzerrt die Forschungs- resp. Publikationsrealität durch Ignorieren der *enormen Bandbreite* an

⁵ Dazu später mehr...

1. quantitativen Aspekten des Zitierens:
 - a. *Umfang* der einbezogenen Fremdstextstellen
 - b. *Frequenz* der Einbeziehung innerhalb des zitierenden Textes
2. qualitativen Aspekten des Zitierens:
 - a. *Lokalisierung* und
 - b. *Funktion* des Einbezugs von Fremdstext

Durch die Tatsache dieser Ignoranz werden Absurditäten möglich, die die gesamte Problematik schlaglichtartig erhellen:

(Fiktiver) Fall 1: *Ein dreiseitiger* Artikel des Journals X wird in *einer* anderen Zeitschrift im *Vorwort* nur durch die *Angabe Autor/Jahr* je *einmal* in einer *Fußnote* genannt.

(Fiktiver) Fall 2: *Fünf* sehr unterschiedliche, *dreißigseitige* Artikel der Zeitschrift Y werden in *hundert* anderen Texten durchschnittlich *zehn* mal im *Hauptuntersuchungsteil* als Element der *zentralen Argumentation* ihrer *Forschungs(hypo)these* zitiert, davon *dreißig* mal durch *wörtliches Zitieren längerer Textstellen*.

Kein/e ernstzunehmende/r Beurteiler/in wissenschaftlicher Qualität würde auf die Idee kommen, was hier Berechnungstatsache wäre: *In beiden Fällen ergibt sich der gleiche numerische Wert für die Errechnung des IF dieser beiden Zeitschriften!*

3. Verzerrungen durch die Ausschaltung unterschiedlicher Publikations- und Zititionstraditionen

Aus den bisher referierten Informationen über die Berechnungsgrundlagen des IF geht allerdings noch nicht hervor, dass ein durchaus gewichtiger Anteil der Absurdität des oben referierten Rechenbeispiels sich einer weiteren Schwäche des IF schuldet: In die Erfassung und Verrechnung der Zitationen gehen ausschließlich solche Zitate ein, die aus Zeitschrif-

tenartikeln⁶ stammen. Monographien, Artikel aus Jahrbüchern (incl. der darin enthaltenen Literaturumschauartikel und Rezensionen) oder Schriftenreihen, HerausgeberInnenbände (incl. der darin enthaltenen einzelnen Fachartikel), Lehrbücher, Tagungsbeiträge, in anderen als Zeitschriften publizierte Forschungsberichte etc. - keiner der darin enthaltenen Texte, egal wo und wie oft sie zitiert wurden (und wie oft und wo sie wen selbst zitieren), wird in irgendeiner Weise bei der Beurteilung der „Wirkmächtigkeit“ von Publikationsorganen berücksichtigt, weder „aktiv“ noch „passiv“.

Auch wenn es noch eine Reihe von weiteren Kritikpunkten an der Usance gibt, den IF bzw. das im SCI wiedergegebene Ranking als Maß für wissenschaftliche Qualität heranzuziehen, werden gerade durch den letztgenannten Umstand die im Folgenden referierten unglaublichen Unterschiede zwischen einzelnen IF-Werten, die keineswegs fiktiv sind, sondern den tatsächlichen IF-Berechnungen entstammen, in ihrem Zustandekommen erklärbarer.

Drei Werte – drei Welten?

Im 2007 veröffentlichten Ranking weist „The Lancet“ einen IF-Wert von 25,900 auf, „Behaviour Research and Therapy“ einen von 2,887 - und die „Zeitschrift für Pädagogik“ den Wert 0,213⁷.

In etwa beobachten wir hier also ein Verhältnis von 100:10:1 ! Wohl kaum zufällig entdecken wir beim näheren Hinsehen eine Systematik: The Lancet rangiert mit vielen anderen verwandten Zeitschriften als Repräsentantin der *Medizin* in den obersten Rängen der Exzellenzliste, der Wert der zweitgenannten

⁶ Es gibt hier noch weitere fragwürdige Einschränkungen, auf die hier nicht näher eingegangen werden soll.

⁷ Zahlen aus Strohmmer 2008

Zeitschrift darf grosso modo als repräsentativ für das Ranking *psychologischer* Periodika gesehen werden - und das Publikationsorgan auf den bedauernswerten untersten Plätzen ist Angehöriger der *Bildungswissenschaft*.

Erweitern wir unseren Blick nochmals auf eine (durchaus umstrittene und hier recht salopp exekutierte) sehr prominente Unterteilung, nämlich jene in Natur- und Geisteswissenschaften, so liegt jene in der Erweiterung der Überschrift dieses Kapitels formulierbare Frage tatsächlich sehr nahe: Stammen diese Publikationsorgane nicht nur aus drei verschiedenen Welten, sondern muss man sich da nicht auch die Frage gefallen lassen, ob eine Wissenschaft, die nur ein Hundertstel (25,900 : 0,213) der Qualität einer anderen Wissenschaft aufweist, überhaupt eine Existenzberechtigung hat, ja, müsste man hier nicht sogar aktiv zur Wahrung der Gesamtqualität diese „Wissenschaft“ aus dem Kanon der anerkannten Wissenschaften ausschließen?

Betrachten wir die Publikationstradition der einen und der anderen scientific community, also jene der MedizinerInnen und jene der BildungswissenschaftlerInnen (ErziehungswissenschaftlerInnen, PädagogInnen), so erkennen wir sehr schnell, dass u.a. der letztgenannte Kritikpunkt am IF hier zum Tragen kommt, nämlich „die Tatsache, dass in den Sozial- und Geisteswissenschaften Bücher weiterhin eine bedeutende Rolle in der wissenschaftlichen Kommunikation spielen“ (Jokic; Ball 2006, 91) Der wohl überwiegende Teil jener Texte, aus denen in der Bildungswissenschaft zitiert wird und die in Fachpublikationen der Bildungswissenschaft „Wirkung zeigen“, entstammt anderen Publikationsformen als jener der Zeitschriftenartikel (vgl. dazu u.a. Heissenberger; Neubauer 2008). Garfield selbst war dieser Umstand sehr wohl bekannt,

wie ein von Stock geführtes Interviews belegt (Stock 2002, 24).

Eine Alternative

Eine vollständige Auflistung und Erörterung aller bisher bereits vorgebrachter und noch vorzubringender Kritikpunkte am Usus wissenschaftlicher Qualitätsbeurteilung, wie er sich im IF und im SCI widerspiegelt, würde noch viel mehr Arbeit leisten müssen, als hier darzustellen möglich ist. Gleichwohl kann an diesem Punkt bereits eines festgehalten werden:

Selbst wenn durch den IF „im Sinne der Erfinder“ nicht direkt „Qualität“ von Forschung beurteilt werden soll, sollte doch angesichts der bisher vorgebrachten Kritikpunkte deutlich geworden sein:

Durch simples Auszählen der Angaben im Literaturverzeichnis anderer Publikationen *ist keine sinnvolle Antwort zu erwarten*, wenn die Frage lautet:

„Welche Art und welches Ausmaß von Wirkmächtigkeit („IMPACT“) erreicht ein/e Autor/in, eine Forschungsarbeit, ein Publikationsorgan im wissenschaftlichen Diskurs einer scientific community?“

Greifen wir Brachmanns Plädoyer für eine „zitationelle Einzelfallanalyse“ auf (Brachmann 2003, 441) und folgen der von Hornbostel postulierten Notwendigkeit des „Aufzeigens einer begründeten Alternative“ (Hornbostel 2004, 87), so könnte zunächst direkt an jener Frage, die der IF zu beantworten imstande ist, überlegt werden, welche Frage sinnvoller zu stellen wäre.

Würde man eine solche Frage aus der von Garfield eröffneten „traditionellen“ Sicht der Zitologie stellen, so könnte diese auf den Punkt gebracht etwa so lauten:

Wie oft steht der Titel einer Forschungsarbeit in den Literaturverzeichnissen anderer Forschungsarbeiten aus den beiden Folgejahren?

