

Die Bedeutung des Bradford's Law of Scattering für die Bibliothekswissenschaft

W. Umstätter

Ist das Bradford's Law of Scattering ein Gesetz oder nur eine
Regelmäßigkeit?

Die Bezeichnung „Bradford's law of scattering“ stammt von
Brian C. Vickery (1948)

Die Bibliothekswissenschaft kennt bisher nicht viele wirklich
nachweisbare Gesetzmäßigkeiten. Den Namen „Gesetz“ tragen
aber mehrere funktionalen Zusammenhänge.

„Bradford's Law of Scattering“?

„Power Law“?

„Zipf's Law“

„Lotka's Law“

In der Wissenschaft sind Naturgesetze Aussagen mit besonders hohem Wissensgehalt.

Ihre „a priori Redundanz“ erfordert

1. möglichst wenig Zufälligkeit (Verhältnis von realer Eintrittswahrscheinlichkeit beim Empfänger / zufälliger E. b. E.)
2. möglichst hohe Zuverlässigkeit (bei gleichen Voraussetzungen, gleiche Ergebnisse)
3. möglichst hohe Präzision (bezogen auf die Signaldefinition)
4. möglichst geringe Schwankungsbreite (Fehlerabschätzung)
5. möglichst gute Begründung (Abstützung auf andere Gesetze, und damit deren Vernetzung, wie in der Physik)
6. möglichst hohe thematischer Reichweite (z.B. Geltungsbereich des Gravitationsgesetzes)
7. möglichst hohe zeitliche Reichweite (in Zukunft oder Vergangenheit)

Das sind die 7 „Dimensionen“ des Wissens. Sie alle sind in Bit zu messen.

"Ein Ereignis ist dann kausal bedingt, wenn es mit Sicherheit vorausgesagt werden kann." ... "In keinem einzigen Falle ist es möglich, ein physikalisches Ereignis genau vorzusagen."

Max Planck (1948)

Bibliothekswissenschaftliche „Gesetze“:

1. Die Halbwertszeit von 5 Jahren (mit Jz. Gültigkeit)

(<http://www.ib.hu-berlin.de/~wumsta/pub18.html>)

2. Die Verdopplungsrate von 20 Jahren (mit Jh. Gültigkeit)

(<http://www.ib.hu-berlin.de/~wumsta/price14.html>)

3. Das Wachstum wiss.-tech. Probleme (mit Jt. Gültigkeit)

(<http://www.ib.hu-berlin.de/~wumsta/infopub/pub1991f/pub93.html>)

4. Das Zipfsche Gesetz

(<http://www.ib.hu-berlin.de/~wumsta/infopub/lectures/leasteffort03a.pdf>)

5. Die Interdisziplinarität des Wissens

(<http://www.ib.hu-berlin.de/~wumsta/lectg.html>)

Anwendungsbereiche des Bradford's Law of Scattering:

1. Zeitschriftenerwerbung
2. Zeitschriftennutzung
3. Zitation („Garfield's Law of Concentration“)
4. Interdisziplinaritätsbestimmung
5. Rechercheausgabe (Bradfordizing)
6. Erfassung von Zeitschriften in Bibliografien (Literaturdokumentation)
7. Entstehung neuer Zeitschriften
8. Entstehung neuer Sachgebiete
9. Wissenschaftsorganisation
10. Scientific collaboration (*scientific assembly line*)

Das Bibliothekswesen kennt seit langem die 80:20 Regel

~80% der Nachfrage sind mit ~20% des Bestandes abzudecken.

Auch Vilfredo Pareto beobachtete, dass 80% des Eigentums in Italien sich in der Hand von nur 20% der Italiener befand.

„Garfield’s Law of Concentration“

Grundaussage im BLS: eine immer gleiche Menge an Literatur verteilt sich auf gerankte Zeitschriften wie $n^0 : n^1 : n^2 : n^3 \dots$

Eigenes Beispiel:

Pro Jahr finden sich ~50 für uns relevante Aufsätze zum Thema „origin of life“ in der Zentralzeitschrift „Origin of Life“ und etwa ~30 Kernzeitschriften werden gebraucht um 100 zu finden, sowie etwa 27.000 (mehrere Datenbanken) um 200 zu finden.

