

# ***Das Forum for Negative Results:*** **Warum wir Forschungs-Misserfolge nicht totschweigen sollten**

Lutz Prechelt  
Freie Universität Berlin, Institut für Informatik

*Faktenstand: 2014-01-30*

- Wissenschaft, altes Modell
- Wissenschaft, neues Modell
  - Falsifikationismus
- Technik
- Rolle von Komplexität
  - Schubladenproblem
  - Plausible-Idee-Problem
  - Folgen in der Informatik
- Publikationsorte für negative oder unklare Resultate

# Wissenschaft, altes Modell

Herkunft von Wissen:

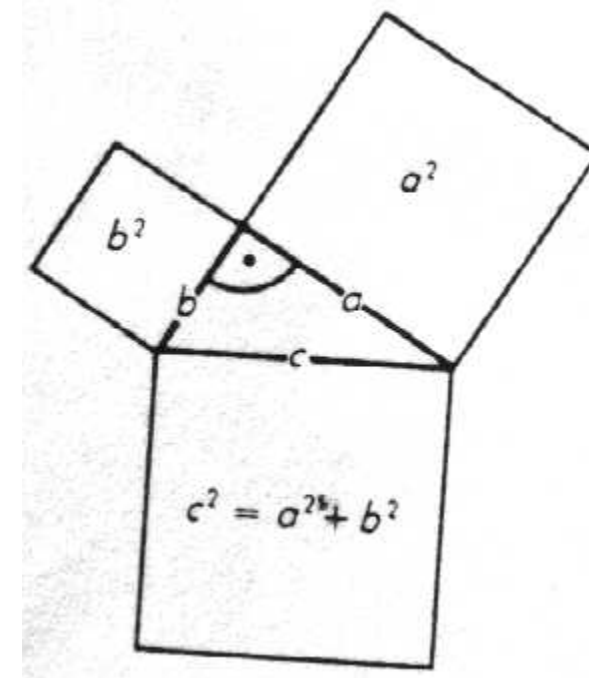
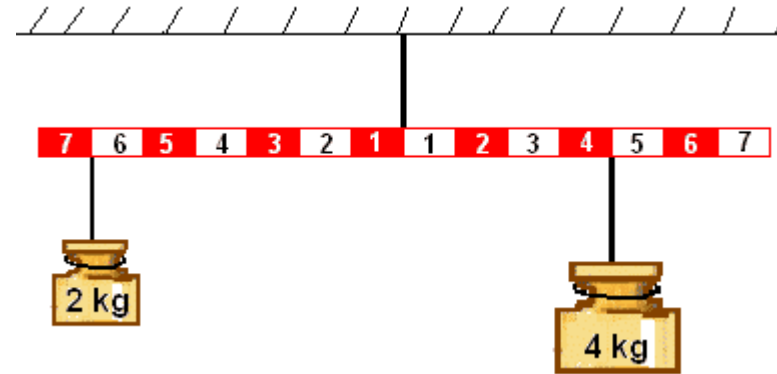
- Tradition
- Autoritäten



# Wissenschaft, neues Modell

## Herkunft & Natur von Wissen:

- Nachvollziehbare Beobachtungen
- Überprüfbare Erklärungen (Modelle, Theorien)