Aus der unverstellten Klarheit dieser Formulierung lässt sich vielleicht am besten erkennen, wie *schmal der Fragenhorizont* ist, an dem sich der IF orientiert - und wie *eng das Spektrum der Antwortmöglichkeiten*. Tatsächlich gibt der Zahlenwert des IF keine andere Information her, als jene, nach der mit dieser Fragestellungsversion gesucht werden kann: Wie oft wurde der Titel einer Arbeit ein bis zwei Jahre später im Literaturverzeichnis von Zeitschriftenartikeln angeführt? Doch ist eine Antwort auf so eine Frage tatsächlich das, wonach wir suchen, wenn wir in Erfahrung bringen möchten, was einen „guten“ wissenschaftlichen Artikel ausmacht - und noch dazu einen, der in der scientific community Wirkung in der Weiterentwicklung des Wissens und der wissenschaftlichen Forschung zeigt?

Wenn wir die möglicherweise sinnvolle Verbindung von „Qualität“ und „Wirkmächtigkeit“ beibehalten wollen, aber bei der Erfassung der Ausprägung beider Faktoren in Bezug auf konkrete wissenschaftliche Texte bzw. Publikationsorgane nicht dabei stehen bleiben wollen *nach der Häufigkeit zu fragen, in der ein abgeschlossener Text als Ganzes in späteren Texten genannt wird* - was ist die Alternative?

Einerseits kritisierten wir bereits die Oberflächlichkeit und Eindimensionalität des IF/SCI-Ansatzes. Andererseits begannen wir im Gegensatz zu der in ihrem logischen Fehlschluss erkannten Idee, „*passive Qualität durch aktiven Impact*“ erreichen zu wollen, die Variante „*aktive Qualität durch passiven Impact*“ zu favorisieren. Das bedeutet, dass wir konsequenterweise nicht nur den zu beurteilenden Text als *multidimensionales* Phänomen

anerkennen müssen, sondern auch erforschen müssten, *wann, wo* und *wie* der zu beurteilende Text *selbst* während der (multidimensionalen) Textproduktion *andere Texte in den Entfaltungsprozess seiner Forschung einbezieht*.

Versuchen wir nun aus den bisher erörterten Kritikpunkten heraus in diesem gerade nochmals erläuterten Sinn eine alternative Frage zu stellen, so ließe sich diese z.B. so formulieren:

Wieviel Raum nimmt in verschiedenen Forschungsarbeiten an welchen Stellen und innerhalb welcher Arten von Aussagen welche Art von Fremdtext ein und für welchen Zweck wird Fremdtext hier in welcher Intensität und welcher Extensität eingesetzt?

Offen bleibt zunächst, ob es möglich sein wird, für diese Frage methodische Zugänge zu finden, die es erlauben, ähnlich wie beim IF/SCI Quantifizierungen durchzuführen - und damit u.a. auch die Erstellung eines Rankings zulassen.

1. Beurteilung von „Qualität“ und „Wirkmächtigkeit“ durch die Bedachtnahme auf „Standards“

Die oben als Alternative formulierte Frage hat einen auf den ersten Blick erkennbaren „Nachteil“ gegenüber der IF-Grundfrage: sie weist eine weit weniger einfache Struktur auf. Dementsprechend können ForscherInnen, die eine dieser beiden Fragen als forschungsleitende für Ihre wissenschaftliche Arbeit nehmen wollen, von vornherein nicht annehmen, dass die notwendigen Schritte zur Einlösung des jeweiligen Forschungsvorhabens in beiden Fällen gleichermaßen einfach oder schwer zu setzen sein werden.

Diese Erkenntnis führt uns unmittelbar zu einer ganz anderen prinzipiellen Frage, nämlich jener nach den Entscheidungskriterien, die

WissenschaftlerInnen ansetzen, wenn sie ihre Wahl treffen: Was möchte ich zum Gegenstand meiner Forschung nehmen? Welche Theorie zu diesem Gegenstand ziehe ich als Bezugssystem für meine Gedankengänge heran? Welche Untersuchungsmethode wende ich als die adäquateste an? Was im wissenschaftlichen Diskurs meiner scientific community will ich mit meinen Ergebnissen verändern können?

Üblicherweise hegt und pflegt jede scientific community auf verschiedenen Ebenen der Bewusstheit, Explikation und Reflektiertheit ihre je eigenen Vorstellungswelten bezüglich der „Standards“, „Kriterien“ (oder wie sie auch immer benannt werden), nach denen die Mitglieder der community ihre Wahlentscheidungen treffen sollten. Einige dieser Entscheidungsgrundlagen weisen den Charakter von „ungeschriebenen Gesetzen“ oder Usancen auf, etliche finden sich in den Lehrbüchern der jeweiligen Disziplin.

Auf welchem Wissenschaftsgebiet bewegen wir uns gerade? Wo ist diese scientific community zu finden, nach deren Üblichkeiten oder auch expliziten Vorschriften wir uns hier richten müssten? Denn wenn wir nicht nach „Lust und Laune“, sondern nach „Recht und Gesetz“ entscheiden wollen, welcher der beiden oben formulierten Forschungsfragen wir den Vorzug geben, müssten wir uns ja auf etwas „berufen“ können, was uns hilft, unsere Entscheidung zu *argumentieren*. Was wiederum bedeutet, dass wir *auf vorhandenes Wissen zurückgreifen* können müssten – *um dieses dann zu zitieren!*

Zitationsforschung hat - wie alle Wissenschaftsgebiete, die Wissenschaft untersuchen – gewissermaßen sich selbst zum Forschungsgegenstand. Und das bringt sie naturgemäß in arge Schwierigkeiten. Denn da sie als ihren

Forschungsgegenstand eine ganz bestimmte Aktivität von WissenschaftlerInnen (nämlich das Zitieren) gewählt hat, und diesen Gegenstand mit zwei anderen Gegenständen (nämlich „Qualität“ und „Wirkmächtigkeit“) in ursächliche Verbindung zu bringen trachtet, *dabei aber selbst zitieren muss und sich sowohl um die Qualität ihrer Forschung kümmern sollte als auch Einfluss auszuüben trachtet*, ist sie letztlich gezwungen, in einer nach wie vor nicht konsensuell gelösten Frage Farbe zu bekennen: Gibt es Entscheidungsgrundlagen für die oben genannten fünf Grundfragen nach dem Gegenstand, der Bezugstheorie, der Methode, den Geltungsbereichen und den zu verändernden Diskursbereichen, die für **alle** Wissenschaften (also sowohl für die Zitatologie als auch für alle von ihr untersuchten Wissenschaften) Geltung haben (können)? Gäbe es darauf eine klare, verbindliche und vor allem leicht auffindbare Antwort, wäre dieses Grundproblem der Zitatologie gelöst. *Kann* es darauf überhaupt eine *für alle* verbindliche Antwort geben? Doch wohl nur, wenn die Unterschiede zwischen den Einzelwissenschaften die Frage der Qualität und der Wirksamkeit nicht tangieren, sondern diesbezüglich nicht nur *Einigkeit der Beurteilenden* herrscht sondern auch *Einheit des Beurteilten*.

Und tatsächlich gibt ja das im SCI veröffentlichte Ranking vor, dass alle im Ranking erfassten Disziplinen eigentlich eine einzige sind, insofern sie alle mit den gleichen Maßstäben für Qualität und Wirkmächtigkeit (und damit auch für Zitation) zu bewerten sind.

Stellen sie sich vor, so würde eine Olympiade organisiert sein: alle TeilnehmerInnen werden in allen Wettbewerben mit der gleichen Methode beurteilt, z.B. nach der für die sportliche Leistung verbrauchten Zeit. Die (einzige) Goldmedaille erhält jene/r Teilnehmer/in,

der/die am wenigsten Zeit verbraucht hat, egal, ob es sich um einen 100-Meter-Läufer, eine Turmspringerin, einen Ringer, einen Stabhochspringer, eine Eiskunstläuferin handelt. So absurd eine derartige „Einheitsrangliste“ in einer solchen „Verrücktheitsolympiade“ wäre, so absurd müsste für die Wissenschaft der SCI in seiner bisher üblichen Form sein.

