Westfälische Wilhelms-Universität Münster

Fraudulent results and failed replications: The effect of flawed research on the public's trust in science





Vertrauen und Kommunikation in einer digitalisierten Welt

Rainer Bromme Institut für Psychologie bromme@uni-muenster.de

BE DE

Special thanks to the following members of our (former) group:



Prof. Dr. Marc Stadtler (Co –PI within the SPP 1409)



Dr. Lisa Scharrer



Dr. Eva Thomm



Dr. Dorothe Kienhues





Trust and Communication in a Digitized World

SCIENCE

Dr. Friederike Hendriks



Kinds of problems which might impact on public trust in science

Scientific misconduct with the intention to decieve

- Fabrication/ falsification of results (for example D. Stapel, Netherlands, Psychology; P. Macchiarini, Sweden, Medicine; A. Wakefield, Medicin,GB)
- plagiarism (for example two former German secretaries of state (defence /science & education)

Failed replications, because

- of questionable research practices
- weakness of theory and/or methods and /or data of the original study or of the replication
- the issue under study is not consistent across time, culture, contexts

	SAKL7 REPORT
Early report	
leal-lymphoid-podular hyperplasis	non-specific colitis and
rear-tympholu-nouular hyperplasia	a, non-specific contis, and
pervasive developmental disorder	r in children
l J Waxafala, S H Murch, A Anthony, J Lensel, D W Cases 1 Harvey, A Volontine, S E Device, J A Wolker Omith	er, M.Maile, M.Banacseritz, A.P.Chillon, M.A.Thomasa,
lummary	Introduction
achipsent he restigated a consecutive series of	We saw several children whe, after a target of apparent
Alber with shorte anterceditic and regressive	They of had gestreintesteal employee, Volng
bretarrerts deorder.	abdominal pair, dischoos, and being and, sponse
Rethods 12 cretoren (moon age 6 years (range 3-20), 11	sance, fixed intoisance. We receive a clinical fittings,
the product of parent devices billing of the	
cound skills, including language, together with diarroos	Patients and metry ts
nd abdominal pass. Choices underwart	12 shillers, considered and is department of
perserversigical, neurological, and developmental	developmental after with loss and the dails and movements
accepted and block satisfies, manufic-memory	symptoms for addression on, Noning and Sec
maging (MOE), electroscophalography (EEG), and Uniter	want lowe street, according to 1 by Forte parameter.
wheture were done under sedation. Barlum fution through	
adiography was done where possible. Diochemical,	took historia including deads of interioritations and
withher and interesting a proving with	expenses to infect on discourse, and assessed the children. In 11
and and the second second second second second	can be busined as obtained by the armor constant (W-S). Next, and the parchieter parameters were denot by
a the section, with measing, margan, and such	combart stat (298, MR) with HMS-4 criteria." Developmental
accristion is eight of the 12 children, with stees	provided a review of prospective developmental receptor
election in one child, and atitis media in space. All 1	shilden del ore underge prediate's assessment in hospital, all
hiden had intestinal abnormalities long from	had been arrented preferringally chereleging, or these assumed to train used in the basis for their behaviored forwards.
interingy showed patienty churche and an and	After heard preparation, ilosedistancepy was performed by
11 children and reactive illeger monor conclusie in	SID4 or MXT under sudgets with midgelast and petitidate based from and formula fauri moreced binner contribution
even, but no granulemes. But yoursi diade as included	fation from the terminal dentity sameding, transverse,
Miler (arte), signing diversities and (arte), or eventing	descending, and approach (white, and been the mittage, the
scal neurological and mailties and and EES tests	compared with images of the provinus seven consecutive
ere normal. Above al laboratory results the significantly	pertiantic columnorphis from normal releasespics and three
alsod urhay exploses and compared with ap-	reported method appendixies in the serviced leven. Bachart
Nilliar, and for the life in the shidow.	Educe through radiography was possible in some cases.
	(MRC), electromorphilisprephy (REC) including visual, breat
to a and explanmental regression in a girty of	more auditory, and sensory avalant presentate (where compliance much three manifold, and hardway terrotomy wave down
which was generally seasciated	non en panie, en ann panie en an.
time pessisie environmental triggers.	Laboratory investigations
ance: 190-191.037-41	archrogital-fuid incare wars manufed to carlade laws
tee Contrient by Dage	searce of childhood neurodigenerative disease. Urinary
dammatory Sinusi Disease shaty Cesure. University Cesartments	motoymations and we measured in random arise samples from sides of the 12 shifteen and 13 are matched and we method
Weditine and Histopathology (4.) Waterfatt Fich, 8. Arthory 10,	normal scenesis, by a multilation of a technique contribut
Environments Are constant entries, a C Disition weather, and the advantation Departments of Pandistric, Santonermanings	permaner/ Commitgeness are sound lightly of
3 In Marghton, 3 M Cosser unpri M Malik enery	and controls. Urmany methodosic-acid concentrations in
herbidey (M Benderda Hitter), Neurology (7 Hanas Hitt), and	pateros and contrais were compared by a two-sample r and.
indexings (A Valuetine mini, Reput Free Hauptine and Settered at	alian
Indicas, Landon West (Ed.), UK.	Glikkya wate nerveral for actualizerynal activulat and been wate setuened for fragin-X if this had our been done