1 18

*Periodensystem der Elemente*

1 H Wasserstoff 1.01	2 He Helium 4.00											13 B Bor 10.81	14 C Kohlenstoff 12.01	15 N Stickstoff 14.01	16 O Sauerstoff 15.999	17 F Fluor 18.998	18 Ne Neon 20.18
3 Li Lithium 6.94	4 Be Beryllium 9.01											5 Al Aluminium 26.98	6 Si Silicium 28.09	7 P Phosphor 30.97	8 S Schwefel 32.07	9 Cl Chlor 35.45	10 Ar Argon 39.95
11 Na Natrium 22.99	12 Mg Magnesium 24.31	3 Sc Scandium 44.96	4 Ti Titan 47.88	5 V Vanadium 50.94	6 Cr Chrom 52.00	7 Mn Mangan 54.94	8 Fe Eisen 55.85	9 Co Cobalt 58.93	10 Ni Nickel 58.70	11 Cu Kupfer 63.55	12 Zn Zink 65.41	13 Ga Gallium 69.72	14 Ge Germanium 72.64	15 As Arsen 74.92	16 Se Selen 78.96	17 Br Brom 79.90	18 Kr Krypton 83.80
19 K Kalium 39.10	20 Ca Calcium 40.08	21 Sc Scandium 44.96	22 Ti Titan 47.88	23 V Vanadium 50.94	24 Cr Chrom 52.00	25 Mn Mangan 54.94	26 Fe Eisen 55.85	27 Co Cobalt 58.93	28 Ni Nickel 58.70	29 Cu Kupfer 63.55	30 Zn Zink 65.41	31 Ga Gallium 69.72	32 Ge Germanium 72.64	33 As Arsen 74.92	34 Se Selen 78.96	35 Br Brom 79.90	36 Kr Krypton 83.80
37 Rb Rubidium 85.47	38 Sr Strontium 87.62	39 Y Yttrium 88.91	40 Zr Zirkonium 91.22	41 Nb Niobium 92.91	42 Mo Molybdän 95.94	43 Tc Technetium (98)	44 Ru Ruthenium 101.07	45 Rh Rhodium 102.91	46 Pd Palladium 106.42	47 Ag Silber 107.87	48 Cd Cadmium 112.41	49 In Indium 114.82	50 Sn Zinn 118.71	51 Sb Antimon 121.76	52 Te Tellur 127.60	53 I Iod 126.90	54 Xe Xenon 131.29