Bei jeder natürlichen Beobachtung in der Fachliteratur würde ein Streit darüber, ob man beispielsweise mit $1 : 30 : 900 : 27.000$ oder $1 : 31 : 931 : 27.931$ rechnet, schon allein durch die auftretenden Streuungen ad absurdum geführt.

Hier, wie beim Ranking von Hochschulen, gilt:

„Der Mangel an mathematischer Bildung gibt sich durch nichts so auffallend zu erkennen, wie durch maßlose Schärfe im Zahlenrechnen.“ *C.F. Gauss.*

Es gibt bekanntlich schon in Bradfords Buch leicht abgewandelte Angaben zu seiner Beobachtung. Insbesondere im Kernbereich wurde daher später über den Groos droop (1967) diskutiert.

Funktionale Beschreibung des Bradford's Law of Scattering

$$p = k^{((a/c) - 1)} \quad a = c ((\ln p / \ln k) + 1)$$

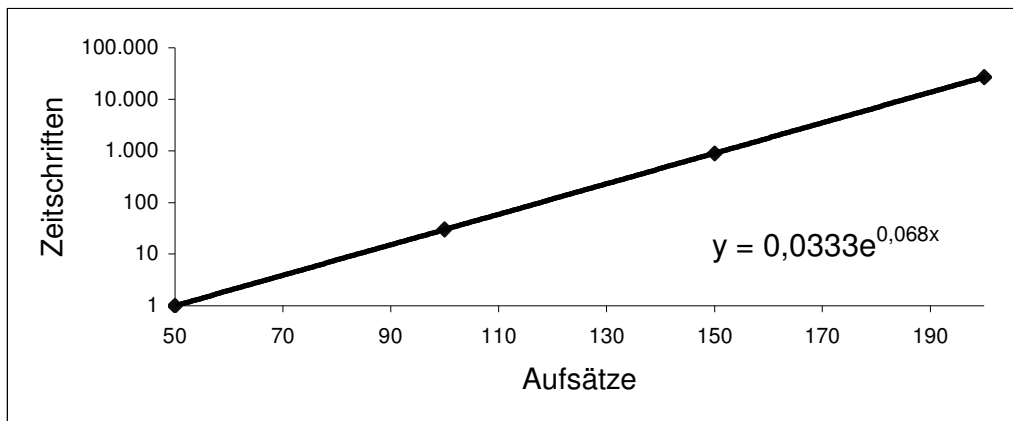
p = Anzahl der zu untersuchenden Periodika
a = Anzahl der Aufsätze in diesen Periodika
c = Anzahl der Aufsätze im 1. Periodikum (Zentralzeitschrift)
k = Anzahl der Kernzeitschriften, die wieder c Aufsätze enthalten.

c und k = Konstanten, die die Funktion bestimmen.

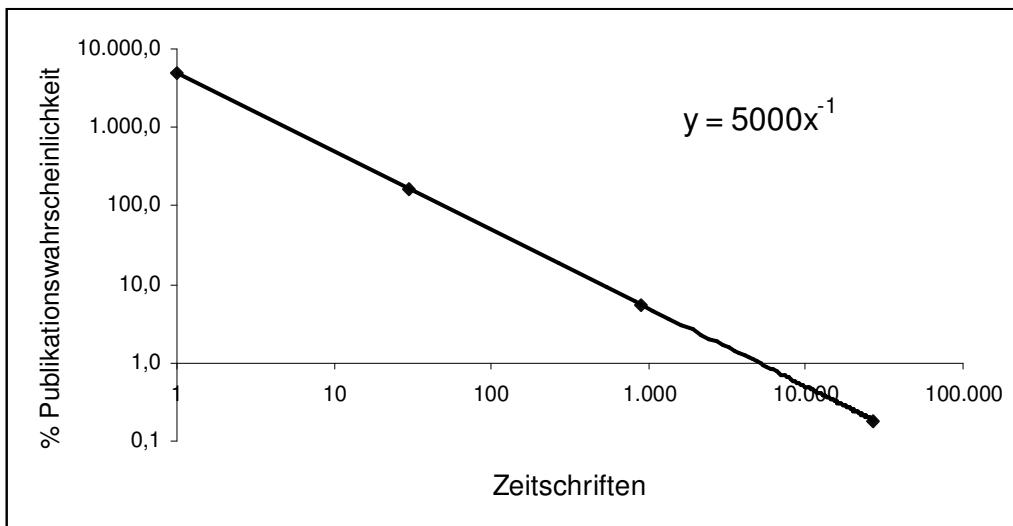
Beispiel für zwei Darstellungsformen:

A	B	C	D	E
50	50	n^0	1	5.000,0
50	100	n^1	30	166,7
50	150	n^2	900	5,6
50	200	n^3	27.000	0,2

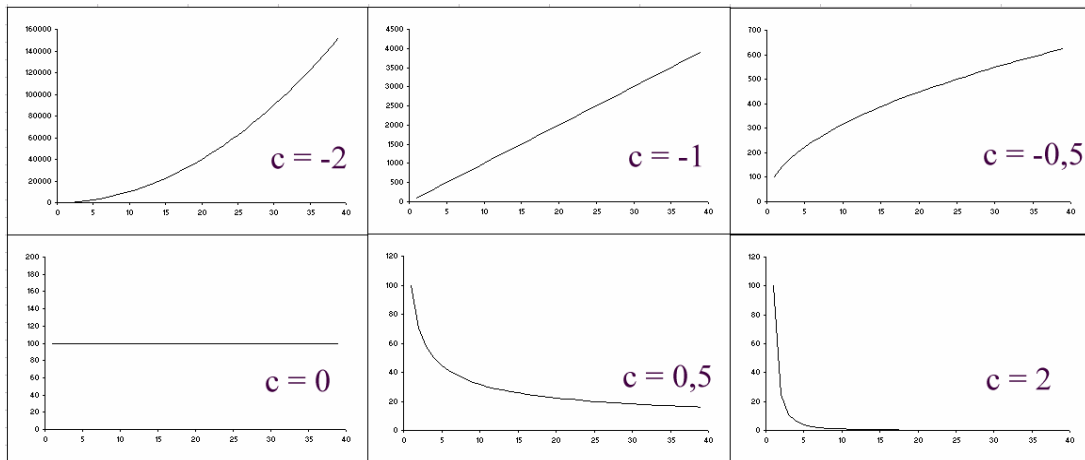
Darstellung von D zu B als **e-Funktion**



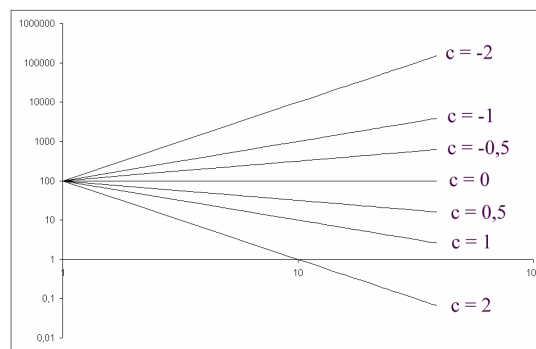
Darstellung von D zu E als **Potenz-Funktion (Hyperbel)**



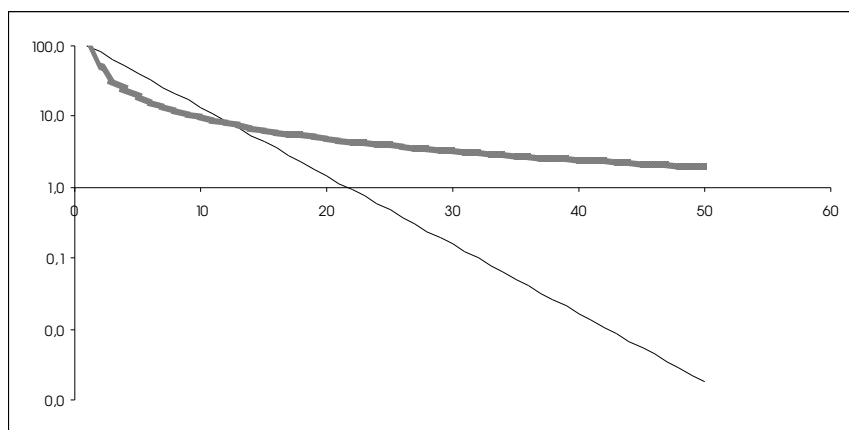
Potenz-Funktionen und e-Funktionen



Das Spektrum der Potenzfunktionen, je nach der Potenz c ($100 / x^c$)



Die Potenzfunktionen in doppelt-logarithmischer Darstellung.



