



Daten frisieren, Ideen klauen – falsche Forscher in der Wissenschaft

Alessandra Rachetti & Wolfhard Wegscheider

Wissenschaftliches Fehlverhalten Themen

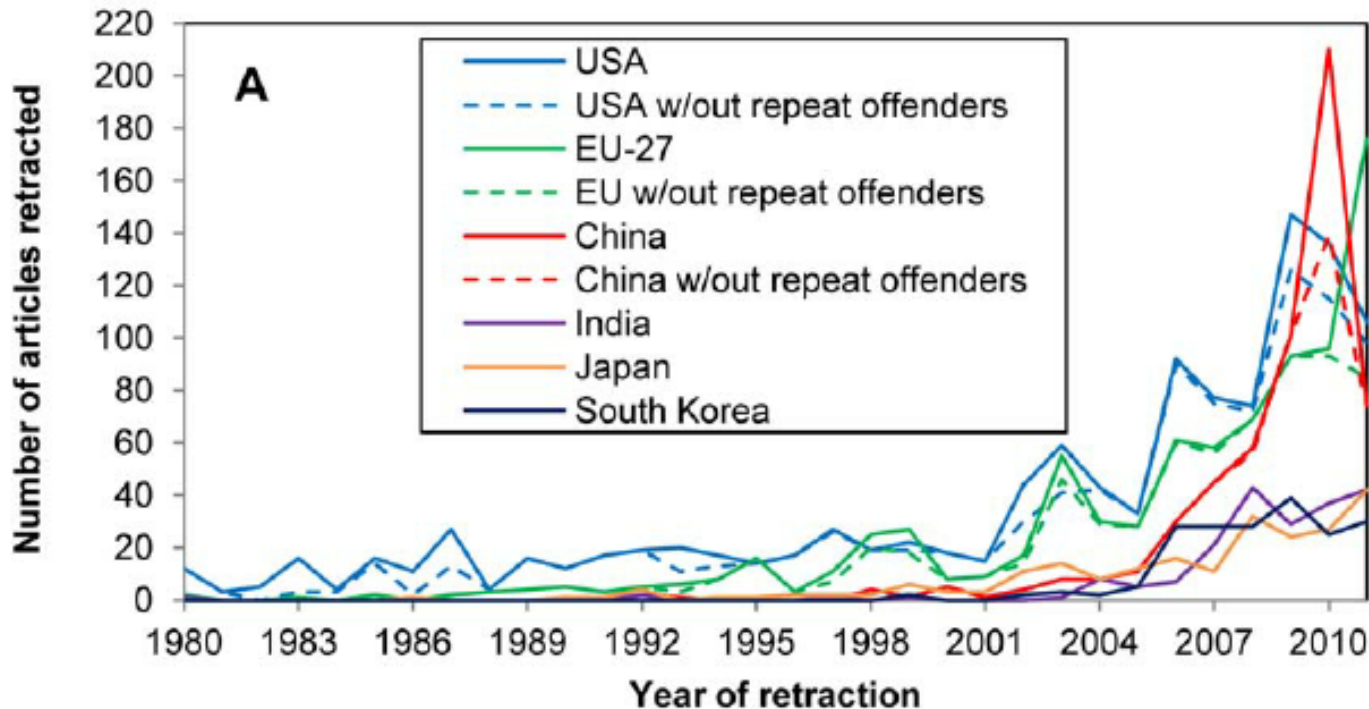
- Häufigkeit
- Definition / Rechtliches
- Vergleich International / Österreich
- Möglichkeiten der Detektion von verfälschten Daten
- Motive, Abschreckung, Auswege

Wissenschaftliches Fehlverhalten Häufigkeit

	Eigenauskunft anonym	Beobachtungen im Kollegenkreis
Datenverfälschung, Plagiiere	2%	14%
Weiteres wissenschaftliches Fehlverhalten	33%	72%

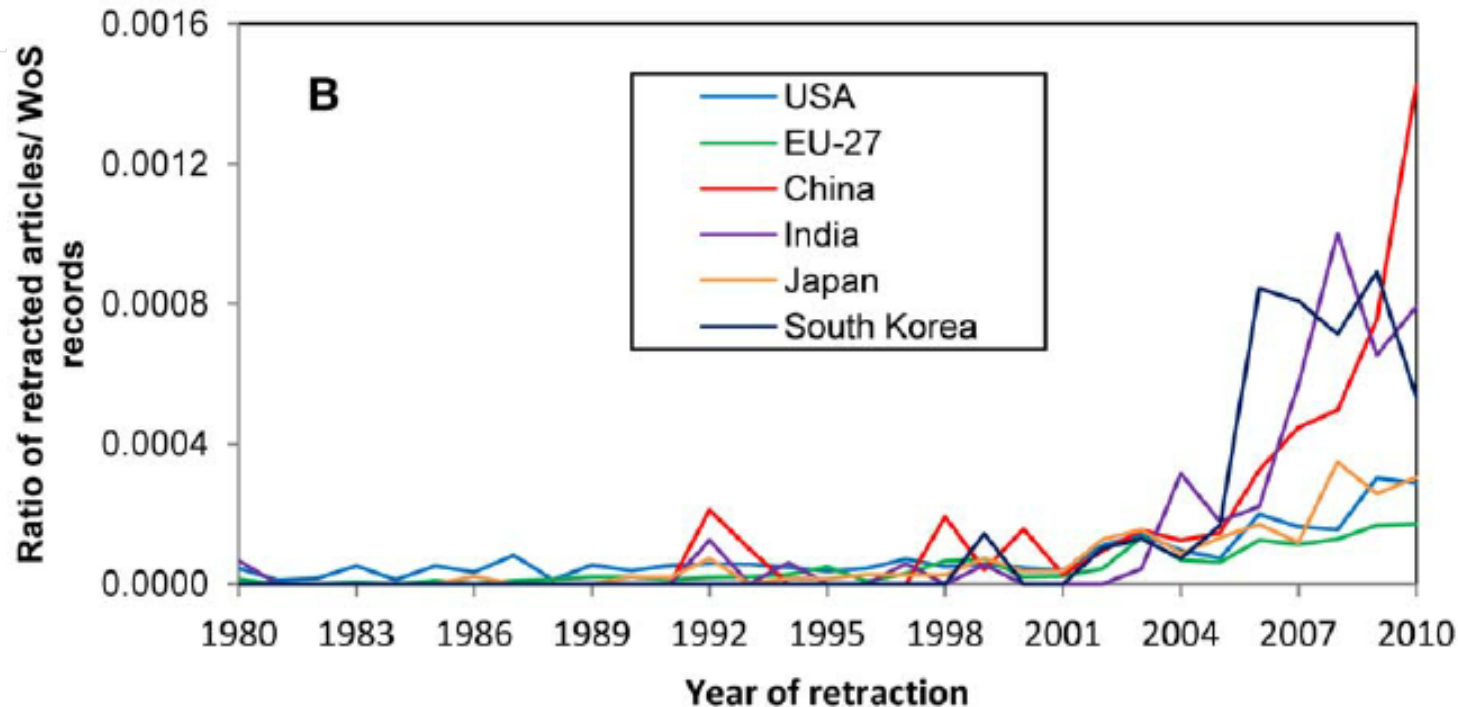
Meta Analyse (2009) Fanelli

Anzahl der zurückgezogenen Publikationen



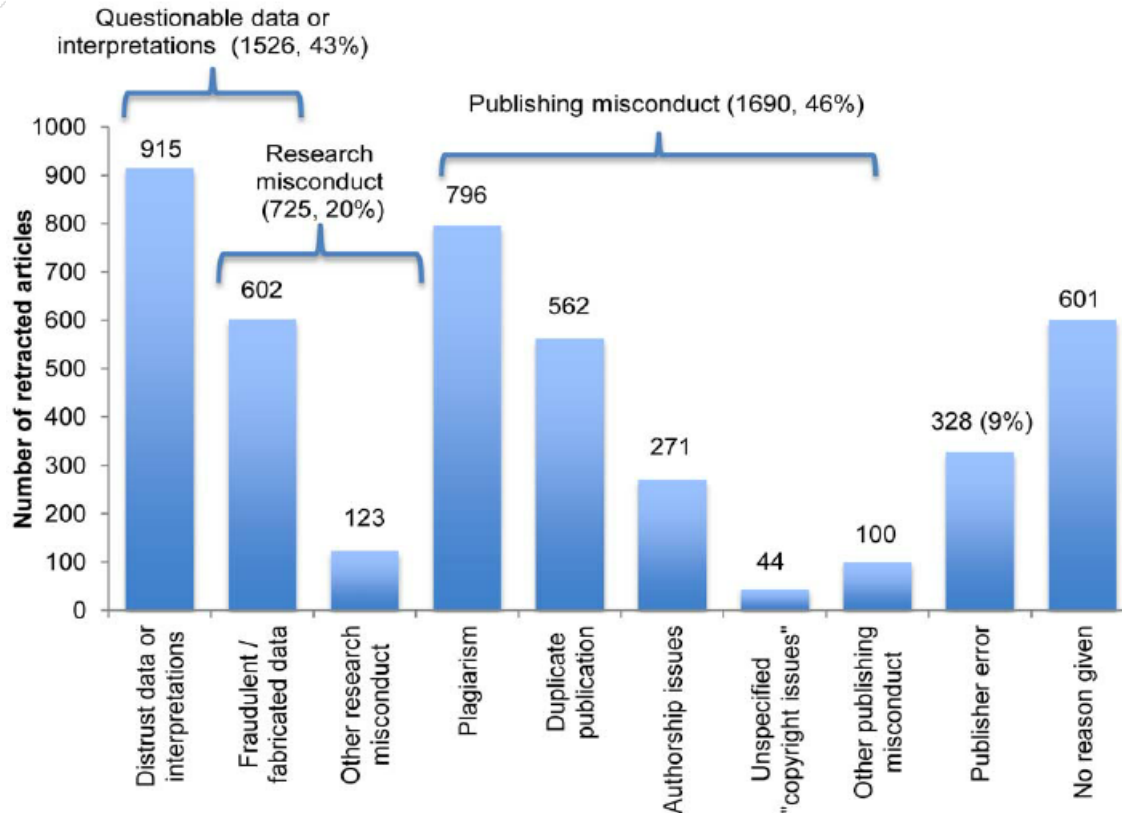
Yearly distribution of retracted articles based on author affiliation countries. (A) Number of articles retracted per year for EU-27, and top 5 non-EU countries. Note that China's large spike in 2010 was mainly due to H. Zhong and T. Liu; and the EU-27 spike in 2011 is due to J. Boldt. Thus far in 2011, values for China (either red line), and EU-27 minus J. Boldt (green, dashed line) remain well below the USA values (either blue line).