55 Cs Cäsium 132.91	56 Ba Barium 137.33	La-Lu	72 Hf Hafnium 178.49	73 Ta Tantal 180.95	74 W Wolfram 183.84	75 Re Rhenium 186.21	76 Os Osmium 190.23	77 Ir Iridium 192.22	78 Pt Platin 195.08	79 Au Gold 196.97	80 Hg Quecksilber 200.59	81 Tl Thallium 204.38	82 Pb Blei 207.2	83 Bi Bismut 208.98	84 Po Polonium (209)	85 At Astat (210)	86 Rn Radon (222)
87 Fr Francium (223)	88 Ra Radium (226)	Ac-Lr	104 Rf Rutherfordium (261)	105 Db Dubnium (262)	106 Sg Seaborgium (263)	107 Bh Bohrium (262)	108 Hs Hassium (265)	109 Mt Meitnerium (266)	110 Ds Darmstadtium (281)	111 Rg Roentgenium (272)							

© Peter Wich - Experimentalchemie.de - Chemie erleben!

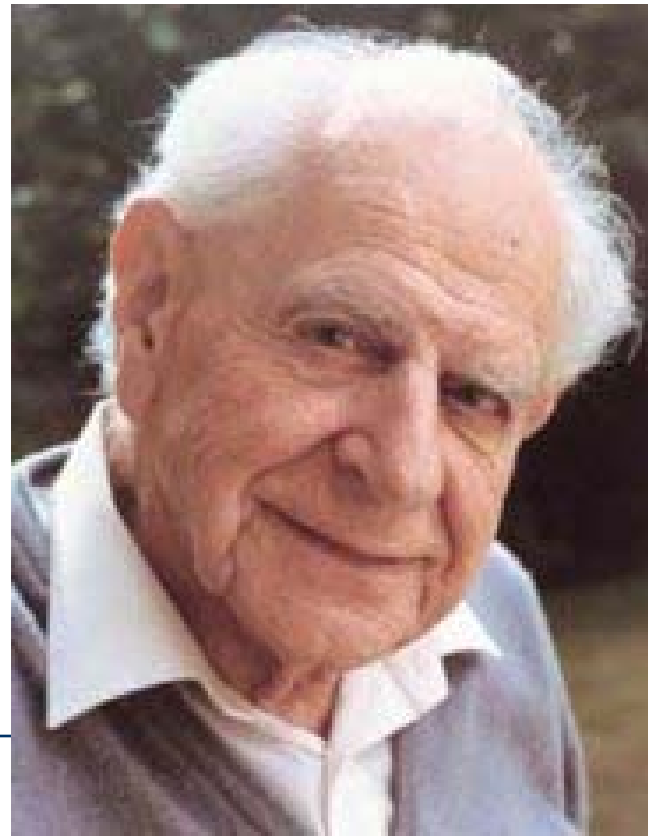
57 La Lanthan 138.91	58 Ce Cer 140.12	59 Pr Praseodym 140.91	60 Nd Neodym 144.24	61 Pm Promethium (147)	62 Sm Samarium 150.36	63 Eu Europium 151.97	64 Gd Gadolinium 157.25	65 Tb Terbium 158.93	66 Dy Dysprosium 162.50	67 Ho Holmium 164.93	68 Er Erbium 167.26	69 Tm Thulium 168.93	70 Yb Ytterbium 173.04	71 Lu Lutetium 174.97
89 Ac Actinium 227.03	90 Th Thorium 232.04	91 Pa Protactinium 231.04	92 U Uran 238.03	93 Np Neptunium (237)	94 Pu Plutonium (244)	95 Am Americium (243)	96 Cm Curium (247)	97 Bk Berkelium (247)	98 Cf Californium (251)	99 Es Einsteinium (252)	100 Fm Fermium (257)	101 Md Mendelevium (258)	102 No Nobelium (259)	103 Lr Lawrencium (262)