Vergleich der e-Funktion (Gerade) und der Potenzfunktion in halb-logarithmischer Darstellung

Bradford's Law of Scattering gilt für:

Themen, *concepts, methods, problems, subjects* oder *words*.

Sie können immer auch durch eine eigene Benennung (ein Wort) zusammengefasst werden.

Bei *one term searches* werden Worte mit klar abgegrenzter Semiotik gesucht, wo geringe Homonymie und Synonymie herrscht.

Das Wort Bibliothek kann z.B. für eine Aufgabe, den Bau, die Bildungseinrichtung, juristische Einrichtung, Funktionseinheit, Konzeption, Organisation, das Problem etc. stehen.

Garfield hat das BLS auf „single disciplines“ und ebenso auf „science as a whole“ angewandt.

Beispiel: Eine Spezialbibliothek mit 900 laufenden Zeitschriften hat bei $c = 50$ Publikationen in der Zentralzeitschrift und $k = 30$ Kernzeitschriften in 20% von

$900 = 180$ Zeitschriften = $50 * (\ln 180 / \ln 30) + 1 = 126$ Aufsätze.

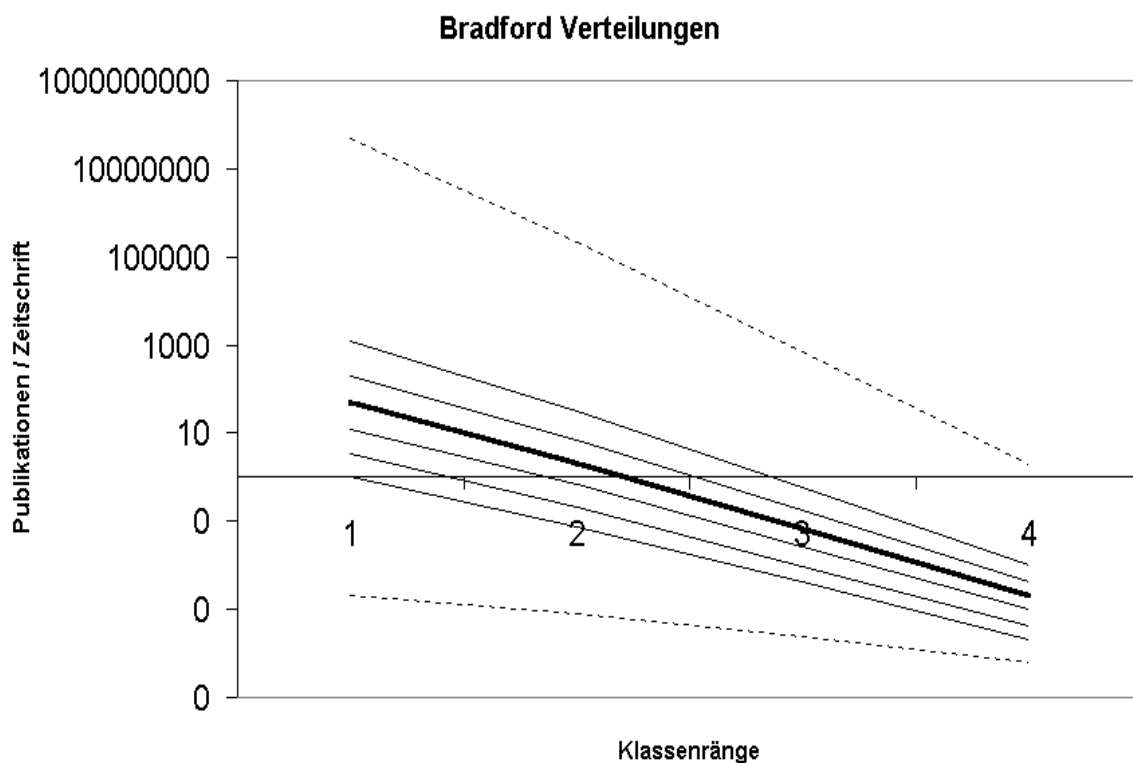
D.h. $126 / 150$ (150 Aufsätze in den 900 Zeitschriften) = 84% des Bibliotheksangebotes werden durch 20% Zeitschr. abgedeckt.