Fraktion der zurückgezogenen Publikationen



Yearly distribution of retracted articles based on author affiliation countries. (B) Ratio of number of retracted articles/number of Web of Science articles for each year and geographic region. Because many retracted articles are not in Web of Science, these values are not true proportions, but allow for country comparisons. doi:10.1371/journal.pone.0044118.g005

Gründe für das Zurückziehen von Publikationen



Justifications for retraction stated in the notices consulted, which accounted for 4,232 retracted articles. Only 20% of articles were retracted due to research misconduct, while more than twice that many were retracted due to publishing misconduct. Note that 42% were retracted because of “questionable data or interpretations.” Percentages are based on the 3,631 (= 4,232 - 601) notices which stated the justification. doi:10.1371/journal.pone.0044118.g003



Richtlinien der Österreichischen Agentur für wissenschaftliche Integrität zur Guten Wissenschaftlichen Praxis (GWP-Richtlinien der OeAWI)

OeAWI - Struktur

Verein mit Vorstand, Mitgliedern, Kommission

Vorstandsvorsitzende

Präsident des Fonds zur Förderung der wissenschaftlichen Forschung (FWF).

Die Kommission für wissenschaftliche Integrität

unabhängiges, mit hochkarätigen WissenschaftlerInnen größtenteils aus dem Ausland besetztes Gremium.

bietet eine neutrale und sachorientierte Plattform, um Fällen (vermeintlichen) wissenschaftlichen Fehlverhaltens objektiv auf den Grund zu gehen.

untersucht auf professionelle Weise die Schwere eines Verstoßes und unterbreitet allenfalls Vorschläge für nachfolgende Maßnahmen.

Ist weder eine Entscheidungsinstanz noch eine rechtssprechende Organisationseinheit.

Verhaltensweisen die als wissenschaftliches Fehlverhalten anzusehen sind

Falschangaben

Verletzung geistigen Eigentums

**Beeinträchtigung der
Forschungstätigkeit anderer**

Wissenschaftliches Fehlverhalten (Österreich)

1. Die Erfindung von Daten („fabrication“)
2. Die Fälschung von Daten („falsification“)
3. Das Plagiiere („plagiarism“)
4. Die unberechtigte Verweigerung des Zugangs zu Primär- und Originaldaten
5. Die Behinderung der Forschungstätigkeit anderer Wissenschaftler/innen
6. Die Sabotage von Forschungstätigkeit
7. Unrichtige Angaben in einem Förderantrag
8. Die Benachteiligung beim beruflichen Fortkommen insbesondere von Nachwuchswissenschaftlerinnen und Hinweisgebern



Country	National policy (Y/N)	Year policy adopted	2014 R&D funding rank	2014 GDP (billions of US\$)	2014 R&D as %of GDP
United States	Yes	2000	1	16,616	2.8
China	Yes	2006	2	14,559	2
Japan	Yes	2006	3	4,856	3.4
Germany	Yes	2011	4	3,312	2.9
South Korea	Yes	2012	5	1,748	3.6
France	No*		6	2,319	2.3
United Kingdom	Yes	2012	7	2,454	1.8
India	No		8	5,194	.9
Russia	No		9	2,671	1.5
Brazil	Yes	2012	10	2,515	1.3
Canada	Yes	2011	11	1,571	1.9
Australia	Yes	2007	12	1,040	2.3
Taiwan	Yes	2000	13	974	2.4
Italy	No**		14	1,842	1.2
Spain	Yes	2011	15	1,418	1.3
Netherlands	Yes	2014	16	712	2.1
Sweden	Yes	2006	17	412	3.4
Israel	No		18	271	4.2
Switzerland	Yes	2008	19	382	2.9
Turkey	No		20	1,227	.9
Austria	No*		21	372	2.8
Singapore	Yes	2013	22	355	2.7
Belgium	No**		23	432	2
Iran	Yes	2011	24	1,014	.8
Mexico	No*		25	1,864	.5
Finland	Yes	2012	26	202	3.5
Poland	Yes	2012	27	844	.8
Denmark	Yes	2008	28	217	2.9
South Africa	No		29	621	1
Qatar	No*		30	211	2.7
Czech Republic	No**		31	295	1.8
Argentina	No		32	803	.6
Norway	Yes	2007	33	293	1.7
Malaysia	No		34	557	.8
Pakistan	No		35	556	.7
Portugal	No		36	248	1.4
Ireland	Yes	2013	37	200	1.7
Saudi Arabia	No		38	997	.3
Ukraine	No**		39	348	.9
Indonesia	Yes	2013	40	1,374	.2

*In the process of developing a national policy.

**National research ethics code but no national misconduct policy

Behaviors Defined as Misconduct in National Policies

Behavior	Number	*Percentage
Fabrication	22	100.0
Falsification	22	100.0
Plagiarism	22	100.0
Unethical Authorship (not Plagiarism)	12	54.6
Unethical Publication Practices	8	36.4
Conflict of Interest Mismanagement	8	36.4
Unethical Peer Review	7	31.8
Misconduct Related to Misconduct Investigations	6	27.3
Poor Record Keeping	6	27.3
Other Deception	6	27.3
Serious Deviations	5	22.7
Violating Confidentiality	5	22.7
Human or Animal Rights Research Violations	5	22.7
Misappropriating Funds	3	13.6
Misrepresenting Credentials	3	13.6
Theft of Physical Property	2	9.1
Interfering with Research	2	9.1
Other	9	40.9

* $n = 22$ countries with misconduct policies.

Research Misconduct - Official Definition

US office of Science and Technology Policy

“fabrication, falsification or plagiarism (FFP) in proposing, performing or reviewing research or in reporting research results”

Research Misconduct - Alternative View

"biased reporting"

"... any omission or misrepresentation of the information necessary and sufficient to evaluate the validity and significance of research, at the level appropriate to the context in which the research is communicated."

"Focus less on what scientists do and more on what they communicate."

Erfinden von Daten (Fabrication)

- die Erfindung von Forschungsergebnissen (Messwerten, Beobachtungsdaten, Statistiken)
- meistens aus finanziellen Gründen
- um mühselige Datenerhebungen zu umgehen oder unvollständige Datensätze zu vervollständigen
- besonders bei klinischen Studien und in der Marktforschung



Verfälschung von Daten (Falsification)

- Manipulation, Abänderung, selektives Weglassen von Daten, die der Forschungsthese widersprechen
- irreführende Interpretation von Daten
- auch “bending” genannt
- mit dem Ziel, ein gewünschtes Ergebnis zu erhalten oder eine bestimmte Hypothese zu bestätigen
- nicht einfach zu erkennen



Merkmale von gefälschten Daten

Verfälschte Daten:

- Werte, die nicht zu den gewünschten Ergebnissen passen, werden gelöscht oder geändert
- Daten gruppieren sich eher um den Mittelwert und die Verteilung ist schmaler

Erfundene Daten:

- sehr geringe Abweichungen
- völliges Fehlen von Ausreißern
- Ziffernpräferenz
- Verteilungen sind tendenziell flach
- Daten gleichmäßig über einen begrenzten Bereich verteilt

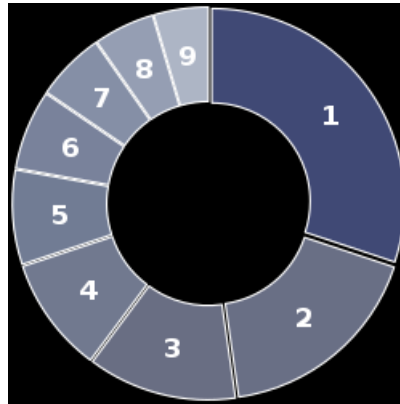
Gefälschte Daten - Vorkommen

- Biomedizinische Wissenschaften
- Marktforschung
- fast nie in Bereichen wie Physik, Astronomie und Geologie

Numerische und statistische Methoden zur Aufdeckung von falschen Daten

- Ziffern- Verteilung und Präferenzen
- Betrachtung von Bereich von Messwerten, Mittelwerten und Standardabweichungen, Perzentile, Schiefe, Wölbung
- Rekonstruktion von plausiblen Datensätzen aufgrund der Deskriptoren der Verteilung
- Untersuchung der zu erwartenden Korrelationen im multivariaten Datenraum

Ziffernverteilung Benfordsches Gesetz



Zahlen mit der Anfangsziffer 1 treten etwa 6,6-mal so häufig auf wie Zahlen mit der Anfangsziffer 9.

Das Benfordsche Gesetz gilt insbesondere für Zahlen von natürlichen Wachstumsprozessen

Benford's Law

Detecting Fraud Using Benford's Law

231

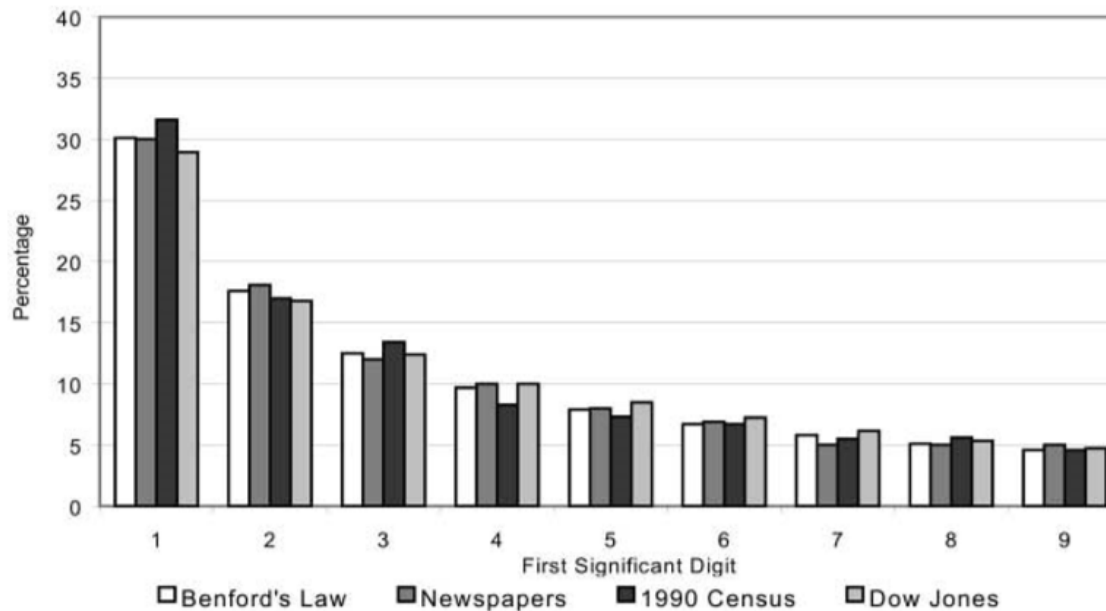


Figure 1. Benford's law predicts a decreasing frequency of first significant digits (Hill, 1998).