Doch die Praxis zeigt etwas anderes: obwohl die Kritik immer lauter und massiver wird, wird jedes Jahr aufs Neue ein Ranking nach dem anderen veröffentlicht - und zeigt seine Wirkungen...

Nun wäre es zwar undenkbar, dass man - um den oben angezogenen Vergleich noch einmal zu bemühen - die SportlerInnen angesichts der so unterschiedlichen Formen, in denen sie ihre jeweiligen großartigen Leistungen erbringen, nur mehr mit den Kriterien bewertet, die für eine oder wenige dieser Formen sinnvoll gelten (also z.B. die Zeitnehmung aus den Laufsportarten), sehr wohl aber gibt es einige Kriterien, die tatsächlich bei *allen* SportlerInnen als verbindlich und wichtig angesehen werden. So müssen z.B. alle TeilnehmerInnen ihre jeweils sehr unterschiedlichen Leistungen ausschließlich aus eigener Kraft bzw. nur mit den jeweils erlaubten Mitteln einbringen. Doping ist in allen Sportarten verboten. Jede/r darf zwar seine ganz individuell zusammengestellten Trainingsprogramme erarbeitet haben, aber es muss immer prinzipiell nachvollziehbar bleiben, wodurch der/die jeweilige Sportler/in zu ihren Höchstleistungen kam. Und die Höchstleistungen müssen in „offiziellen Zusammenkünften“ erbracht werden. Wenn ein 100m-Läufer bei sich im Garten die Strecke in Weltrekordzeit läuft, wird dies nicht als Grundlage für die Eintragung in die entsprechenden „Rankings“ akzeptiert.

Etwas Vergleichbares gibt es auch in den scientific communities: Plagiate sind verboten, die Anerkennung einer wissenschaftlichen Leistung ist an ihre Repräsentation im offiziellen Diskurs gebunden - und der Weg, auf dem WissenschaftlerInnen zu jenen Erkenntnissen kommen, die sie als ihren eigenen Beitrag zum wissenschaftlichen Diskurs der jeweiligen scientific community beisteuern, muss für alle nachvollziehbar (und damit auch kritisierbar) sein.

Das bedeutet, dass wissenschaftliche Erkenntnisse bis zu einem hohen Grad personalisiert sind, Publikationsorte und -zeiten eine gewichtige Rolle spielen - und intersubjektive Nachvollziehbarkeit für alle Wissenschaften ein Kriterium für Qualität darstellt.

Alle drei Faktoren spiegeln sich im Phänomen des Zitierens wieder: Zitieren ist per se eine Personalisierung von Wissen, die Aussagen werden mit den Namen ihrer AutorInnen wiedergegeben. Für jedes Zitat muss im gleichen Artikel ein Ort reserviert sein, an dem der Publikationsort und das Publikationsdatum genannt werden. Und eben diese Angaben, also die „Daten“, die angeben, von wem, woher und welcher Zeit ich diese oder jene Aussage entnehme, ermöglicht es LeserIn den Beitrag des Artikelschreibers bzw. der Artikelschreiberin zum Diskurs der scientific community besser nachvollziehen und beurteilen zu können.

„Intersubjektive Nachvollziehbarkeit“ und „Diskursanbindung“, diese zwei Kriterien sollen im Weiteren als „kleinster gemeinsamer Nenner“ aller Wissenschaften, wenn sie auf Qualität Wert legen, besonders berücksichtigt werden.

Welches Kriterium aber können wir sinnvoll für die Frage nach „Wirkmächtigkeit“ einsetzen?

zen, wenn wir auf „Standards“ Bezug nehmen wollen?

2. Beurteilung von „Qualität“ und „Wirkmächtigkeit“ über „Intersubjektive Nachvollziehbarkeit“ und „Diskursanbindung“

Eben diese beiden Aspekte wissenschaftlicher Textproduktion, die Einbindung in bestehende Diskurse der scientific community auf der einen Seite und die Deklaration aller Wege, die von der (wissenschaftlichen Forschungs-)Frage zur (wissenschaftlichen) Antwort führen, auf der anderen Seite - beide bieten uns nicht nur für die Frage nach Qualität Antwortmöglichkeiten, sondern - via Zitatologie - auch für unser Nachdenken über Wirkmächtigkeit.

Wenden wir uns noch einmal unserem Beispiel aus dem Kapitel „Drei Werte - Drei Welten?“ zu: Strohmer (2008) hat je einen Musterbeispielsartikel aus jeder der drei genannten Zeitschriften („The Lancet“, „Behaviour Research and Therapie“, „Zeitschrift für Pädagogik“) bezüglich des Zitierverhaltens der AutorInnen näher untersucht. Die Zahlen, die im Folgenden Grundlage sowohl der Tabellen als auch der Interpretationen sind, entstammen dieser Arbeit.

Zunächst lohnt ein Blick auf die Verteilung von Eigen- und Fremdttext in den drei Artikeln:

The Lancet		Behaviour Research and Therapy		Ztschr.f. Pädagogik	
Fremdttext	Eigentext	Fremdttext	Eigentext	Fremdttext	Eigentext
16,11 %	83,89 %	14,13 %	85,87 %	80,31 %	19,69 %

Tabelle 1: Verteilung von Eigen- und Fremdttext in den drei Musterbeispielsartikeln (Zahlen aus Strohmer 2008)

Auf den ersten Blick ist zu erkennen, das die ersten beiden Artikel sich massiv von dem letzten in der Verteilung Eigentext-Fremdttext unterscheiden, ja sich das Verhältnis nahezu umdreht: 16:84 und 14:86 auf der einen Seite und 80:20 auf der anderen!

Der pädagogische Artikel weist also 80% Fremdttext auf! Wie ist diese Tatsache zu erklären? Das Verhältnis der anderen beiden Artikel wäre vielleicht von einer Vorabschätzung der LeserInnen nicht so weit entfernt gewesen, aber würde man einen wissenschaftlicher Artikel erwarten, der nahezu *nur* Fremdttext enthält?

Ein erster Hinweis zur Lösung dieses Rätsels liegt in den Daten der nächsten Tabelle:

	The Lancet	Behaviour Research and Therapy	Ztschr.f. Pädagogik
Verweis	14,10	11,23	1,97
Wiedergabe wörtlich	0,00	5,47	18,97
Wiedergabe sinngemäß	75,07	62,38	57,87
Wiedergabe erweiternd	4,73	1,32	13,14

Tabelle 2: Bezüge auf Fremdttext (Zahlen aus Strohmer 2008)

Hier sehen wir mehrere Gegenläufigkeiten: Einerseits verwenden die ersten beiden Artikel wesentlich mehr Verweise als der dritte, reduzieren also das zu zitierende auf die bloße Angabe von Autor und Jahr des zitierten Werkes, ohne Angabe irgendwelcher Inhalte. Umgekehrt sehen wir beim dritten Artikel ungleich mehr wörtliche Wiedergaben als bei den ersten beiden. Ebenso finden wir dort deutlich mehr „erweiternde Wiedergaben“.

Der pädagogische Artikel setzt also weniger auf „isolierte“ Angaben anderer Werke, sondern entnimmt aus diesen mehr und direktere Informationen als die anderen beiden - *und er*

„bearbeitet“ (anders als der medizinische und der psychologische Artikel) diese direkt entnommenen Informationen im Sinne von „Erweiterungen“.

Die letztgenannte Interpretation beinhaltet allerdings einen weiteren Auswertungsschritt:

Die Verwendungen von Fremdtext wurden in den drei Artikeln auch bezüglich ihres Auftretens im Textfluss untersucht. Dieser Vorgangsweise liegt die Idee zugrunde, es könnten sich „Muster“ der Fremdtextverwendung finden lassen, die in einer bestimmten Abfolge der unterschiedlichen Bezugnahmen bestehen.