Bei 300 Zeitschriften würden 20% = 60 Zeitschr. = 82% abdecken.

Bei 100 Zeitschriften würden 20% = 20 Zeitschr. = 80% abdecken.

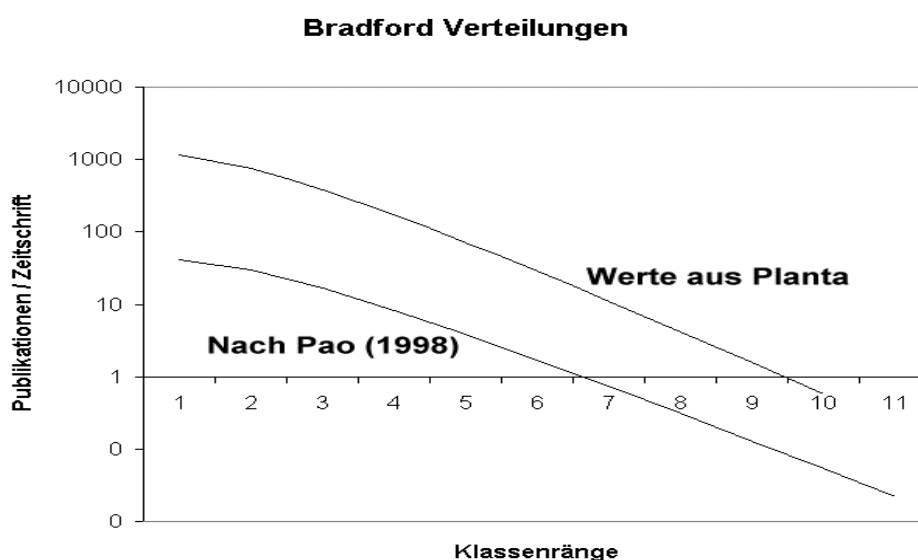
Verallgemeinerung des Bradford's Law of Scattering

Vergleicht man idealtypisch die Abnahme der Publikationen pro Zeitschriftentitel im Sinne des Bradford's Law of Scattering über einen Bereich, von 8 bis 200.000 Publikationen / J (-----) zu jeweils einem Thema, so zeigt sich, dass bei etwa 200 Publikationen / J (hervorgehobenen Kurve) eine neue Zeitschrift entsteht.



Das Bradford's Law of Scattering bei Zitierungen

Vergleich der Gesamtheit der Referenzen, die sich aus der Zeitschrift *Planta* ergeben, mit einem Beispiel von M. Pao aus dem musikalischen Themenbereich (*Concepts of Information Retrieval* Englewood, Colorado (1989))



Dass die Kurve im Kernbereich abgeflacht ist, beruht möglicherweise auf der Uncitedness III.

In anderen Fällen darauf, dass wir bei einem wachsenden Literaturangebot wählerischer werden.

Diese wiederholte Abweichung („Groos droop“) im Kernbereich kann als Widerlegung des BLS verstanden werden, aber auch als wichtiges Indiz für interpretierbare Systemabweichungen.

E. Garfield unterscheidet 3 Formen der Uncitedness

Uncitedness I: irrelevant

Uncitedness II: vergessen bzw. nicht gefunden

Uncitedness III: zu bekannt (*par excellance*)

Hinzuzufügen ist die vielleicht wichtigste Form der

Uncitedness IV: absichtlich ignoriert

Gründe für Uncitedness IV:

1. Die Falsifizierung der Publikation ist z.Z. nicht möglich.
2. Die Falsifizierung einer Publikation kann in der gebotenen Kürze nicht erfolgen.
3. Die Publikation ist aus vielfältigen Gründen unverständlich.
4. Man müsste eigene Fehler zugeben.
5. Man möchte bestimmte Autoren bzw. Institutionen möglichst nicht aufwerten.
6. Man ignoriert anscheinend oder scheinbar minderwertige Wissenschaft (z.B. Zeitschriften der Dritten Welt)
7. Die enthaltene Theorie enthält ebenso viel richtiges wie fehlerhaftes.
8. Die enthaltene Methode ist zu stark abgewandelt worden.
9. Man möchte fehlerhafte Hypothesen aussterben lassen. Planck: Fehlerhafte Theorien werden nicht widerlegt, sie sterben aus.
10. Man möchte die eigene Arbeit nicht durch zu viele Gegenargumente belasten. Nur 7% der Zitationen im SCI sind negativ bzw. Falsifikationen. Der Rest soll belegen, dass die eigene Theorie richtig ist.