Benford's Law

Detecting Fraud Using Benford's Law

233

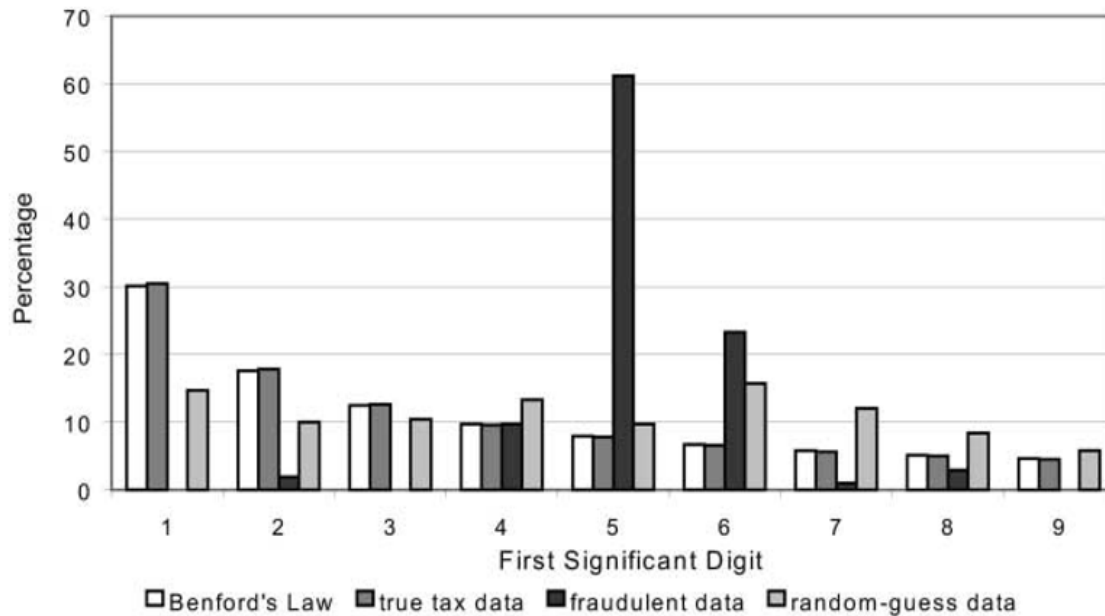


Figure 2. Percentages of the first digits of Benford's distribution, of true tax data, of fraudulent data and of Hill's random-guess data (Hill, 1998).

Ziffernverteilung

Gleichverteilung der letzten Stelle

- Ziffern sind umso mehr gleichverteilt, je kleiner ihr Stellenwert ist.
- Die am wenigstens signifikante Stelle einer Zahl enthält meistens Rauschen

Gleichverteilung der letzten Stelle Stamm Blatt Diagramm

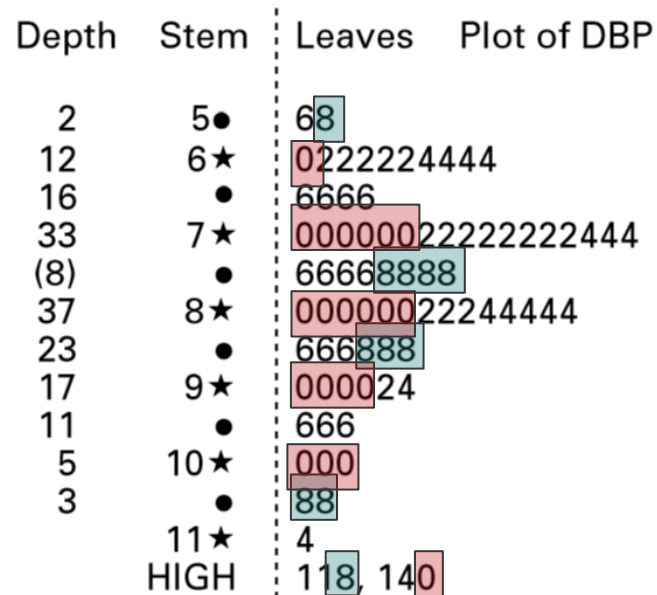



Figure 14.2 Stem and leaf plot of diastolic blood pressure. The first two values are 56, 58. Note that there are 21 values ending in zero, while 11 values end in 8. This is only slight digit preference. (The values are not expected to be measured to better than 2 mmHg.)



**“If you copy from one author, it’s
plagiarism, but if you copy from many,
it’s research.”**

Wilson Mizner

Opossums – echte und erfundene Daten



104 australische Gebirgsopossums,
9 Variable insgesamt

2 Variablen wurden ausgewählt um Daten mit einem vergleichbaren
Mittelwert / Standardabweichung zu simulieren

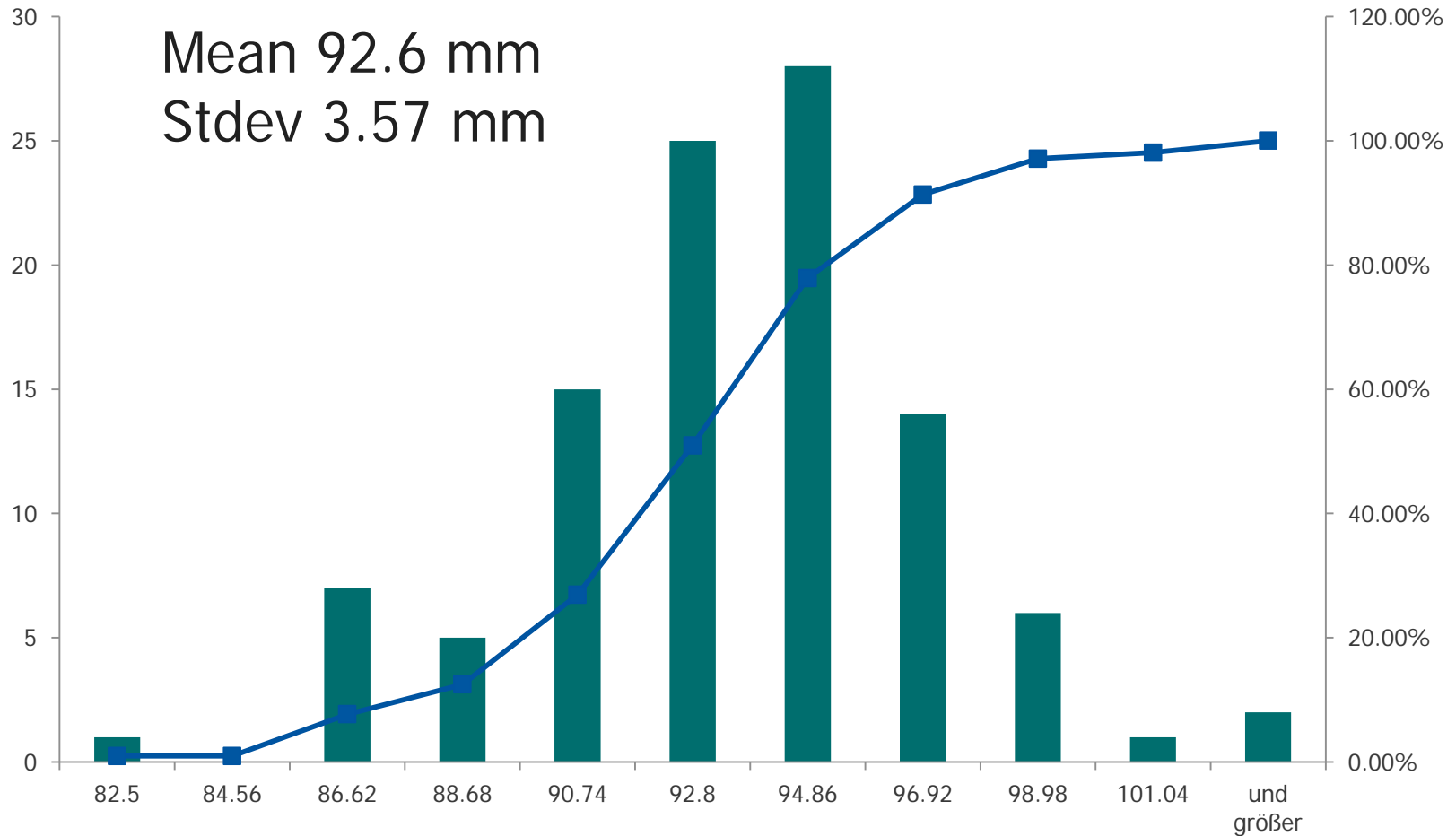
- Skullw skull width
- Hdlngth head length