Diese Idee soll an einem Ausschnitt aus einem anderen - psychoanalytisch-pädagogischen Artikel illustriert werden:

„Als Freud im Jahre 1895 in den „Studien über Hysterie“ vier Falldarstellungen publiziert, beginnt er seine Nachbemerkungen zur Falldarstellung „Elisabeth v. R.“ mit den viel zitierten Sätzen:

„Ich bin nicht immer Psychotherapeut gewesen, sondern bin bei Lokaldiagnosen und Elektroprognostik erzogen worden wie andere Neuropathologen, ..., während eine eingehende Darstellung der seelischen Vorgänge, wie man sie vom Dichter zu erhalten gewohnt ist, mir gestattet, bei Anwendung einiger weniger psychologischer Formeln doch eine Art von Einsicht in den Hergang einer Hysterie zu gewinnen“ (Freud 1895, 131).

Freud bringt an dieser Stelle zum Ausdruck, dass seine „Krankengeschichten“ von gängigen Formen der Darstellung wissenschaftlicher Ausführungen abweichen, zieht aber zugleich in Erwägung, dass die Veröffentlichung von novellenartig verfassten Falldarstellungen durchaus der Eigenart des Gegenstandes geschuldet sein könnte, dem Freuds wissenschaftliches Interesse zusehends gilt.

Mit anderen Worten: Der Rückgriff auf novellenartig gehaltene Fallberichte könnte geradezu nötig sein, wenn Leserinnen und Lesern deutlich gemacht werden soll, welche Schritte der Erforschung der „inneren Welt“ von Menschen Freud veranlassten, sich von alt-hergebrachten Theorien des Psychischen zu lösen, um statt dessen neue Theorien zu entwickeln und zu publizieren.

Heute, mehr als hundert Jahre später, gehen selbst innerhalb der psychoanalytischen Fachgesellschaften die Auffassungen darüber auseinander, welche spezifische Relevanz solchen Einzelfalldarstellungen zukommt.

Dessen ungeachtet werden Fallberichte nach wie vor publiziert - in psychoanalytisch-therapeutischen Büchern und Zeitschriften ebenso wie in psychoanalytisch-pädagogischen Publikationsorganen, die nur am Rande psychotherapeutisch orientiert sind.

Umso bemerkenswerter ist der Umstand, dass in psychoanalytisch-pädagogischen Kontexten über den Stellenwert solcher Fallberichte nur gelegentlich in expliziter Form diskutiert wird.“ (Datler u.a. 2004, 7)

Der Autor⁸ beginnt seine Ausführungen mit einem (durch einen kommentierten Verweis auf die zitierte Stelle eingeleiteten) Zitat. Dieses Zitat hat aber, so wird in den nächsten Sätzen klar, einen besonderen Stellenwert: es wird weder als „Zierde“ noch als nahtlos in einen Gedankengang eingefügtes „Zitiergut“ eingesetzt. Wäre es ersteres, so würde es wie üblich abgesetzt vom Fließtext „über“ den Artikel gestellt und im Folgenden mit keinem Wort mehr erwähnt werden, da es ja nur als „Aufputz“ des ganzen Artikels dient, als kleiner

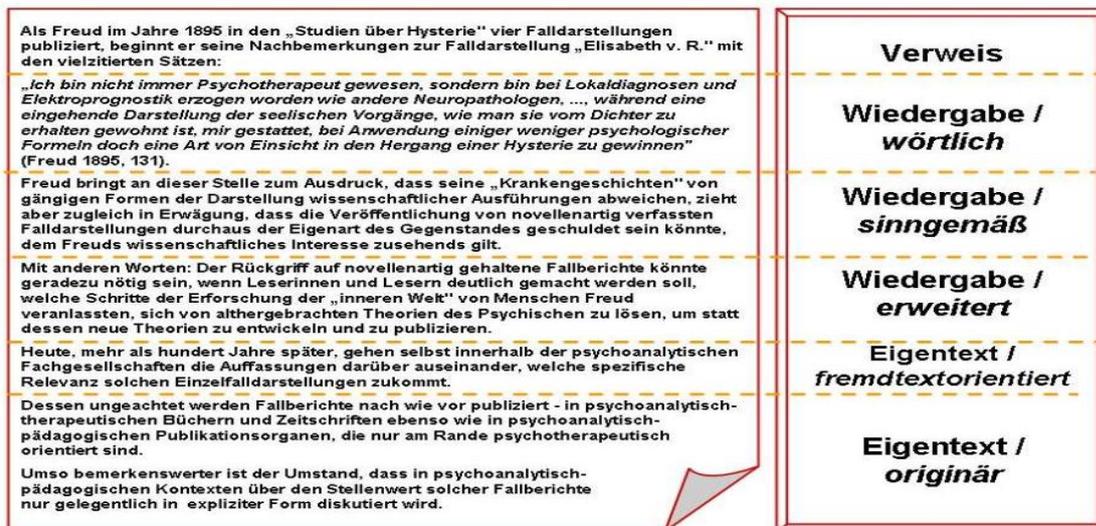
⁸ Es soll hier zunächst außer Acht gelassen werden, dass es sich in diesem Fall um eine AutorInnengruppe handelt, der Einfachheit halber wird im Folgenden auch immer nur der Erstautor (Datler) benannt.

Hinweis auf die Belesenheit des Autors und/oder als „Auftakt“ einer Themenbearbeitung. Wäre zweiteres der Fall, würden nur mehr die Anführungsstriche darauf hinweisen, dass es sich um „Fremdtext“ handelt, das Zitierte würde sich aber als eine Aussage unter mehreren, die alle dem gleichen Ziel zustreben, lesen lassen. Im Gegensatz dazu ist das, was der Autor hier eröffnet, ein Gedankengang, der sich auf den Autor des Zitates bezieht, *sodass der Zitierte (bzw. seine im Zitat aufscheinenden Gedanken) gewissermaßen zum Untersuchungsgegenstand des Zitierenden wird*. Gleich anschließend an das wörtliche Zitat „entfaltet“ der zitierende Autor das „Zitiergut“ weiter, indem er es gewissermaßen *nochmals* zitiert, diesmal aber nur mehr sinngemäß, d.h. er gibt nicht etwas anderes, neues wieder, sondern interpretiert das bereits Wiedergegebene ein Stück weit, indem er es *noch einmal mit eigenen Worten* den LeserInnen vor Augen führt. Kennt man allerdings den ganzen Artikel, fällt es nicht schwer zu erkennen, dass die Wortwahl und Formulierung, mit der der Autor das Zitierte „verdeutlicht“, nicht

zufällig ist, sondern vielmehr dem „Ganzen“ dient, indem sie das Zitierte gewissermaßen dem Ziel des Artikels anpasst.

Im nächsten Absatz ist dann zwar noch immer das Zitat Inhalt der Aussagen des Autors. Allerdings führt er jetzt etwas ein, was im Text des Zitierten nicht mehr so ohne weiteres zu finden ist. Vielmehr *erweitert* der Autor die Aussagen des Zitierten beträchtlich. Gleichzeitig ermöglichen die dargestellten Überlegungen, die nun gewissermaßen teilweise vom Zitierten, teilweise vom Zitierenden stammen, einen nahtlosen Übergang zu *eigenen* Aussagen des Autors, auch wenn diese im Zitat gewissermaßen ihren Ausgangspunkt finden. Der auf diesen Absatz des Autors folgende Satz, in dem er sich langsam vom Zitierten zu lösen beginnt, stellt dann die erste „eigentliche“ Aussagen des Autors dar - also ein Stück „originärer“ Eigentext.

In der folgenden Graphik 4 werden die soeben erfolgten Kategorisierungen der Textteile überblicksartig und direkt auf die einzelnen Textblöcke bezogen wiedergegeben.



Graphik 4: Abfolge von Bezugstypen als Diskursivierungsmuster

Die hier aufgefundenen Bezugstypen, die vor den Eigentext in den Text eingebracht werden, können *in ihrer Abfolge* als ein mögliches *Muster der Eigentext-Fremdtext-Relation* im Sinne eines *Diskursivierungstypus* angesehen werden. Eine Art der Diskursivierung, die als *eine auf Fremdtext aufbauende, systematische und schrittweise Entfaltung von Eigentext* (bei hoher intersubjektiver Nachvollziehbarkeit) angelegt ist.