# Wissenschaft: Rolle negativer Beobachtungen

Karl Popper:

- Wissenschaftlich sind nur Aussagen, die widerlegt werden können
  - "Experiment"
  - Trennt W. von Religion
- Theorien können nicht bewiesen werden
  - allenfalls gestützt
- Aber sie können widerlegt ("falsifiziert") werden
  - Einmal reicht völlig!
- Möglichkeit 3:
  - Experimentresultat bleibt mehrdeutig, also unklar

- Experimentresultate sind also entweder
  - positiv: stützend
  - negativ: widerlegend
  - unklar: weder noch



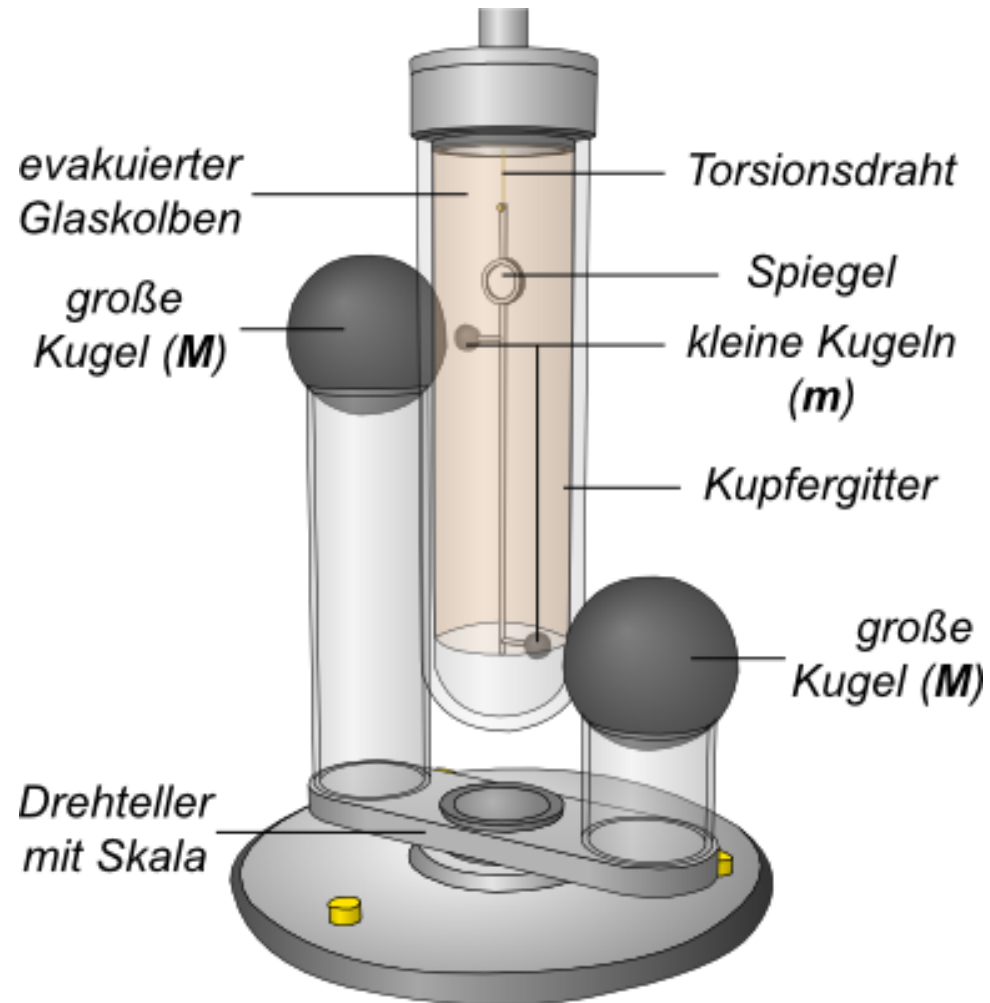


# Beispiel: Gravitationsgesetz

- Person A behauptet

$$\vec{F}_{m_1} = G \frac{m_1 m_2}{r^2}$$

- Alle anderen glauben das nicht
- A erfindet ein überzeugendes Experiment
- Am Anfang sind die Ergebnisse unklar
- Irgendwann ist
  - das Experiment ausgereift,
  - sein Ergebnis stützend und
  - die Zweifler lassen sich überzeugen



- Ähnliche Arbeitsweisen beim Wissenserwerb
  - Dieser ist aber nicht das Ziel
  - sondern es wird **Nützlichkeit** angestrebt.
  - Deshalb gehen die Experimente oft weit über den Wissensstand hinaus
- z.B. Ingenieurwesen
  - z.B. Medizin



- Wenn die Situation komplex wird
  - bei Naturwissenschaft oft der Fall (z.B. Biologie, Geologie, Psychologie u.a.)
  - bei Technik sogar fast immer (z.B. Bauingenieurwesen, Mikroelektronik, Medizin, Software Engineering u.a.)
- treten viel öfter negative oder unklare Ergebnisse auf als erhoffte
- weil man die Zusammenhänge zu wenig überschauen kann
- **Wissensfortschritt sollte sich deshalb nicht nur aus positiven Ergebnissen speisen**

# Hauptproblem bei Wissenschaft: Das Schubladenproblem (1)

- In den meisten Bereichen (z.B. allen mit Menschen) schwanken Experimentergebnisse probabilistisch
- Beispiele:
  - Wirksamkeit eines Medikaments (gegenüber Placebo)
  - Wirksamkeit einer Psychotherapie (gegenüber einer Etablierten)
  - Existenz von Hellseherei (gegenüber Nichtexistenz)
  - Die Geschwindigkeit von Neutrinos (gegenüber Lichtgeschw.)
- Klare Ergebnisse erreicht man nur (und nur annähernd) über statistische Signifikanz:
  1. Ich rechne nach: Falls der behauptete Effekt nicht existiert, müsste das beobachtete Gesamtergebnis unwahrscheinlich sein
    - Wie unwahrscheinlich es ist, berechnet die inferentielle Statistik
    - Nun setzt man eine willkürliche Schwelle, meist 5% oder 1%, und sagt:
  2. Sehe ich das Ergebnis trotzdem, glaube ich an den Effekt