Source

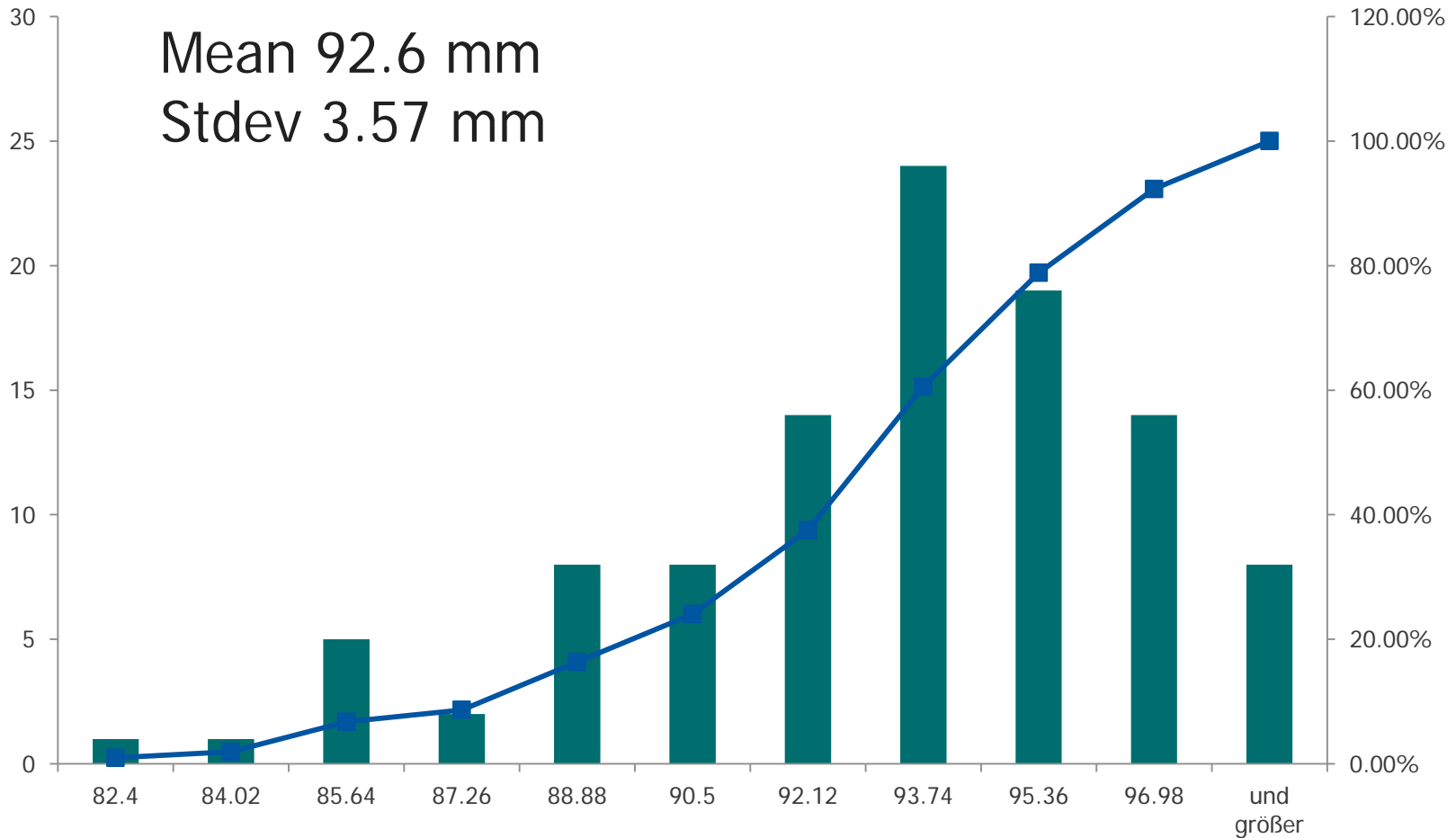
Lindenmayer, D. B., Viggers, K. L., Cunningham, R. B., and Donnelly, C. F. 1995. Morphological variation among columns of the mountain brushtail possum, *Trichosurus caninus* Ogilby (Phalangeridae: Marsupiala). *Australian Journal of Zoology* 43: 449-458.