Dieses Beispiel sollte illustrieren, was mit „*Beurteilung von „Qualität“ und „Wirkmächtigkeit“ über „Intersubjektive Nachvollziehbarkeit“ und „Diskursanbindung“ im Sinne „aktiver Qualität durch passiven Impact“*“ markiert wurde: Die besprochene Textstelle verleiht der zitierten Arbeit Freuds einen (relativ) hohen Grad an Wirkmächtigkeit, indem *größere Teile des Artikels an sie als Diskurselement anbinden*. Im „Gegenzug“ erreicht die Arbeit des Autors durch den Einsatz des oben genannten Diskursivierungstypus einen (relativ) hohen Grad an „Qualität“, insofern die *intensive und extensive Diskursanbindung* mit einer *hohen intersubjektiven Nachvollziehbarkeit der Argumentation* einhergeht.

3. Beurteilung von „Qualität“ und „Wirkmächtigkeit“ über Textanalysen

Im eben angerissenen Beispiel wurde das Augenmerk nicht auf die im Literaturverzeichnis aufgelisteten Werke gelegt, sondern dem eigentlichen Werk des Autors ein Stück nachgegangen, das ja nicht in der Nennung eines anderen Werkes besteht, sondern in der „*Einarbeitung“* der *in diesem anderen Werk enthaltenen Aussagen in die Aussagenstruktur des eigenen Textes*. Hier wird also davon ausgegangen, dass „Impact“ nicht eine Bezeichnung dafür ist, wie oft ein Text in anderen Texten genannt wird, sondern *inwiefern und in welcher Weise er **tatsächlich** Einfluss auf die Produktion wissenschaftlicher Erkenntnisse in*

anderen Texte genommen hat (bzw. ihm die Gelegenheit dazu gegeben wurde).

Um die Quantität und Qualität dieser „Wirkmächtigkeit“ erfassen zu können, wird hier also ein essentiell anderer Zugang zum Problem der Einflussmessung eröffnet, als wir ihn aus den Berechnungen des IF kennen: Nicht Literaturverzeichnisse anderer Texte sind der Ort, an dem Information zur Beantwortung der Frage nach Einfluss und Qualität gesucht und gefunden wird, sondern *der zitierende Text selbst!* Und die Suche nach einschlägigen Informationen in diesem Text geht in zwei Richtungen: In Richtung der *zitierenden* AutorInnen einerseits und in Richtung der *zitierten* AutorInnen andererseits. Die Berücksichtigung der Aspekte eines zitierenden Textes ermöglichen also sowohl eine Zuteilung von Werten zum zitierenden Text als auch zum zitierten Text: der zitierte Text „besetzt“ einen mehr oder minder großen (und mehr oder minder bedeutsamen) Ausschnitt des zitierenden Textes - und der zitierende Text *bindet* in mehr oder minder extensiven und intensiven Weise an relevanten Fremdtext an.

Wie beide Blickrichtungen in dieser hier vorgestellten Form der alternativen Berechnung von „Impact-Faktoren“ eines Textes oder eines Publikationsorgans in die Konstruktion eines Analyseinstruments eingehen, wird weiter unten angedacht. In jedem Fall aber geht es nicht mehr um das Abzählen der Häufigkeit von Namen und Jahren in einer Liste, sondern um die Erforschung eines hochkomplexen Phänomens: des wissenschaftlichen Textes! Und damit betreten wir das Gebiet der Textanalyse, eines Instrumentariums, das üblicherweise zur Erforschung verschiedenster Arten von Texten eingesetzt werden kann und wird, so gut wie nie aber zur „Selbsterforschung“, also zur *wissenschaftlichen Untersuchung wissenschaftlicher Texte*. An anderer

Stelle werde ich näher auf dieses Forschungsfeld eingehen, zunächst möchte ich nur auf jenen methodischen Ansatz direkt Bezug nehmen, der in dem hier vorgestellten alternativen Zugang zur Berechnung von Impact-Faktoren eingesetzt werden soll. Es kann aber bereits festgehalten werden, dass der Anspruch, dem Einfluss von Texten *tatsächlich* nachzugehen die Aktivierung eines komplexen methodischen Untersuchungsinstrumentes zur Folge haben muss!

These 1:

Die Erforschung des Einflusses von wissenschaftlichen Texten auf andere wissenschaftliche Texte zwingt Empirie und Hermeneutik in die Kooperation

Wenn es darum gehen soll, über das Abzählen der Nennung eines Werkes im Literaturverzeichnis anderer Werke hinaus- und in den Text selbst hineinzugehen - und trotzdem die Möglichkeit aufrecht zu erhalten, über *Quantifizierungen* auch *Vergleiche* zwischen der Wirkmächtigkeit verschiedener Texte anzustellen, so ergibt sich eine Notwendigkeit, der auszuweichen ganze ForscherInnengenerationen größte Anstrengungen unternommen haben: die Zuteilung von „Empirie“ und „Hermeneutik“ zu zwei verschiedenen und nicht miteinander verbindbaren Welten muss so weit wie möglich aufgegeben werden. Dabei müssen die Eigenart und die „wissenschaftliche Potenz“ beider Zugänge erhalten bleiben, wenn sowohl die Quantifizierbarkeit des Faktors „Wirkmächtigkeit“ erhalten bleiben soll als auch das Eingehen auf das komplexe Phänomen „wissenschaftlicher Text“, in dem das Phänomen „Impact“ gesucht und gefunden wird.

Was hier also notwendig wird, ist der Einsatz einer „Empirisch-Hermeneutischen Textanalyse“ (Stephenson 2003), die als Haupt-

instrument zur Erfassung hochkomplexer Texte für die hier vorgestellte Version von Zitatalogie fungieren soll.

These 2:

Qualifizierung und Quantifizierung von Zitaten müssen Kategorisierungen von „Textsegmenten in Kontexten“ ermöglichen

In Graphik 4 wird bereits ein Grundprinzip dieser „Empirisch-Hermeneutischen Textanalyse“ ersichtlich: Die Kategorisierungen („Verweis“, „Wiedergabe/wörtlich“, usw.) sind ganz bestimmten und klar begrenzten *Textausschnitten* zuordenbar. Diese Textausschnitte weisen einen bestimmten Umfang auf - und eine bestimmte Abfolge. Damit sind die aus dem Gesamttext „ausgeschnittenen“ Textteile, die im Folgenden allgemein als „Textsegmente“ bezeichnet werden, einerseits sowohl - über die Kategorisierungen - *qualifizierbar*, als auch - über den Umfang - *quantifizierbar*. Andererseits sind die Kategorisierungen immer wieder auch *kontextbezogen*: Ob ein Segment als „Wiedergabe erweitert“ oder als „Eigentext originär“ signiert wird, hängt *auch* davon ab, ob und in welcher Weise es sich auf im unmittelbaren Kontext vorfindbaren Fremdtext oder Eigentext beziehen lässt.

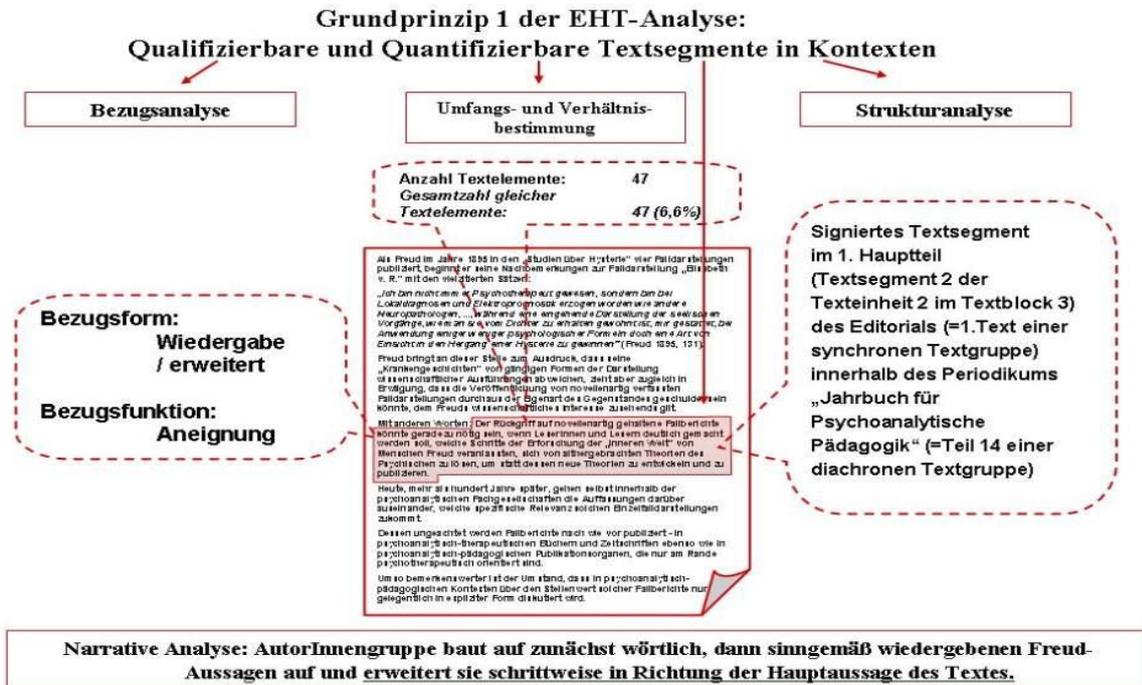
These 3:

Impactberechnungen über die Erfassung „kategorisierter Textsegmente in Kontexten“ setzen eine Verbindung von Bezugsanalysen und Strukturanalysen voraus

Die oben erwähnte Möglichkeit, ein (wörtliches) Zitat könne ja auch nur als „Zierde“ vor dem Beginn des eigentlichen Textes stehen, war nur ein erster Hinweis auf einen weiteren zu berücksichtigenden Umstand, der für den „tatsächlichen“ Einfluss eines Textes auf einen

anderen ebenfalls maßgeblich ist. Die folgende Graphik zeigt die Grundlage der Verbin-

dung von These 2 und 3:



Graphik 5: Grundprinzip 1 der EHT-Analyse

In der Graphik wird *ein* Textsegment rot hervorgehoben und in den drei „Sprechblasen“ (links, rechts und oberhalb) näher bestimmt. Die links wiedergegebenen Bestimmungen stellt die „Qualifizierung“ dar, die rechte die *Kontextbestimmung* und die obere die „Quantifizierung“. Der hier in den Artikel (Datler u.a. 2004) eingearbeitete Fremdtext (Freud 1895) nimmt also im rot markierten Textsegment in Form einer erweiterten Wiedergabe und im Umfang von 6,6% des erweiterten Kerntextes (s.u.) Einfluss auf den (1.) Hauptteil eines Editorials der 14. Ausgabe eines Periodikums der Psychoanalytischen Pädagogik.

Um zu diesem Ergebnis kommen zu können, wurde der Text selbst in Bezug auf zwei Di-

mensionen⁹ (die wiederum ein bis mehrere Unterdimensionen beinhalten) und zwei Auswertungsaspekten untersucht:

In der *Bezugsanalyse* werden die Textsegmente hinsichtlich der *Form* und der *Funktion* kategorisiert, in denen und über die Bezüge hergestellt werden.

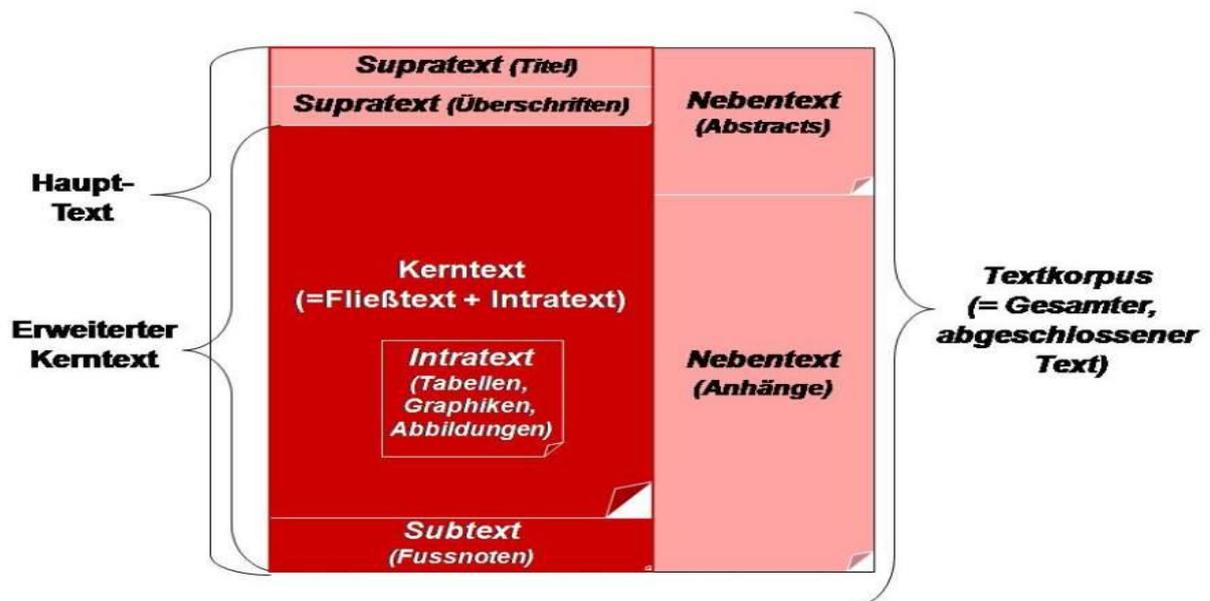
In der *Strukturanalyse* werden die *Orte*, von denen aus und zu denen Bezüge erfolgen, erfasst. Eine Liste der (Unter-)Dimensionen, durch die eine differenzierte Erfassung bzw.

⁹ Siehe dazu auch die im Anhang vollständig aufgelisteten Dimensionen und Kategorien der aktuellen Version der Empirisch-Hermeneutischen Textanalyse EHT 5.1.

Kategorisierung solcher „Orte“ erfolgen kann, ist im Anhang wiedergegeben (s.v.a. Anhang 2.2, aber auch Anhang 1.8; 1.13; 2.1; 2.3-2.6). Im Prinzip wird davon ausgegangen, dass Texte nach innen und nach außen hin „verschachtelt“ sind: Vom kleinsten Textelement (1 Wort) bis hin zum „Textkosmos“ als Gesamt aller Texte gibt es verschiedene Rahmensetzungen, die diese „Textholons“ wie die Teile einer russischen Puppe ineinander verschachtelt zeigen: Ein Wort „steckt“ in einem Satz, dieser in einem Absatz, der wieder in einem Kapitel enthalten ist, das wiederum Teil eines Artikels ist. Dieser Artikel wiederum ist Teil einer bestimmten Zeitschriftenausgabe, die wieder Teil eines Jahrganges als Teil eines ganzen Publikationsorgans ist usw.

Damit können Aussagen über den Einfluss, den ein wissenschaftlicher Text auf andere ausübt, für klar definierte Rahmensetzungen getätigt werden: In welcher Weise und in welchem Ausmaß übt ein bestimmter einzelner Text eines Autors (oder das gesamte Werk dieses Autors) auf einen einzelnen Text, eine Zeitschriftenausgabe, ein ganzes Periodikum aus?