Modell:

Wir haben es heute mit

~100.000 laufenden Zeitschriften und

~100.000 Spezialgebieten, die von

10 Mio. Wissenschaftlern betreut werden,

zu tun. (Das entspricht 100 Wissenschaftler / Spezialgebiet)

Wenn jede wissenschaftliche Person rund 100

Kernzeitschriften regelmäßig sichtet, dann werden diese im

Durchschnitt von (1 Mrd. Durchsichten / 100.000 Zeitschr.)

~10.000 Wissenschaftlern weltweit gesichtet. (nicht gelesen!)

Ein Wissenschaftler schaut

~10.000 Titel / J durch. ($10.000 * \sim 1 - 2 \text{ Min.} = 200 \text{ h}$)

Er liest genauer

~100 Publikationen / J ($100 * \sim 2 \text{ h} = 200 \text{ H}$)

Das sind 20% von 2.000 Arbeitsstunden / J = 400 h für das

Literaturstudium.

Wiederholt kritisch studiert werden nur ~10 Quellen / J.

Beispiel: Seit dem ich „Little Science Big Science“ von D.J. de Solla Price wiederholt gelesen habe, sah ich mich über die Jahrzehnte hinweg gezwungen immer mehr zu falsifizieren.

Das Bradford's Law of Scattering ist eine funktionale Darstellung der Interdisziplinarität.

Während NATURE oder SCIENCE von fast allen

Naturwissenschaftlern frequentiert werden, gibt es Tausende von Zeitschriften, die nur wenige Leser haben.

Wenn Hjørland und Nicolaisen meinen, dass „the social sciences are generally considered very interdisciplinary“, so müsste die Konstante c vergleichsweise klein und k entsprechend groß sein. Dies ist aber schwer nachweisbar, da in den Sozialwissenschaften Monografien stark dominieren

Fast alle Disziplinen haben den Eindruck besonders Inter-, Trans- oder Multidisziplinär zu sein.

Eine durchschnittliche Disziplin hat in einer deutschen UB ~150 Zeitschriftentitel. Das dürfte etwa die Grenze der wissenschaftlichen Kapazität eines Spezialisten sein.

Aus dieser Begrenzung menschlicher Kapazität ergibt sich die „Gesetzmäßigkeit“ in der Bibliothekswissenschaft.

Das BLS ist so gesehen ein Maß für die *collaboration* in der Wissenschaft.

Hjørland, B. and Nicolaisen, J.: Context: Nature, Impact, and Role,
 Proceedings Lecture Notes in Computer Science 3507: 96-106
 (2005)

“This paper introduces a distinction between Lexical scattering,
 Semantic scattering, and Subject scattering.”

“However, it is hard to find actual reports that describe how
 Bradford's law has been applied in practical library and
 information services.”

Das „lexical scattering“ von Hjørland und Nicolaisen folgt nicht dem
 „Zipf's law“, weil dort Worte in ihrer Syntax gesucht werden,
 weil beim BLS Worte über Zeitschriften verteilt sind.

Die „Gesetzmäßigkeit“ des Zipfschen Gesetzes ergibt sich bekanntlich
 aus dem „*principle of least effort*“. Beim BLS geht es aber
 grundsätzlich um die Verteilung von Begrifflichkeiten auf
 Dokumente.

Hjørland und Nicolaisen: “Umstätter, 2005, who misunderstood
 what was claimed in the paper he discusses”
 This can be proved. *Umstätter*

Die Aussage „nobody have thus far tried to outline the consequences
 of different conceptions of 'subjects' for Bradford's law.“
 übersieht die Diskussion über Relevanz, Precision, Recall ratio,
 Noise und Pertinenz.

Bei „Onlinern“ wird das „Bradfordizing“ häufig eingesetzt. Es
 erleichtert nicht nur das Auffinden von Aufsätzen, es hilft auch
 Anfängern rasch zu erkennen welche Zeitschriften für sie
 besonders wichtig sind.