<https://vincentarelbundock.github.io/Rdatasets/datasets.html>

Opossum Kopflänge/Originaldaten Histogramm

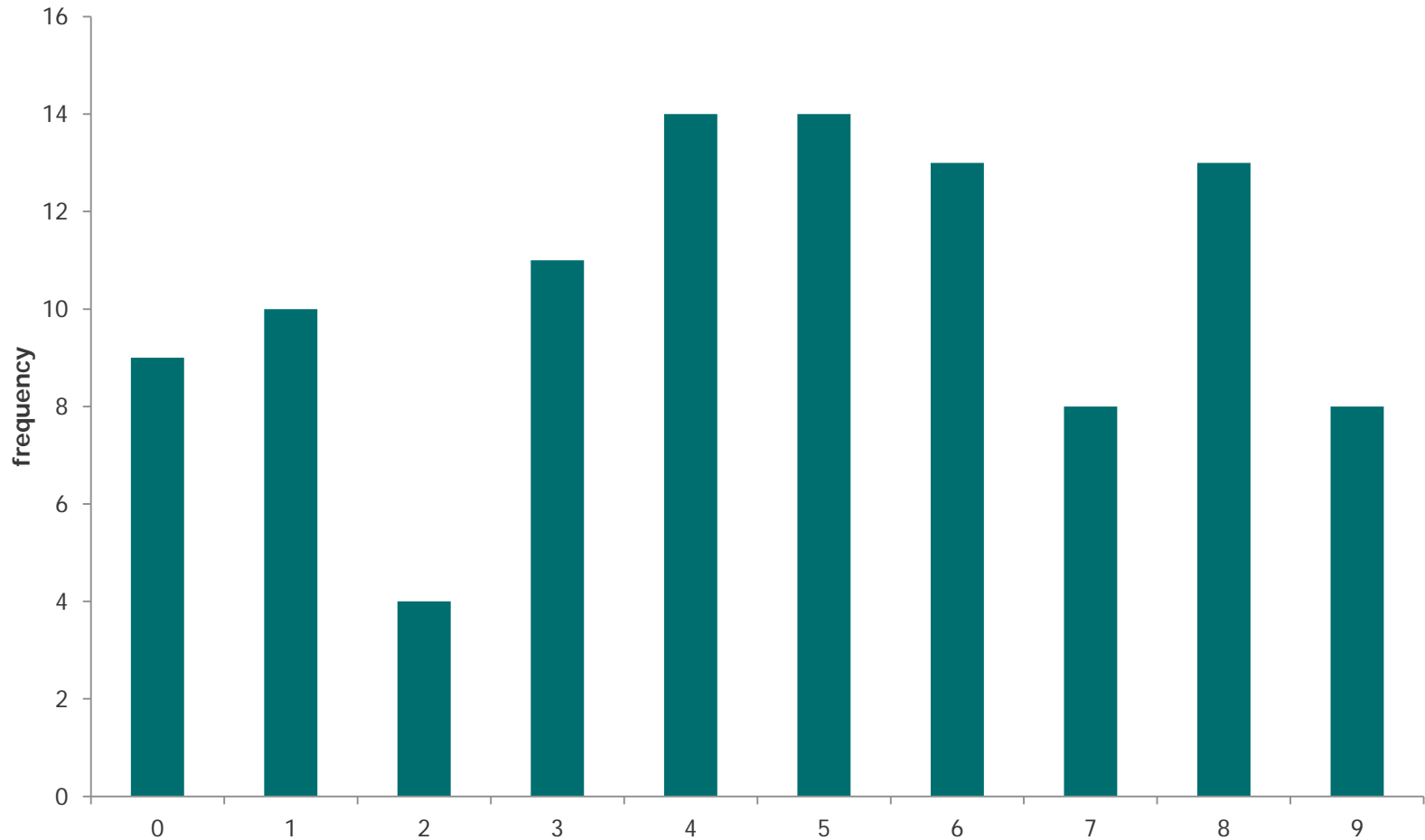


Opossum Kopflänge/gefälschte Daten Histogramm



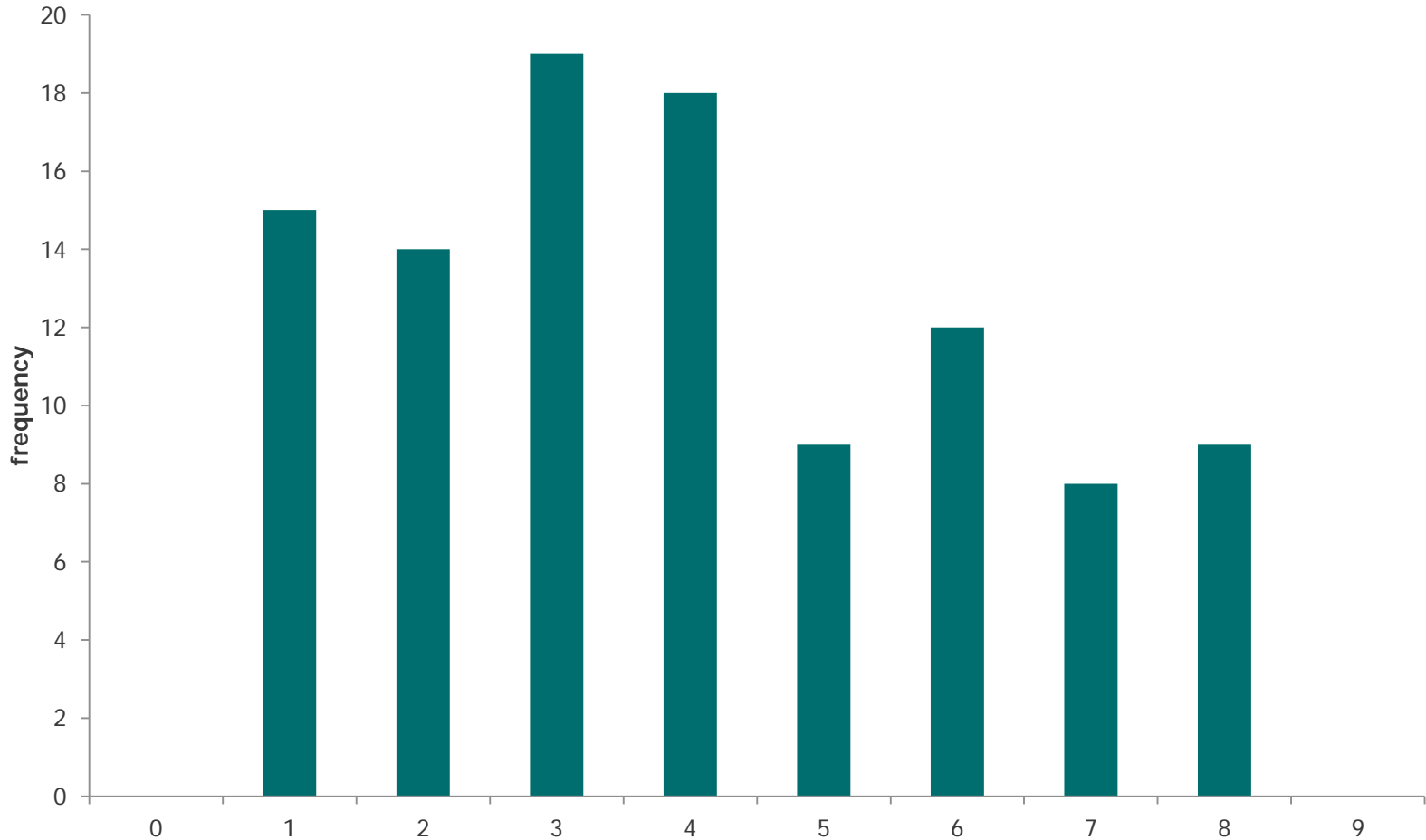
Opossum Kopflänge/Originaldaten

Verteilung der letzten Ziffern

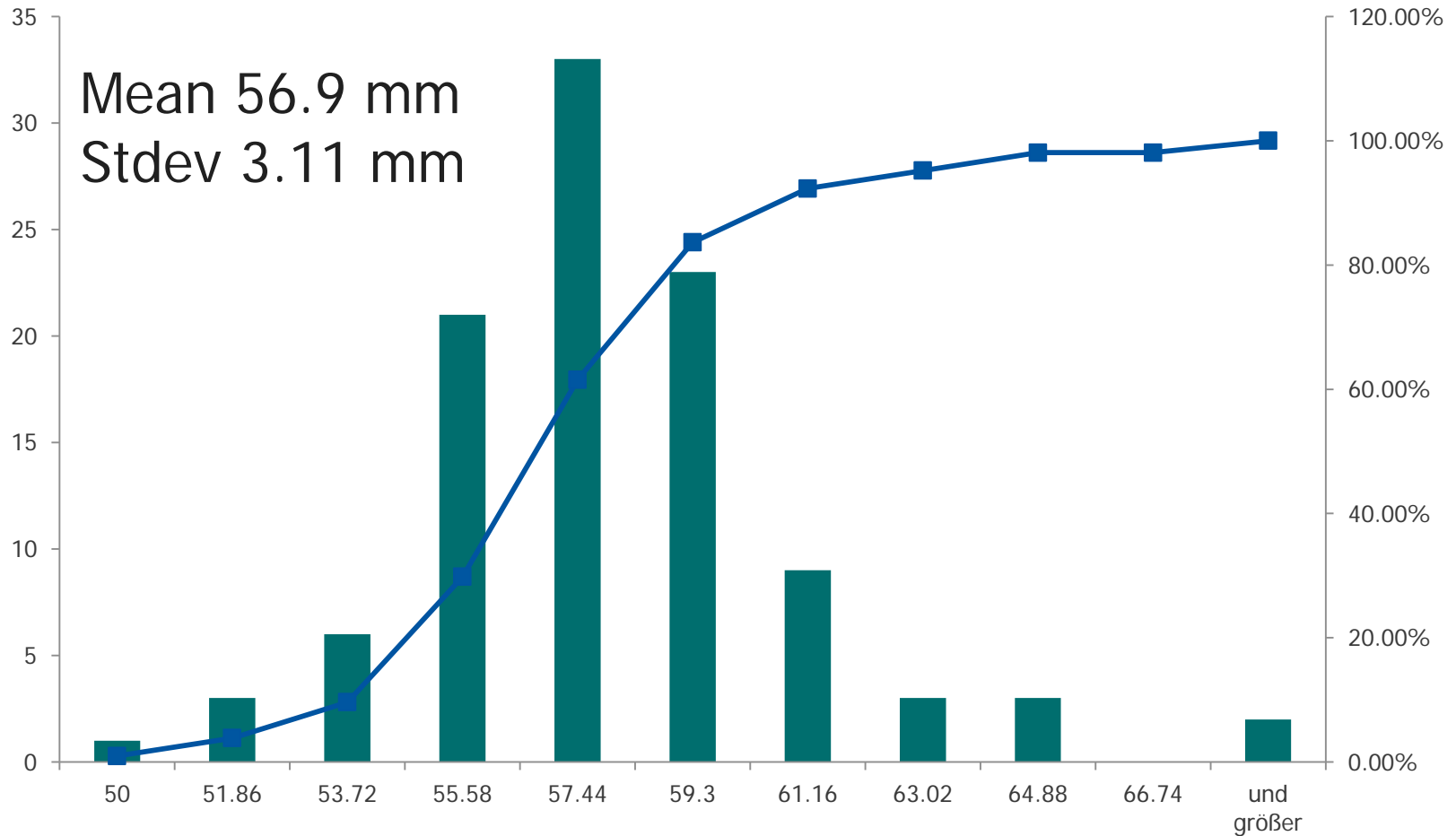


Opossum Kopflänge/gefälschte Daten

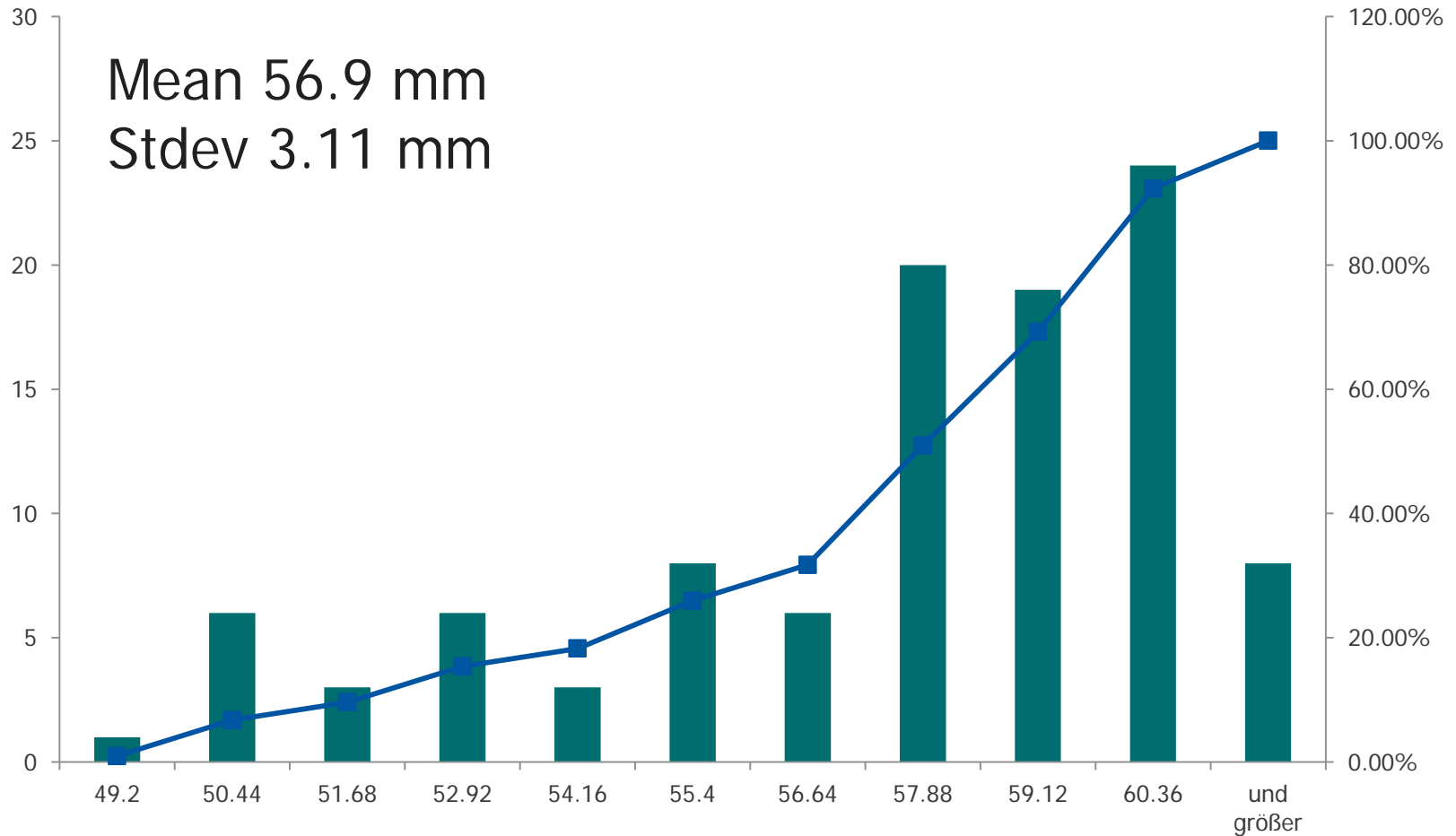
Verteilung der letzten Ziffern



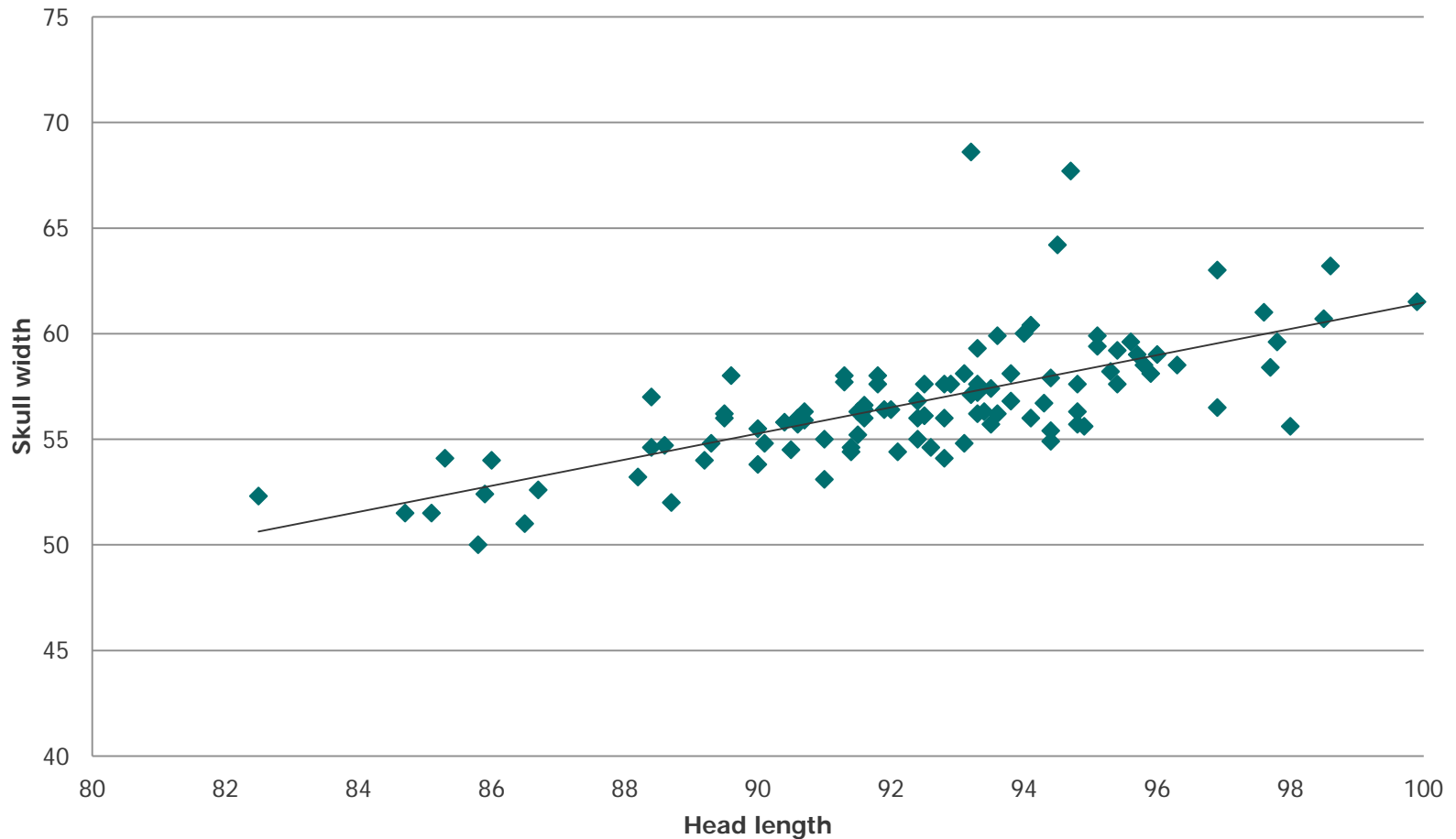
Opossum Schädelbreite/Originaldaten Histogramm