Der einzelne Artikel ist in sich wiederum nicht nur durch kleinere und größere Textsegmente strukturiert, sondern auch durch eine Reihe von *Textarten*, die unterschiedliche Funktionen erfüllen:



Graphik 6: Textarten

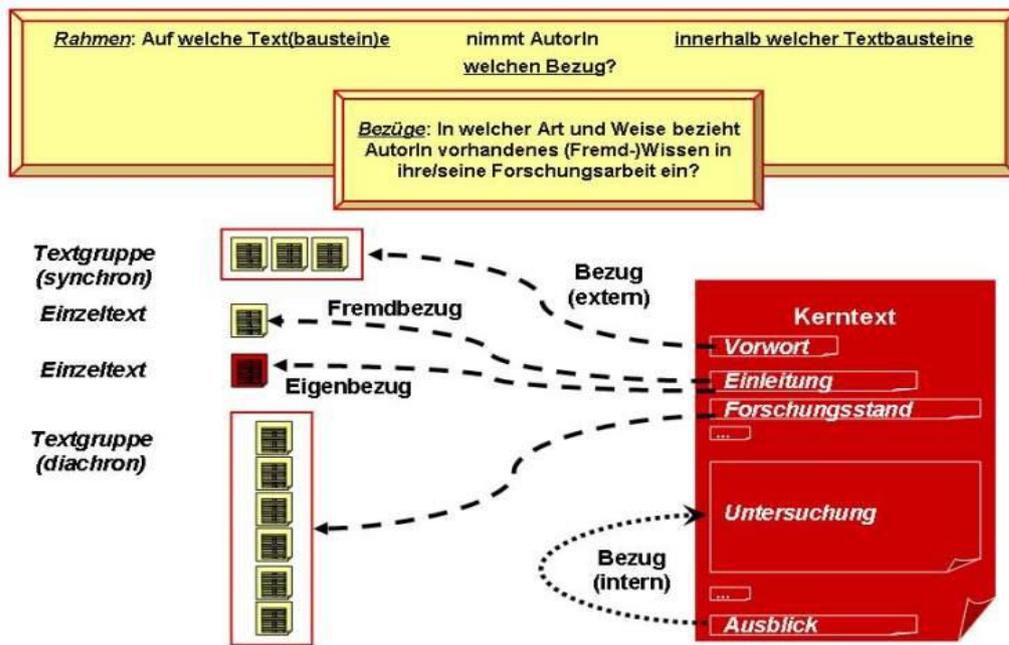
Die Erfassung der Wirkungen, die zitierte Texte auf die zu untersuchenden Texte ausüben, konzentriert sich bei der hier dargestellten Textanalysemethode auf den *erweiterten Kerntext*¹⁰, wobei hier wiederum das Augenmerk auf die einzelnen *Textbausteine* eines wissenschaftlichen Textes gelegt wird (siehe Anhang 2.2) Die Liste der Textbausteine enthält die im Wissenschaftsbetrieb allgemein üblichen Teile eines Fachartikels wie „Referieren des Forschungsstandes“, „Formulieren der Forschungsfrage“, „Darstellung der Methodik“ etc.

Nicht nur das Ausmaß¹¹, indem ein bestimmter Fremdtext in das Werk eines anderen Autors einbezogen wird, sondern auch die Antwort auf die Frage, ob der Bezug innerhalb des Vorwortes, des Forschungsstandes oder der Wahl der Methodik hergestellt wird, ist für die Bestimmung des Einflusses, den er ausübt, von entscheidender Bedeutung: Der Beitrag der einzelnen Textbausteine zur Wirkmächtigkeit, die ein wissenschaftlicher Text eines Autors bzw. einer Autorin für den Diskurs einer scientific community erhält, ist sehr unterschiedlich. Das „Vorwort“ hat diesbezüglich nur marginale, die „Rückbindung der Ergebnisse an den disziplinären Diskurs“ hingegen maximale Bedeutung. Der Einfluss auf ihr Werk, den ein/e Autor/in einer/einem anderen Autor/in zugesteht, ist daher auch daran zu messen, in welchen Baustein seines/ihres Textes er/sie den Fremdtext integriert!

¹⁰ Also den Fließtext (inklusive „Intratext“ im Sinne anderer *Textformen* (s. Anhang 2.4) wie Abbildungen, Graphiken, Tabellen etc.) des Kerntextes und den Text in den (inhaltlichen) Fußnoten.

¹¹ Also der Gesamtumfang all jener Textteile, die durch Verweise, aber v.a. über verschiedene Arten der Wiedergabe bis hin zur Kategorie „fremdbezogener Eigentext“ den Text eines Autors durch „Bezug auf Fremdtext“ bestimmen.

Grundprinzip 2 der EHT-Analyse: Bezüge in und zwischen unterschiedlichen Rahmensetzungen



Graphik 7: Grundprinzip 2 der EHT-Analyse

In der Graphik 7 ist an dem fiktiven Beispiel zu sehen, dass Fremdtex te unterschiedlicher Art¹² in den zitierenden Text eingebunden sind. Für jede/n dieser „KandidatInnen“ gibt es unterschiedliche *Gewichtungen*, die den Umfang, in dem der Fremdtex t auf die eine oder andere Weise „wirkmächtig“ wird, noch einmal relativieren: Selbst wenn der Umfang und die Form der Bezugnahme für alle Fremdtex te gleich wäre, würde die Erfassung des „Impacts“ je nach *Textbaustein* unterschiedliche Ergebnisse erbringen!

¹² Eine synchrone Textgruppe (also gleichzeitig publizierte Texte wie es bei einer einzelnen Zeitschriftenausgabe der Fall ist), ein von einem/einer anderen Autor/in geschriebener Text, ein vom gleichen Autor geschriebener Text und eine diachrone Textgruppe (also eine Ansammlung von Texten aus verschiedenen Publikationszeitpunkten wie bei einer Schriftenreihe).

These 4:

Sowohl der zitierende als auch der zitierte Text können über die Einführung „aktiver“ und „passiver“ Impact-Faktoren bezüglich des Ausmaßes des in ihnen enthaltenen „Einflusspotentials“ erfasst werden

Kehren wir zurück zu dem Beispiel aus Graphik 5: Hier eröffnete ein Autor (Datler) in seinem Artikel (2004) einen bestimmten Raum, in dem er einen anderen Autor (Freud) mit seinem Artikel (1895) Einfluss auf sein Denken, dass er in diesem Artikel entfaltete, einräumte. Betrachten wir nur den in der Graphik wiedergegebenen Ausschnitt, und berechnen wir das *Verhältnis zwischen den fremdtextbezogenen Segmenten und den Segmenten mit originärem Eigentext* über die Anzahl der kleinsten Textelemente (Worte), so ergibt sich ein Ver-

hältnis von 80,89 : 19,11. Dieses entspricht (ob zufällig oder nicht sei dahingestellt) weitestgehend dem in Strohmayer 2008 angegebenen Verhältnis des Musterbeispiels aus der Zeitschrift für Pädagogik (80,31 : 19,69). Damit ist dies ein besonders geeignetes Beispiel, an einem kurzen Textausschnitt zu illustrieren, was denn ein solches Verhältnis bedeuten kann, von dem wir uns anfangs gefragt haben, wie es überhaupt zustande kommen kann:

Es dreht sich zwar nahezu der ganze Text(ausschnitt) um Erkenntnisse, die der Arbeit eines anderen Autors entnommen wurden, aber gleichzeitig verleihen diese „Fremd“-Aussagen - und der Beitrag, den der zitierende Autor selbst leistet (Interpretation des Zitierten, Entfaltung verschiedener Aspekte des Zitierten, Erweiterung der zitierten Aussagen) der abschließenden Aussage in Form eines originären Eigentextes ein ganz besonderes Gewicht. Vor allem aber führt uns der zitierende Autor dadurch eine Form wissenschaftlichen Denkens vor, die in höchstem Maße dem wissenschaftlichen Diskurs dient: Ein Thema wird dem aktuellen Diskurs entnommen, eine Position eines Autors aus der scientific community wird zitiert, erläutert und entfaltet - und entlang dieser Entfaltung baut der zitierende Autor seine eigene Position auf, die er zu diesem dem aktuellen Diskurs entnommenen Thema einnimmt. Gleichzeitig wird für LeserIn genau nachvollziehbar, wie er zu dieser eigenen Position kommt.

Mit einer entsprechenden Gewichtung, die auf die bisher und weitere anderenorts darzustellenden Faktoren Rücksicht nimmt, können nun die errechneten Werte in zwei Richtungen „verbucht“ werden: Für den zitierenden Autor auf die Seite seines „passiven“ Impacts (an welchem Ort, in welcher Weise und in welchem Umfang gibt der Autor anderen Autoren Raum in seinem eigenen Text) - und für den zitierten

Autor auf die Seite seines „aktiven“ Impacts (an welchem Ort, in welcher Weise und in welchem Umfang geben andere AutorInnen ihm Raum in ihrem Text). Die entsprechenden Werte können nun bezüglich verschiedener Rahmensetzungen aggregiert werden: z.B. könnten Gesamtwerte für den Artikel Freud 1895 in der Zeitschrift für Pädagogik errechnet werden - oder Gesamtwerte für die gesammelten Schriften Freuds (in denen der zitierte Artikel publiziert wurde) in ihrem Einfluss auf die Artikel des Autors Datler in einer bestimmten „Schaffensperiode“ - die Auswertungsmöglichkeiten sind nur durch die jeweiligen Forschungsinteressen begrenzt!