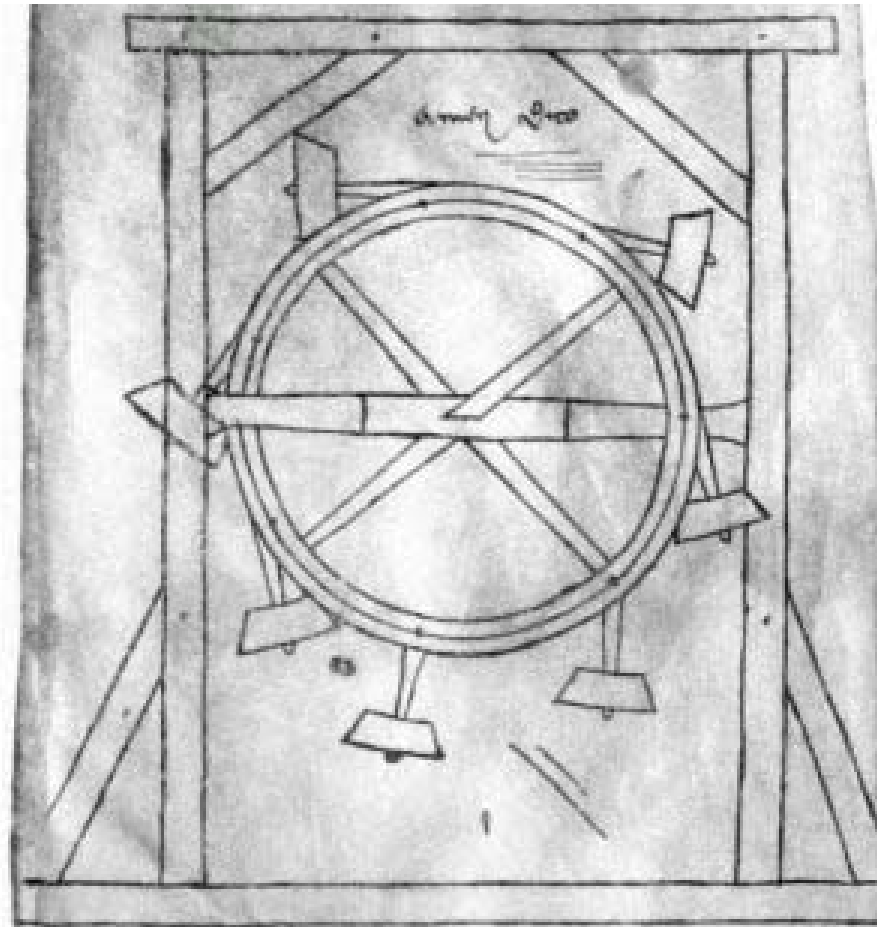


# Hauptproblem bei Wissenschaft: Das Schubladenproblem (2)

- Doch auch unwahrscheinliche Ergebnisse treten gelegentlich ein!
  - Wenn also viele Experimente den selben nicht vorhandenen Effekt prüfen
  - und nur die wenigen zufällig stützenden publiziert werden
    - der Rest wandert in die Schublade
  - dann bildet man sich ein positives Gesamtergebnis ein
- (Beachte: Negative Ergebnisse gibt es hier selten, sondern meist nur stützende oder unklare
  - Ein unklares kann am Fehlen des Effekts liegen oder an einem zu schwachen Experiment.
  - Das kann man aber nicht unterscheiden
    - falls nicht auch noch eine Effekt**stärke** behauptet war)







# Hauptproblem bei Technik: Das Plausible-Idee-Problem

- Ideen auf dem Weg zu Nützlichkeit:  
"So müsste es doch gehen!"
  - Viele Leute haben den gleichen Gedanken
    - und probieren ihn mit evtl. hohem Aufwand aus
  - Manche früher, manche später
- Wenn es aber so nicht geht und das nicht publiziert wird, muss jeder seine schlechte Erfahrung nochmal neu machen
  - Riesenverschwendung!
- Hier sind neben unklaren auch oft negative Ergebnisse möglich



## 2 Hauptmodi:

- Konstruktion
  - gehört hauptsächlich zur Nützlichkeits erzeugung
- Empirie
  - gehört eigentlich hauptsächlich zur Wissenserzeugung
- Aber: Wissenschaft und Technik sind daran nicht verlässlich zu unterscheiden
  - Fred Brooks in CACM 39(3), 1996:  
**"The scientist builds in order to study;  
the engineer studies in order to build."**
- Beide Modi haben also Ihren Platz, aber Endzweck müsste immer Erzielung und Nachweis von Nützlichkeits sein.

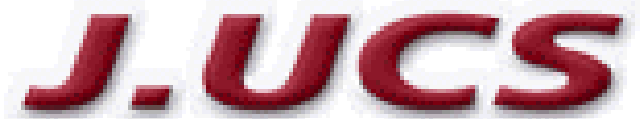
- Konstruktion: Erfindung oder Weiterentwicklung
  - A: ganz neuartig, deshalb keine Nutzenmessung 
  - B: Nutzen verbessert und das demonstriert 
  - C: Nutzen nicht verbessert oder unklar/schöngeredet 
  - D: Nutzenmessung fehlt
- Empirie: reine Untersuchung, ohne Erfindung/Entwicklung
  - E: neutrale Wissenserzeugung (ohne Vorerwartung) 
  - F: erwartetes Ergebnis eingetreten 
  - G: unerwartetes oder unklares Ergebnis eingetreten 