Opossum Schädelbreite/gefälschte Daten Histogramm



Opossum Schädelbreite /Kopflänge Originaldaten



GRIM, GRIMMER, SPRITE

GRIM: Granularity-Related Inconsistency of Means

GRIMMER: Granularity-Related Inconsistency of Means Mapped to Error Repeats

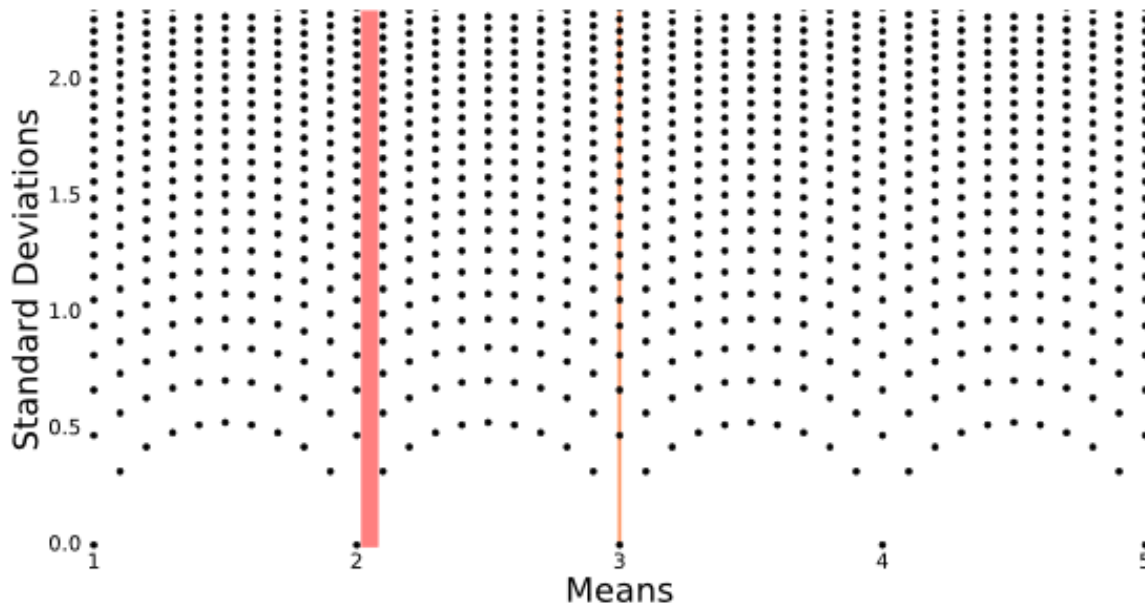
SPRITE: Sample Parameter Reconstruction via Iterative Techniques