SCPENCE ND THE PUBLTC

MISCONDUCT AND CONFIDENCE – A MEDIA ANALYSIS

Does media coverage of scientific misconduct affect public confidence in science and scientists? And how is research reported in the Swedish media? These are the questions that Swedish non-profit organisation **Vetenskap & Allmänhet, VA, (Public & Science)** has been investigating, together with the **SOM Institute** at the University of Gothenburg, in the study Misconduct and **confidence – a study of media coverage of scientific misconduct and public confidence in research** (VA report 2014:3). The main findings are presented here. You can download the complete report (in Swedish only) or a summary in English at www.v-a.se

Vetenskap & Allmänhet



- 359 articles and TV clips published between 1 January 2002 to 31 December 2013.
- Scientific misconduct: A conscious intent to deceive
 - plagiarism
 - fabrication
 - manipulated (falsification).

Andersson, U. (2015). Does media coverage of research misconduct impact on public trust in science? A study of news reporting and confidence in research in Sweden 2002 -2013. *Observatorio* 9 (4) 015-030, Available at http://obs.obercom.pt.

FIGURE 1: CONFIDENCE IN RESEARCH AND SCIENTISTS (PERCENT) AND THE NUMBER OF ARTICLES/ITEMS ABOUT SCIENTIFIC MISCONDUCT, 2002–2013





Comment: The solid lines show the percent that have confidence in scientists as a profession, as well as confidence in research in medicine and the social sciences, respectively. The respondents stated that they have very or fairly high confidence in the national SOM survey 2002–2013. Confidence in scientists was not measured in 2009. The dotted lines show the number of articles about scientific misconduct per year and the number of articles about scientific misconduct relating to medicine and the social sciences, respectively.





Figure 2. Public trust in medical research and number of news reports on research misconduct in medicine



- There is only a small decrease of the general trust in science (Swedish study)
- Trust in science /scientists is not related to the media coverage of scientific misconduct. (Swedish study)
- The general trust in science /scientists is in most countries still high or there is only a slight decrease.
- The Nielsen Report from 2014 on Attitudes about Science in NZ does not ask for *trust* in science, but it reveals a high *appreciation of science*



Public confidence in institutional leaders, by selected institution: 1973-2014

NOTE: Responses to As far as the people running these institutions are concerned, would you say that you have a great deal of confidence, only some confidence, or hardly any confidence at all in them? Figure shows only responses for "a great deal of confidence."

SOURCE: University of Chicago, National Opinion Research Center, General Social Survey (1973–2014). See appendix table 7-23.

Science and Engineering Indicators 2016

SCPENCE

RND THE

FIGURE 37: COMPARING THE IMPORTANCE OF THREE AREAS RELATING TO THE IMPORTANCE OF SCIENCE FOR THE ECONOMY AND THE ENVIRONMENT – 2010 VS. 2014





Strongly agree Moderately agree Neither agree nor disagree Moderately disagree Strongly disagree Don't know

*Note that wording in 2010 survey was "New Zealand needs to develop science in order to enhance our international competitiveness"

Next are some statements some people have made about science and technology. Please indicate how much you agree or disagree with them...? Base CATI 2010 n= 600, 2014 n= 500; ONLINE 2010 n= 600, 2014 n= 2,504

> Nielsen. (2014). Report on Public attitudes towards science & technology. Retrieved from http://www.curiousminds.nz/assets/Uploads/report-on-public-attitudes-towards-scienceand-technology.pdf



- There is only a small decrease of the general trust in science (Swedish study)
- Trust in science /scientists is not related to the media coverage of scientific misconduct. (Swedish study)
- The general trust in science /scientists is in most countries still high or there is only a slight decrease.
- The Nielsen Report from 2014 on Attitudes about Science in NZ does not ask for trust in science, but it reveals a high and positive appreciation of Science.