- Im Nachhinein erscheint ...
  - ...ein Experiment mit unklarem Ausgang oft schwach
  - ...eine Idee, die nicht funktioniert hat, oft als unplausibel.
- Deshalb lehnen Gutachter solche Publikationen oft ab,
- deshalb reden Autoren Ihre Studien oft schön
- und die oben beschriebenen Probleme treten kräftig auf
- Ein Kulturwechsel tut not:



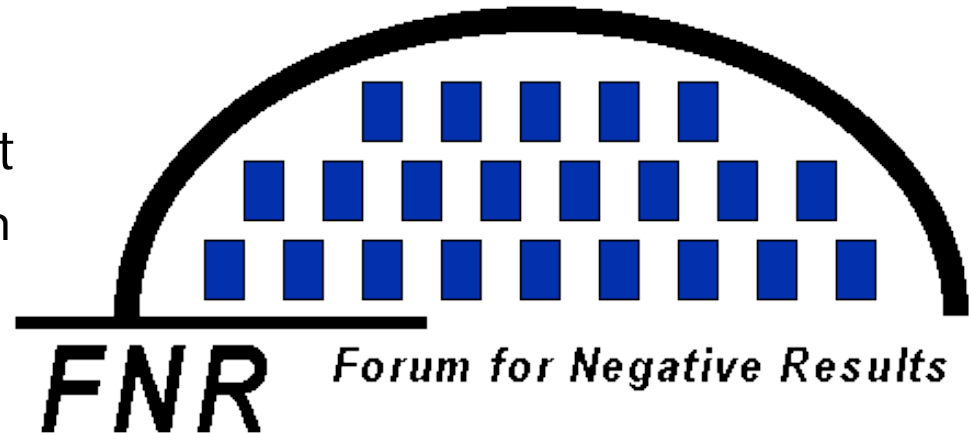
## J.UCS:

- gegründet 1994 von Maurer (Graz)
- eine der ersten elektr. Zeitschriften
  - recht gute Qualität
- anfangs Springer Verlag, jetzt Open Access
- breites Themenspektrum der ganzen Informatik



## FNR:

- Gegründet 1997 von Prechelt
  - "a permanent special section of J.UCS"
- Sucht explizit Beiträge mit unerwarteten oder negativen Resultaten
  - mit zumindest einer teilweisen Erklärung dazu



# FNR Akzeptanzkriterien

1. Klares und explizites Forschungsziel
2. Gute und plausibel begründete Erfolgchance
3. Forschung durchgeführt, Ziel dennoch nicht erreicht
  - Und das wurde erst bei der Evaluation sichtbar
4. Gefahr, dass jemand den gleichen Irrweg nochmals beschreitet
5. Durchführung und Auswertung haben hohe Qualität
6. Misserfolgsgründe sind zumindest teilweise verstanden und erklärt

- In welchen Ländern hat FNR hauptsächlich Erfolg?
- In welchen nicht?
- Wie viele Einreichungen hat es pro Monat?

## Medizin/Biologie:

- Journal of Errology
- Journal of Negative Results in Ecology and Evolutionary Biology
- Journal of Negative Results in BioMedicine
- Journal of Negative Observations in Genetic Oncology

## Psychologie:

- Journal of Articles in Support of the Null Hypothesis

## Sozialwissenschaften:

- Journal of Spurious Correlations

## Diverse:

- Journal of Unsolved Questions
- The All Results Journals
  - Biol, Chem, Nano, Phys

## Mathematik:

- *Rejecta Mathematica*

## Informatik:

- J. of Interesting Negative Results in Natural Language Proc. & Machine Learning
- Journal of Negative Results in Speech and Audio Sciences
- Forum for Negative Results