Nick Brown, James Heathers
The „Data Thugs“

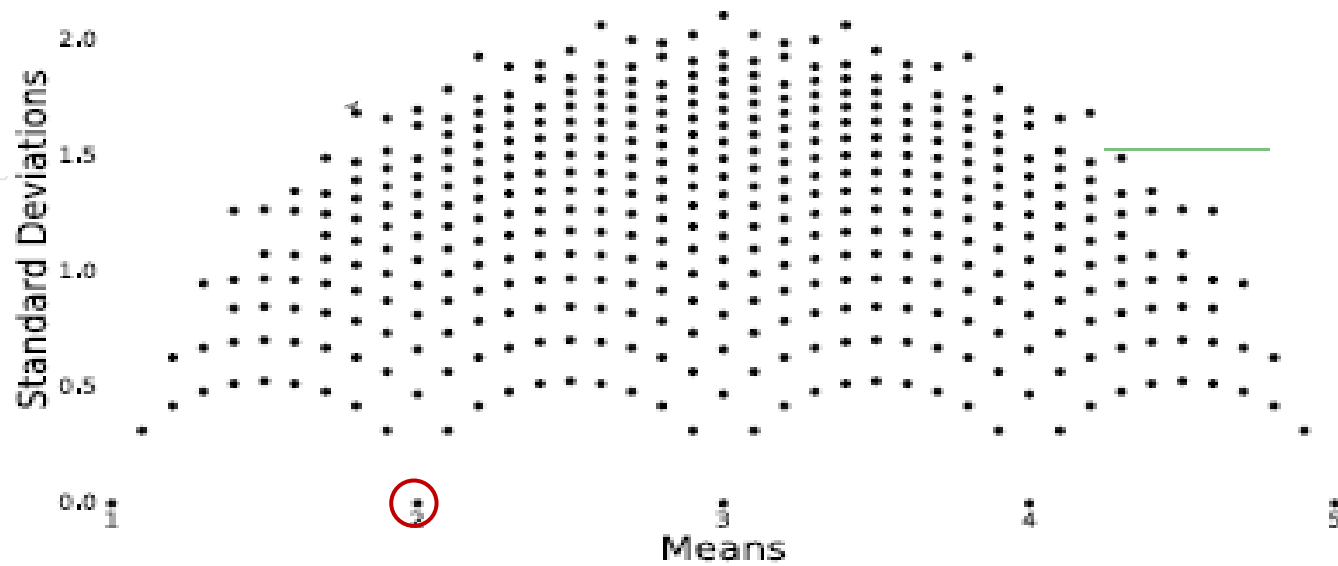
GRIM

granularity-related inconsistency of means

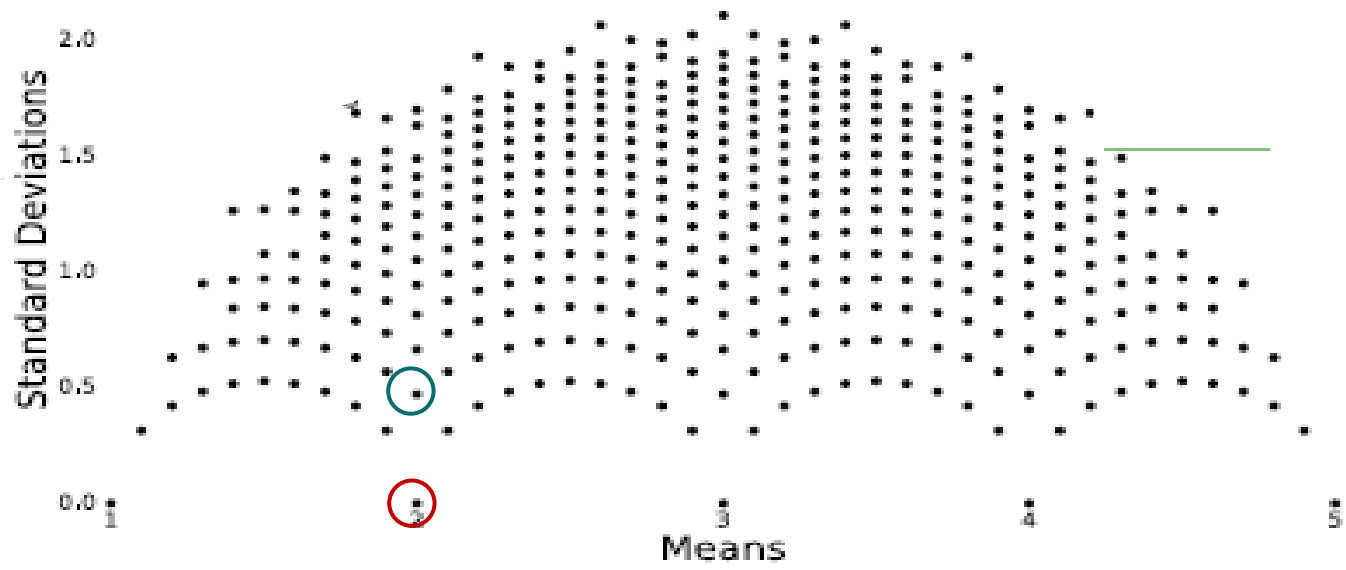


SPRITE

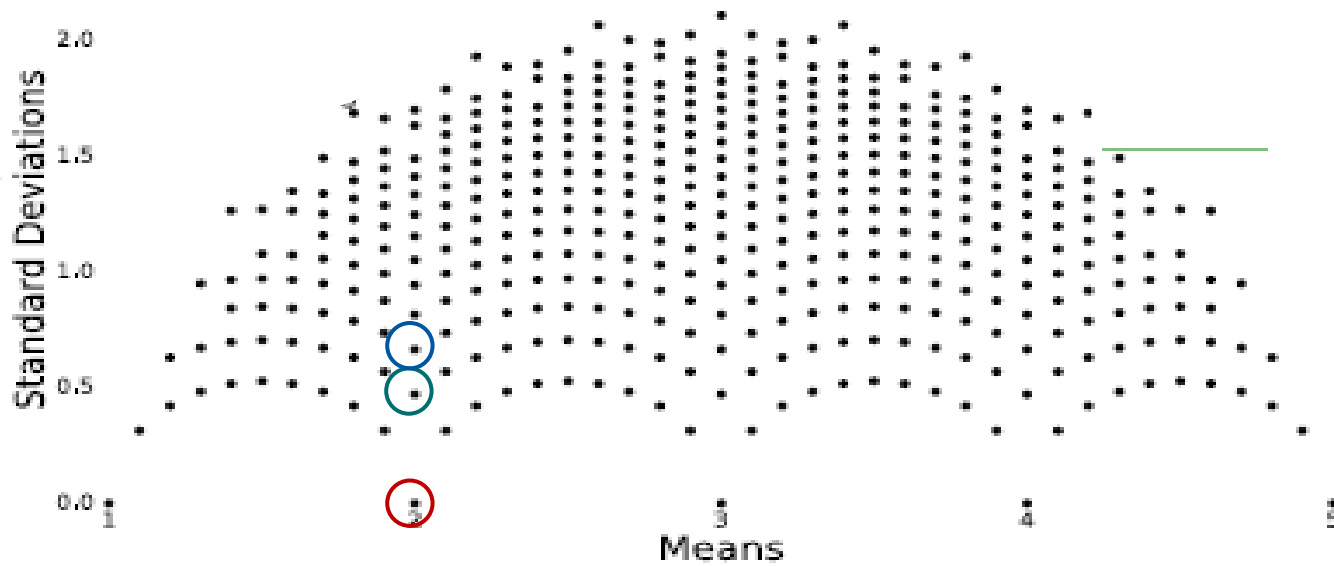
- Sample Parameter Reconstruction via Iterative Techniques
- SPRITE ermöglicht das Ableiten von statistisch möglichen Datensätzen aus den in einer Studie gemeldeten Mittelwerten und Standardabweichungen.
- Reverse Engineering



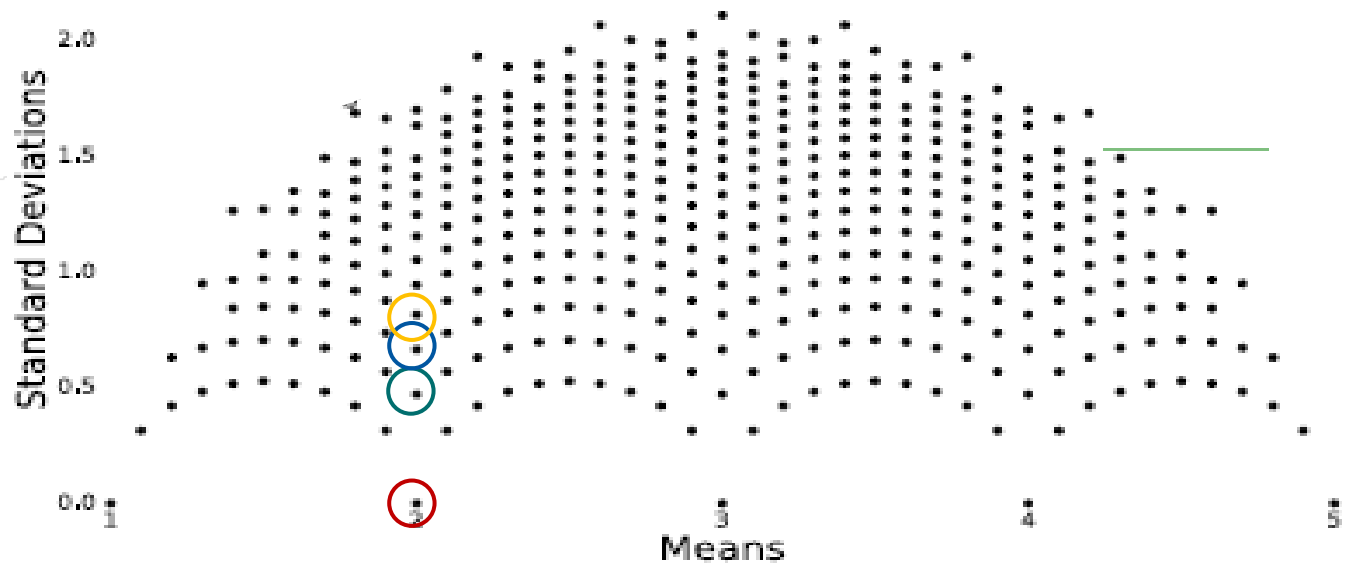
1	2
2	2
3	2
4	2
5	2
6	2
7	2
8	2
9	2
10	2
Mean	2
Sdev	0.00



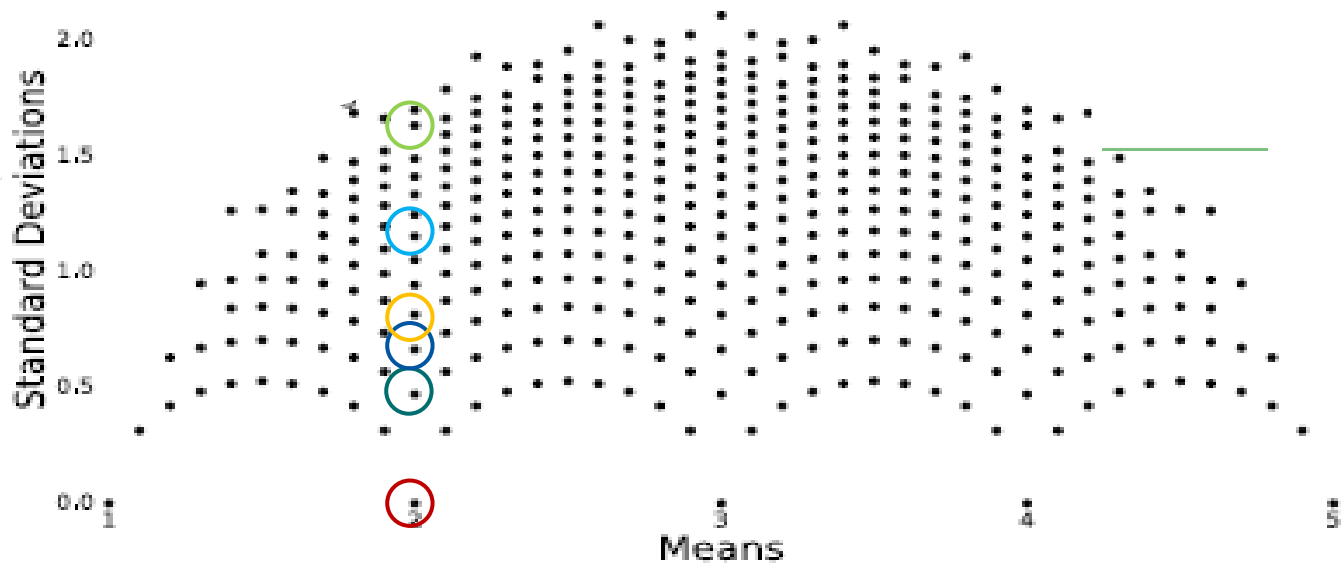
1	2	1
2	2	2
3	2	2
4	2	2
5	2	2
6	2	2
7	2	2
8	2	2
9	2	2
10	2	3
Mean	2	2
Sdev	0.00	0.47



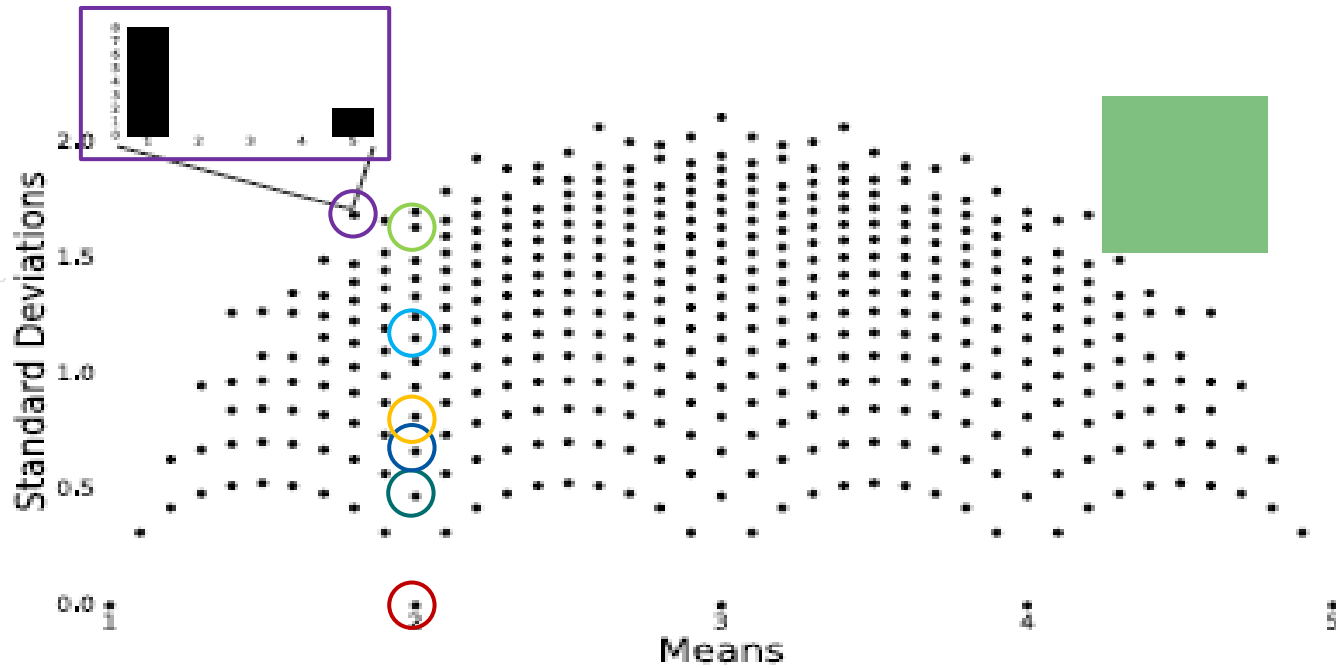
1	2	1	1
2	2	2	1
3	2	2	2
4	2	2	2
5	2	2	2
6	2	2	2
7	2	2	2
8	2	2	2
9	2	2	3
10	2	3	3
Mean	2	2	2
Sdev	0.00	0.47	0.67



1	2	1	1	1	1
2	2	2	1	1	1
3	2	2	2	1	2
4	2	2	2	2	2
5	2	2	2	2	2
6	2	2	2	2	2
7	2	2	2	2	2
8	2	2	2	3	2
9	2	2	3	3	2
10	2	3	3	3	4
Mean	2	2	2	2	2
Sdev	0.00	0.47	0.67	0.82	0.82