Everything is fine?



- There is only a small decrease of the general trust in science (Swedish study)
- Trust in science /scientists is not related to the media coverage of scientific misconduct. (Swedish study)
- The general trust in science /scientists is in most countries still high or there is only a slight decrease.
- The Nielsen Report from 2014 on Attitudes about Science in NZ does not ask for trust in science, but it reveals a high and positive appreciation of Science.



Everything is fine?

Let's talk not only about the weather, but also about the climate.



Let's not talk only about the *weather (trust)*, but also about the epistemic *climate*.

The pattern is different and less positive,

- if the public is asked with regard to specific topics, as climate change, nuclear energy or genetically modified food.
- if scientific topics are related to political (and in some countries: religious) preferences / attitudes.
- if there are political and or religious public debates /controversies.

Then even alleged scientific misconduct could matter, as the *climate gate (hacked emails from climate researchers)* example has shown.



Epistemic climate

My preliminary definition: The generally accepted practices of establishing what could be known as ,true' about the natural, social and cultural world and the generaly accepted practices for the discourse about and with this knowledge.

This includes the distinction between those questions which could -at least in principle - be answered by research and those which are answered by personal values. (In this vain, it does not make sense to ,believe' in the human causes for climate change.)



Epistemic climate

My preliminary definition: The generally accepted practices of establishing what could be known as ,true' about the natural, social and cultural world and the generaly accepted practices for the discourse about and with this knowledge.

Is there a change of the epistemic climate?





Is there a change of the *epistemic climate*? This is an open question.

The answer will presumably be different for different regions of the world.

Survey data on public trust in science are measurements of single weather parameters.

These parameters (together with many others) make up the weather, not the climate.

Researching into the publics' actual understanding of the epistemic and discoursive practices of science could help to establish broader perspective on ,public trust' than surveys can provide. On the long run it could help to reveal if there is an *epistemic climate change*.

Examples from the work from our research group:

- The structure of laypersons' trustworthiness judgements about science experts
- Effects of overhearing a controversy among scientists

Study Example 1: The structure of trustworthiness judgements

Hendriks F, Kienhues D, Bromme R (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). PLoS ONE 10(10): e0139309. doi:10.1371/journal.pone.0139309

Friederike Hendriks, Dorothe Kienhues





Study Example 1: The structure of trustworthiness judgements

Hendriks F, Kienhues D, Bromme R (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). PLoS ONE 10(10): e0139309. doi:10.1371/journal.pone.0139309

To measure laypeople's evaluations of science experts (encountered online), we constructed an inventory to assess epistemic trustworthiness.

Exploratory (n = 237) and confirmatory factor analyses (n = 345) showed that judgments about the trustworthiness of science experts entail these three dimensions: **Expertise, integrity, and benevolence**.

METI items rated on a 7point scale like a semantic differential.

The Muenster Epistemic Trust Inventory (METI)

Factor	Item
	competent – incompetent
	intelligent – unintelligent
Exportico	well-educated – poorly educated
Expertise	professional – unprofessional
	experienced – inexperienced
	qualified – unqualified
	sincere – insincere
Intogrity	honest – dishonest
integrity	just – unjust
	fair – unfair
	moral – immoral
Denovelence	ethical – unethical
Benevolence	responsible – irresponsible
	considerate – inconsiderate

Study Example 1: The structure of trustworthiness judgements

Hendriks F, Kienhues D, Bromme R (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). PLoS ONE 10(10): e0139309. doi:10.1371/journal.pone.0139309

To measure laypeople's evaluations of science experts (encountered online), we constructed an inventory to assess epistemic trustworthiness.

Exploratory (n = 237) and confirmatory factor analyses (n = 345) showed that judgments about the trustworthiness of science experts entail these three dimensions: **Expertise, integrity, and benevolence**.

METI items rated on a 7point scale like a semantic differential.

The METI development is based on experimental studies (for testing the validity of the scales) and on well educated samples. But a very recent (2017) representative German survey has corroberated the results.



Here are some reasons why you might trust scientists. To what extent do you personally agree with them?