These 5:

Für den zitierten Text kann das Ausmaß an „Anschlussfähigkeit“ berechnet werden, für den zitierenden das Ausmaß an „Diskursanbindung“

Die beiden soeben genannten Auswertungsmöglichkeiten können nun in eine bestimmte Richtung ausgebaut bzw. systematisiert werden. Die erste Variante bringt den Begriff der „Anschlussfähigkeit“ ins Spiel: wie weit ist ein Autor bzw. eine Autorin in seinem/ihrer Werk (bzw. ein bestimmtes seiner/ihrer Werke) für andere AutorInnen „attraktiv“, d.h. inwieweit können seine/ihre Gedanken in aktuelle Diskursausschnitte eingepasst werden? Und umgekehrt: inwieweit bindet sich der zitierende Autor an aktuelle Diskurse an? Für beide Aspekte entsprechende Werte zu errechnen, ist nach dem bisher Entfalteten nur mehr eine Gewichtungs- und Berechnungsaufgabe.

Jedenfalls bieten sich hier einer „alternativen“ Zitations- resp. Rezeptionsforschung eine Reihe von systematisch zu entwickelnden Möglichkeiten, „Qualität“ und „Wirkmächtigkeit“ in ihren spezifischen Bedeutungen und ihren qualitativen wie quantitativen Erfassungsmöglichkeiten

auf eine theoretisch und methodisch wohl begründete Basis zu stellen.

In der hier skizzierten Form der Zitations-, Rezeptions- und Diskursforschung setzt die Untersuchung der Wissensproduktion der jeweiligen scientific communities direkt an den wissenschaftlichen Originaltexten als spezifischem „Gewebe“ von Fremdtext und Eigentext an. Die Realisierung eines solchen komplexen textanalytischen Forschungsansatzes soll jene Tendenzen unterwandern helfen, die „wissenschaftliche Qualität“ einerseits in einer fatalen Reduktion als eine für alle scientific communities gleiche und gleich zu erfassende Eigenschaft wissenschaftlicher Forschung definieren und andererseits die Erfassung ihrer jeweiligen Ausprägung auf die Fokussierung von „Randscheinungen“ der wissenschaftlichen Textproduktionen einzuschränken.

Im 2. Teil von „Talität versus Qualität“ wird näher auf die skizzierten textanalytischen Ansätze der Qualitätsprüfung in der Zitations-, Rezeptions- und Diskursforschung eingegangen und deren mögliche Zielsetzungen und Schwerpunktbildungen an Beispielen ausgeführt.

Literaturverzeichnis

Brachmann, J. (2003). Gezählte Kompetenz. Zur Zitationsanalyse in der Erziehungswissenschaft. - In: Zeitschrift für Erziehungswissenschaft, 2003, Heft 3, 432-448.

Daniel, H.-D. (Hg.) (1988). Evaluation von Forschung. Konstanzer Beiträge zur Sozialwissenschaftlichen Forschung. Band 4. - Konstanz: Universitätsverlag Konstanz GmbH 1988.

Datler, W. (2004). Wie Novellen zu lesen...: Historisches und Methodologisches zur Bedeutung von Falldarstellungen in der Psychoanalytischen Pädagogik. In: Datler, Müller, Finger-Trescher (2004), 9-41.

Datler, W., Müller, B., Finger-Trescher, U. (2004) (Hrsg.). Sie sind wie Novellen zu lesen... Zur Bedeutung von Falldarstellungen in der Psychoanalytischen Pädagogik. Psychosozial-Verlag, Gießen.

Garfield, E. (1998). From Citation Indexes to Informetrics. Is the Tail Now Wagging the Dog? – Libri, 48(2), 67-80.

Garfield, E. (2006). Commentary: 50 years of citation indexing. – In: International Journal of Epidemiology 2006/9.

Gstach, J. (1991). Die psychoanalytische Identität der Individualpsychologie im Spiegel der Literaturangaben individualpsychologischer Autoren. – Z.f.Individualpsychol. 16: 39-53.

Gstach, J. & Brinskele, H. (2005). Zur individualpsychologischen Identität – revisited. Eine neuerliche Analyse individualpsychologischer Literaturangaben – Z.f.Individualpsychol. 30: 115-142.

Heissenberger, E. & Neubauer, T.: (2008). Zur Zitationskultur der Bildungswissenschaft und der Psychologie als wesentlicher Hinweis auf unterschiedliche Arten des Rezipierens von Veröffentlichungen. Eine Untersuchung mit besonderem Verweis auf sonder- und heilpädagogische Veröffentlichungen. - Diplomarbeit: Universität Wien 2008.

Hornbostel, S. (2004). Der Drahtseilakt - Evaluationskriterien zwischen fachinterner Relevanz und externer Definition. - In: Merkens, H. (Hg.) 2004, 77-88.

Jakobs, E.-M- (1999). Textvernetzung in den Wissenschaften. Zitat und Verweis als Ergebnis rezeptiven, reproduktiven und produktiven Handelns. – Tübingen: Max Niemeyer-Verlag 1999.

Jokic, M. & Ball, R. (2006). Qualität und Quantität wissenschaftlicher Veröffentlichungen. Bibliometrische Aspekte der Wissenschaftskommunikation. - Jülich 2006.

Merkens, H. (Hg.) (2004). Evaluation in der Erziehungswissenschaft. - Wiesbaden: Verlag für Sozialwissenschaften 2004.

Nicolaisen, J. (2002). The J-shaped distribution of citedness. – Journal of Documentation, 58(4), 383-395.

Schneider, W. L. (1988). Grenzen der Standardisierbarkeit bei der Bewertung von Forschungsergebnissen. Einige Überlegungen aus der Sicht der Wissenschaftstheorie. - In: Daniel, H.-D.; Fisch, R. (Hg.) 1988, 433-447.

Stephenson, T. (2003). Paradigma und Pädagogik. Wissenschaftsanalytische Untersuchung im Spannungsfeld zwischen Pädagogik, Therapie und Wissenschaft. – Wien: Empirie Verlag 2003.

Stock, W.G. (2002b). Citation Consciousness. - In: Password, 2002, 6, 22-25.

Strohmer, J. (2008). Zitatählung am Fließband oder Analyse von Eigentext-Fremdtext-Relationen? Über die Verwendung von Zitaten in den Humanwissenschaften am Beispiel der Erziehungs- resp. Bildungswissenschaft (im Speziellen der Psychoanalytischen Pädagogik), der Psychologie und der Medizin: Ein Beitrag zu einer neuen Sicht auf die Thematik der Messung von Forschungsqualität. – Universität Wien: Diplomarbeit 2008.

Sykora, A. & Frühwirth, K. (2009): Die „Impact“-Faktoren psychoanalytisch-pädagogischer Texte: Zur Aufdeckung der Eigenart und Qualität psychoanalytisch-pädagogischer Forschung als Ergebnis des Projekts „Zitationsforschung“. In: Beiträge zur Zitations-, Rezeptions- und Diskursforschung - <http://thomas-stephenson.at>.

Weingarten, P. & Winterhager, M. (1984): Die Vermessung der Forschung. Theorie und Praxis der Wissenschaftsindikatoren. - Frankfurt a.M., New York: Campus Verlag 1984.

Autor

Univ.Do. Dr. Thomas Stephenson

Psychotherapeut und klinischer Psychologe, Mitglied des Staff der Sigmund Freud Privat-Universität, habilitierte sich mit einer wissenschaftstheoretischen Arbeit zur textanalytischen Paradigmenforschung im Überschneidungsbereich von Psychotherapie und Bildungswissenschaft.

thomas.stephenson@sfu.ac.at