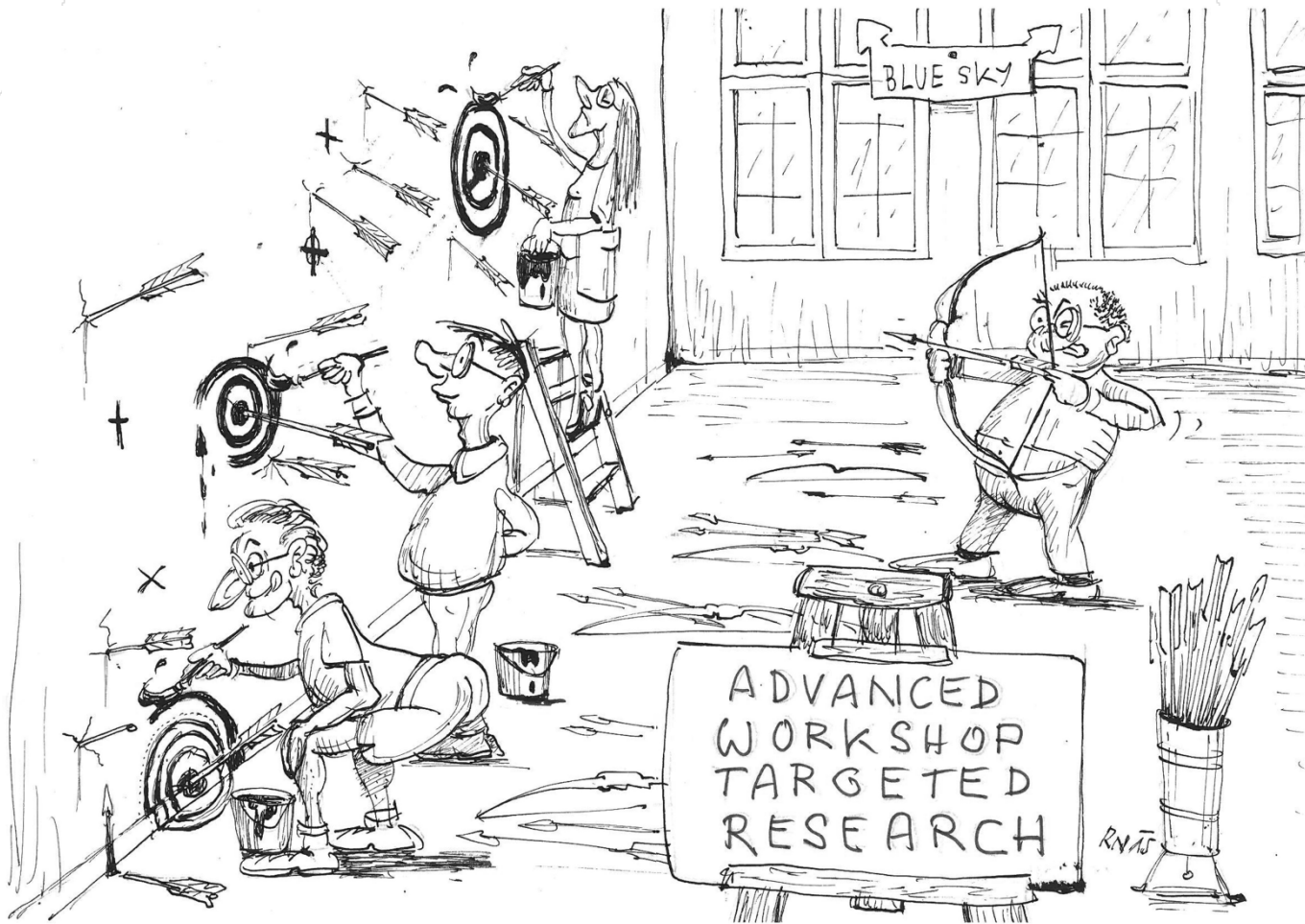
1	2	1	1	1	1	1	1
2	2	2	1	1	1	1	1
3	2	2	2	1	2	1	1
4	2	2	2	2	2	2	1
5	2	2	2	2	2	2	1
6	2	2	2	2	2	2	1
7	2	2	2	2	2	2	2
8	2	2	2	3	2	2	2
9	2	2	3	3	2	2	5
10	2	3	3	3	4	5	5
Mean	2	2	2	2	2	2	2
Sdev	0.00	0.47	0.67	0.82	0.82	1.15	1.63



1	2	1	1	1	1	1	1	1
2	2	2	1	1	1	1	1	1
3	2	2	2	1	2	1	1	1
4	2	2	2	2	2	2	1	1
5	2	2	2	2	2	2	1	1
6	2	2	2	2	2	2	1	1
7	2	2	2	2	2	2	2	1
8	2	2	2	3	2	2	2	1
9	2	2	3	3	2	2	5	5
10	2	3	3	3	4	5	5	5
Mean	2	2	2	2	2	2	2	1.8
Sdev	0.00	0.47	0.67	0.82	0.82	1.15	1.63	1.69

Weitere fragwürdige Praktiken in der Wissenschaft

- post-hoc-Analysen
- selektives Veröffentlichen der "interessantesten" Ergebnisse
- Nichtveröffentlichung negativer Befunde



Fragwürdige Daten

- unbeabsichtigte Fehler
- vermeidbare Irrtümer
- absichtliches Verfälschen
- arglistige und umfassende Datenerfindung

Eine perfekte Analyse von falschen Daten kann viel schädlicher sein als eine unvollständige Analyse korrekter Daten.

Aufdeckung von problematischen Daten

- Peer-Review
- voller Zugriff auf Originaldaten
- öffentliche zugängliche Datenbanken
- Statistische Tests
- Replikation der Messungen
- automatisiertes Scannen von Publikationen auf Inkonsistenzen

Das System des Peer Review

- Die Kontrolle der Qualität in wissenschaftlichen Publikationen wird traditionell durch das Peer-Review-System sichergestellt
- Behauptungen von Wissenschaftlern müssen professioneller Kritik standhalten und nachvollziehbar sein

Datenaufbewahrung, Zugang zu Datenbanken, Haltbarkeit der Daten

- Daten müssen gut dokumentiert und geschützt sein
- Verantwortung für die Dokumentation und die Nachvollziehbarkeit liegt beim Ersteller
- Kurator: verantwortlich für die Haltbarkeit, Zugang und Manipulationssicherheit

Aufdeckung durch statistische Analyse der Daten

- Bei der Ausbildung von Statistikern muss der Aspekt der Aufdeckung von Betrug stärker betont werden.
- Statistiker sollten nicht zu viel über die Methoden zur Aufdeckung von Betrug veröffentlichen, damit unehrlichen Wissenschaftler nicht das Handwerkszeug gegeben wird um den Betrug zu verschleiern.

Hilfreiche Software

- **ContentMine:**
Datenextrahierungssoftware konvertiert beliebige Grafiken in einzelne Datenpunkte
- **Statcheck:**
R-package kann jedes HTML / PDF-Dokument nach Daten scannen und unpassende statistische Parameter erkennen
- point and click-Web-App unter <http://statcheck.io>.
(Artikel hochladen und die statcheck-Ergebnisse in einer CSV-Datei herunterladen)

John Antonakis

Die fünf Krankheiten der Wissenschaft

- **Significosis**, ein übermäßiger Fokus auf statistisch signifikante Ergebnisse
- **Neophilia**, eine übermäßige Wertschätzung für Neuheit
- **Theorrea**, eine krankhafte Fixierung auf neue Theorien
- **Arigorium**, ein Mangel an Disziplin in theoretischer und empirischer Forschung
- **Disjunctivitis**, eine Neigung zur Produktion großer Mengen überflüssiger, trivialer und inkohärenter Publikationen



Was kann man besser machen ?

- Veröffentlichungsdruck reduzieren
- negativen Ergebnissen in Veröffentlichungen Raum geben
- Ethik Unterricht auf der Universität
- einen erheblichen Betrag an Forschungsgeldern für Replikationen zur Verfügung stellen
- öffentliche Datenrepositorien einfordern
- Strafen ?

Plagiated resources

OSTP Federal Policy on Research Misconduct, cited from Martinson, B. C., Anderson, M. S., de Vries, R. (2005). Scientists behaving badly. *Nature*, 435, 737-738
DOI: 10.1038/435737a

Goodstein, D. (2005). Conduct and Misconduct in Science. Retrieved from <http://www.physics.ohio-state.edu/~wilkins/onepage/conduct.html>
[Friday, 11-Nov-2005 10:42:26 EST]

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly*, 29, 5-21.
DOI: 10.1016/j.leaqua.2017.01.006

Evans, S. (2001). Statistical aspects of the detection of fraud.
In Lock, S., Wells, F., Farthing, M. (Eds), *Fraud and Misconduct in Biomedical Research (186-203)*. London, England: BMJ Books,

Peter Doorn, Ingrid Dillo and René van Horik (2013). Lies, Damned Lies and Research Data: Can Data Sharing Prevent Data Fraud?
In *The International Journal of Digital Curation* Volume 8, Issue 1
DOI: 10.2218/ijdc.v8i1.256

Fanelli, D. Redefine misconduct as distorted reporting (2013). *Nature* 494, 149. DOI: 10.1038/494149a

Plagiated resources

The International Journal of Digital Curation Volume 8, Issue 1 | 2013

Hochschulrektorenkonferenz. *Zum Umgang mit wissenschaftlichem Fehlverhalten in den Hochschulen*; HRK: Bonn, Germany, 1998.

David B. Resnik, Lisa M. Rasmussen & Grace E. Kissling (2015) An International Study of Research Misconduct Policies, Accountability in Research, 22:5, 249-266, DOI: 10.1080/08989621.2014.958218

Michael L. Grieneisen, Minghua Zhang, A Comprehensive Survey of Retracted Articles from the Scholarly Literature Article *in* PLoS ONE · October 2012
DOI: 10.1371/journal.pone.0044118 · Source: PubMed

Heathers JA, Anaya J, van der Zee T, Brown NJ. 2018. Recovering data from summary statistics: Sample Parameter Reconstruction via Iterative TEchniques (SPRITE) PeerJ Preprints 6:e26968v1
DOI: 10.7287/peerj.preprints.26968v1

Brown NJL, Heathers JAJ. 2016. The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology. PeerJ Preprints 4:e2064v1
DOI: 10.7287/peerj.preprints.2064v1