Because scientists are experts in their field.17523Because scientists work according to rules and standard procedures.11341637281134Because scientists do research in the public interest.1382

● stimme voll und ganz zu ● stimme eher zu ● unentschieden ● stimme eher nicht zu ● stimme nicht zu ● weiß nicht, keine Angabe



Here are some reasons why you might distrust scientists. To what extent do you personally agree with them? Because scientists are strongly dependent on the funders of their research.





Reasons for trust

Because scientists do research in the public interest.

14	26		37		13	8
Because scier	ntists work acco	rding to rules	and standard pr	ocedures.		
16		37		28		34
Because scie	ntists are expert	s in their field	i.			
31			41		17	5 2 3

Reasons for distrust

Because scientists are strongly dependent on the funders of their research.

		46		30	11	5	5	2		
Because scientists often adjust results to their own expectations.										
1	13	27	31		17	8		4		
Because scientists often make mistakes.										
6	12	43	3	26		11		3		

completely agree
somewhat agree
undecided
somewhat disagree
don't know, missing answer



Reasons for trust

Reasons for distrust

Because scientists do research in the public interest.					Because scientists are strongly dependent on the funders of their research.							
14	26		37		Benevo	lence	46		30	11	5	52
Because scient	ists work accordir	ng to rules and sta	indard procedures.			Because scientis	ts often adjust results	s to their own expe	ectations.			
16	3	17	28		_ uInte g	rity	27	31		17	8	4
Because scient	tists are experts in	n their field.				Because scientis	ts often make mistak	es.				
31 41			17	Exper	tise12	43		26		11	3	

completely agree
somewhat agree
undecided
somewhat disagree
don't know, missing answer

Study Example 2: Effects of overhearing a controversy among scientists



- **First assumption**: For epistemic trustworthiness ascriptions to an expert, it matters **who** discloses new evidence (Jensen, 2008).
 - A scientist blogger adds new evidence to his blog entry himself.
 - Another expert is responsible for this addition.
- Second assumption: For epistemic trustworthiness ascriptions to an expert, it matters what (i.e. which kind of evidence) is disclosed.
 - A blog's commentary might entail scientific (content or method related) critique

(First experiment)

- Or underlying societal or ethical aspects of an issue (Second experiment).

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Disclose your flaws! Admission positively affects the perceived trustworthiness of an expert science blogger. *Studies in Communication Sciences*, 16(2), 124-131. doi:10.1016/j.scoms.2016.10.003

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Evoking vigilance: Would you (dis)trust a scientist who discusses ethical implications of research in a science blog? *Public Understanding of Science*, 25(8), 992-1008. doi: 10.1177/0963662516646048 28

Materials



- A science blog (fictitious)
 - about a study investigating a neuroenhancing drug.
- A **comment** authored either by the responsible scientist blogger, or by another expert, entailing...
 - Scientific critisism (Experiment 1):
 - Optimism bias pertaining to the conclusiveness of results.
 - Ethical aspects (Experiment 2):
 - Arguments for and against neuroenhancement.



Medikamenter im Institut

für Neuropharmakologie

Educt

Neue erfolgversprechende Ergebnisse im Forschungsfeld "Neuro-Enhancement"!

Ich möchte Ihnen im heutigen Blog-Eintrag von unserer neuesten Studie mit einem Medikament zur kognitiven Leistungssteigerung berichten: Serofinil.

Das Forschungsfeld des Neuro-Enhancement - oder umgangssprachlich Gehim-Doping - beschäftigt sich seit einigen Jahren mit der Entwicklung eines Mittels, das die Konzentration verbessern und dadurch die Denkleistung steigern kann. Ich persönlich forsche seit einigen Jahren mit dem Medikament Serofinil, das ursprünglich im Rahmen der Migrane-Vorsorge verschrieben und eingesetzt wurde. Patienten, die es aufgrund von starker Mioräne verschrieben bekommen hatten, berichteten, konzentrationsfähiner und wacher zu sein (Müller & Schweizer, 2006)