- Vorstellung folgt jetzt in Reihenfolge des Alters

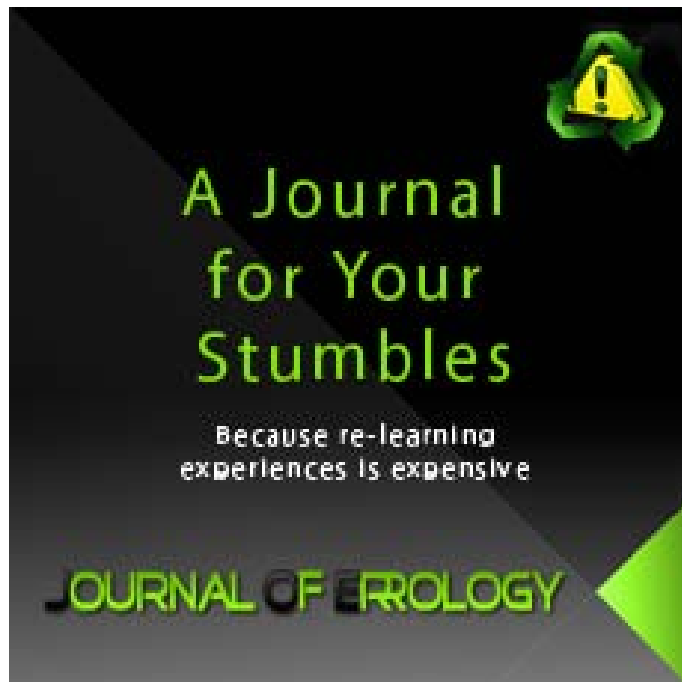


- <http://www.journalofnullresults.com>
  - Gegründet: noch gar nicht?
  - 0 Artikel
- 
- <http://nullresults.wordpress.com/>
  - Gegründet 2010
  - 0 Artikel
- Januar 2012: Webauftritt ist ein unmodifiziertes MediaWiki
    - d.h. ohne eigenen Inhalt
  - Januar 2014: Domain existiert nicht mehr
- 
- Webauftritt ist ein Wordpress-Blog mit nur einem "Hello, World!"-Eintrag.

- <http://junq.info>
- Gegründet 2011
- 5 Artikel bis Jan 2012
- Jan 2014: Homepage geht, Restfunktionen nicht
- *Our vision is establishing the publication of null-results as an important cornerstone for the advancement of knowledge and scientific understanding in all disciplines thus contributing to overcome biases and fraud in research.*
  - für unfertige Forschung



- <http://bioflukes.com/JoE>
- Gegründet 2011
- 2011-2014: 3 Artikel



- *JoE [...] enables sharing and discussions on those unpublished futile hypothesis, errors, iterations, negative results, false starts and other original stumbles that are part of a larger successful research in biological sciences.*

- <http://www.arjournals.com>
- Gegründet 2010
  - 4 Ausgaben für verschiedene Fächer

2010-2014:

- Biol: 10 Artikel
- Chem: 7 Artikel
- Nano: 0 Artikel
- Phys: 1 Artikel

- *The main objective of The All Results Journals focuses on recovering and publishing negative results, valuable pieces of information in Science.*



THE ALL RESULTS JOURNALS

- <http://math.rejecta.org>
- Gegründet 2009
- 2009 bis 2011: 12 Artikel
- Januar 2014:  
Artikelabruf defekt
- Rejecta Mathematica is an open access, online journal that publishes only papers that have been rejected from peer-reviewed journals (or conferences with comparable review standards) in the mathematical sciences.
  - mapping blind alleys, insights on known results, interesting despite a flaw, controversial premises





- <http://jinr.site.uottawa.ca/>
- Gegründet 2008
- 2008 bis 2011: 1 Artikel
  - aber 17 Mitglieder im Editorial Board
- *Papers present thoroughly conducted experiments in which an intuitive and plausible idea has failed.*
  - Das entspricht FNR

J I - R

- <http://www.jspurc.org>
- Gegründet 2005
- 2005 bis 2011: 0 Artikel
  - aber eine Liste diverser Artikel über JSpurC
- *The mission of JSpurC is to provide a legitimate venue for exploring pure and applied methodological questions in the social sciences in the company of colleagues without fear of professional embarrassment or reprisal.*

## THE JOURNAL OF SPURIOUS CORRELATIONS

*Qualitative and Quantitative Results in the Social Sciences*

- <http://www.jnr-eeb.org/>
- Gegründet 2004
- Publizierte Artikel:
  - 2004-2010:  
2, 2, 1, 2, 2, 1, 1
  - 2011-2014 (Januar):  
2, 2, 1, 0
- *The primary intention of Journal of Negative Results is to provide an online-medium for the publication of peer-reviewed, sound scientific work in ecology and evolutionary biology that may otherwise remain unknown.*

JOURNAL OF NEGATIVE RESULTS

- ECOLOGY & EVOLUTIONARY BIOLOGY -

- <http://www.jnrsas.org>
- Gegründet 2003
- *January 2012:*  
*DB Error: Can't connect to local MySQL server through socket*  
*'/var/lib/mysql/mysql.sock'*  
*(111)*
- *Januar 2014:*  
*Verzeichnisliste mit*  
*\_vti\_bin/, \_vti\_inf.html,*  
*cgi-bin/*  
*aber keine Inhalte*

- <http://www.jnrbm.com/>
- Gegründet 2002
- 2002 bis 2011: ~105 Artikel
- 2012 bis 2013: 36 Artikel
- *JNRBM is an open access, peer-reviewed, online journal that promotes a discussion of unexpected, controversial, provocative and/or negative results in the context of current tenets.*
  - Sehr breit also



JOURNAL OF NEGATIVE  
RESULTS IN BIOMEDICINE



- <http://www.jasnh.com/>
  - Psychologie
- Gegründet 2002
- 2002-2013:  
7, 4, 3, 1, 4, 2, 1, 2, 3, 7,  
4, 5 Artikel
- *We [...offer] an outlet for experiments that do not reach the traditional significance levels ( $p < .05$ ). Thus, reducing the file drawer problem, and reducing the bias in psychological literature.*



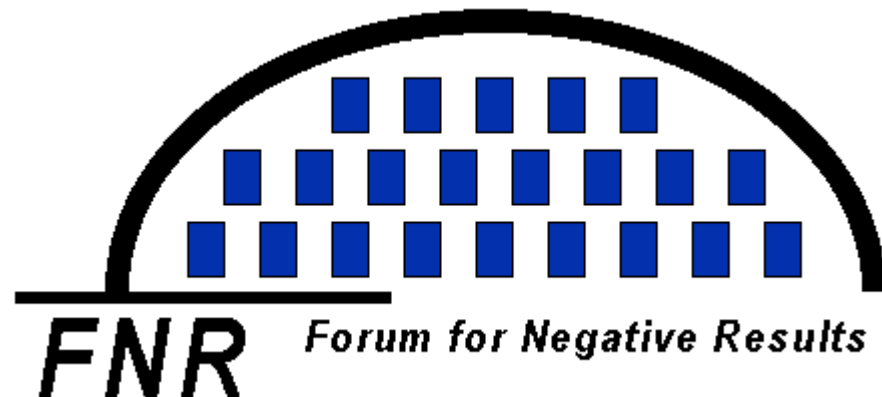
- <http://www.path.jhu.edu/NOGO/>
- gegründet ca. 1997
  - Mehr eine Datenbank als eine Zeitschrift
  - Als einziges Organ etwas älter als FNR
- Hatte wenig Einreichungen
- Website nicht mehr erreichbar
  - `/UNDEF/SEND+7^WEBUTLTY:::4:0:***` *WebLink: HTML Error*
  - offenbar 2008 eingestellt
  - vermutlich ist Gründer Scott Kern im Ruhestand

# Forum for Negative Results (in the J. of Universal Computer Science)

- <http://page.mi.fu-berlin.de/prechelt/fnr/>
- Gegründet 1997
- Also nochmal die Frage:
  - Wie viele Einreichungen?
  - Welche Länder?
- 1997-2014: 1 Einreichung
  - sehr gut
  - erschienen Ende 2012



**J.U.C.S.**



- siehe FNR Homepage:  
<http://page.mi.fu-berlin.de/prechelt/fnr>
- L. Prechelt: Why we need an explicit forum for negative results.  
J.UCS 3(9), 1997
- <http://themonkeycage.org/blog/2011/06/26/clairvoyance-and-really-bad-journal-policies/>
  - Experiment mit Ergebnis "Hellsehen funktioniert" wurde 2011 publiziert, die Widerlegung dann aber vom selben Journal dreimal abgelehnt, weil Wiederholungsstudien dort grundsätzlich nicht erscheinen.

**Danke!**