In einer kurzen Studie haben wir uns nun noch einmal der Frage gewidmet, ob die Denkleistung durch das Medikament Serofinil bedeutend gesteigert werden könne. Dabei wurde Patienten für vier Wochen täglich jeweils eine Dosis von 0.23mg Serofinil verabreicht, eine weitere Gruppe bekam kein Medikament. In nachfolgenden Tests der Konzentration und Aufmerksamkeit (Jacobs, 2013; Brickenkamp, 1962) schnitt die Gruppe unter Serofinil deutlich besser ab als die Gruppe ohne neurologische Unterstützung Hieraus kann man schlussfolgern, dass sich Serofinil für die gezielte Einnahme bei geistig fordernden Tätigkeiten, zum Beispiel beim Lemen vor Prüfungen, eignet. Wir planen dieses Medikament im Laufe des Jahres 2014 durch die Zulassung beim Bundesinstitut für Arzneimittel und Medizinprodukte bestätigen zu lassen, sodass auch gesunde, d.h. nicht an Migrane leidende Menschen, von den positiven Effekten des Medikaments Serofinil profitieren können. Allerdings müssen noch weitere Studien folgen, bevor Nebenwirkungen und tatsächliche Wirksamkeit von Serofinil abschließend geklärt sind, damit Serofinil für gesunde Menschen eine Zulassung erhalten kann

+Lite

Be the first to like this

This entry was posted in Cognitive Enhancement, Neuropharmakologie, Serofinil by Dr. Gustav Vieratz. Bookmark the permatine

ONE THOUGHT ON INFUE ERFOLGVERSPRECHENCE ERGEBNISSE IN FORSCHUNGSFELD INFURO-ENHANCEM

OUSTAV VIERATZ

wissenschafture

Mitarbeiter und Teil

der Forschungsgrupp Leistungssteigemde

Medicamente' im Institut

für Neuronharmakologie

Erfurt

Dr. Gustav Vieratz on September 3, 2013 at 4:16 pm said Ich muss hinzufügen, dass ethische Argumente im meinem Blogeintrag bisher fehlen, aber dringend diskutiert werden müssen. Daher fasse ich nun die wichtigsten Argumente für und gegen die Nutzung von kognitiv leistungssteigernden Medikamenten zusammen.

Für die Nutzung kognitiv leistungsteigender Medikamente spricht, dass Menschen ständig dabei sind, ihre Denkleistung durch äußere Einwirkung zu erhöhen: Schlaf Zufühnung von Nährwerten durch Ernährung sowie Sport sind gatürliche Mittel un die Leistung unserer Gehime zu verbessem. Auch trinken die meisten von uns täglich Kaffee oder Tee - diese Getränke enthalten Stoffe, die der Leistung des Gehirns positiv zu Gute kommen. Zweitens gibt es auch ein gesellschaftliches Argument: Die Menschheit strebt seit jeher nach ihrer Verbesserung und

29

Procedure of the two Experiments





Epistemic trustworthiness on the dimensions expertise, integrity and benevolence Decision to recommend the drug to a (fictitious) friend

Experiment 1: Scientific Critique







Experiment 1: Scientific Critique







Experiment 1: Scientific Critique





Experiment 1: Summary of Results



• Who?

- Integrity and Benevolence are rated higher if the responsible scientist blogger admits critical aspects himself (vs. Other's Comment).
- What?
 - Expertise ratings are lower, if another expert's comment is criticizing the conclusiveness of results (vs. No Comment).

Experiment 2: Ethical Aspects





Experiment 2: Ethical Aspects





Experiment 2: Ethical Aspects









- The experiments provide evidence on effects of open criticism and debate (in the comment section of a blog) on laypeople's inferences of expert's epistemic trustworthiness.
- Pragmatic expectations are only one example for conversation related heuristics used by citizens when monitoring or even when only overhearing conversations among scientists.
- Inferring about scientists trustworthiness from *conversation* related cues does not require *content* related scientific expertise.
 - It is an example for a heuristic which is not /less constrained by citizens' bounded understanding of science. It is not about electrons, bacteria or genetic expressions. Instaed it is about the quality of sources and discourse.



Is there a change of the *epistemic climate*? This is an open question.

The answer will presumably be different for different regions of the world.

Survey data on public trust in science are measurements of single weather parameters.

These parameters (together with many others) make up the weather, not the climate.

Researching into the **publics' actual understanding of the epistemic and discoursive practices of science** could help to establish broader perspective on ,public trust' than surveys can provide. On the long run it could help to reveal if there is an *epistemic climate change*.



tēnā koutou Many thanks

bromme@uni-muenster.de

Most of our publications could easily be retrieved from <u>www.uni-muenster.de/PsyIFP/AEBromme/</u> veroeffentlichung/veroeffentlichung.html



