

Jochen Kruppa & Björn Kiehne

## Statistik lebendig lehren durch Storytelling und forschungsbasiertes Lernen

### Zusammenfassung

Dieser Artikel stellt die Frage, wie Statistik so gelehrt werden kann, dass die Studierenden sie als wichtiges Werkzeug der Forschung, aber auch als Mittel zur Kommunikation zwischen den Berufsgruppen begreifen. Beispiele mit Elementen des Academic Storytellings und der forschungsbasierten Lehre aus der Bioinformatik und Biostatistik dienen dabei der Anschauung. Sie geben Anregungen dafür, Statistik sinnlich ansprechend, kognitiv anspruchsvoll, praktisch begreifbar und kontextbewusst anzubieten. Statistik zu erlernen ist mehr als den sicheren Umgang mit Zahlen zu lernen. Es geht um ein wichtiges konstituierendes Element von empirischer Forschung, das individuelle Erkenntnis und kollegiale Zusammenarbeit gleichermaßen prägt. Wir zeigen an verschiedenen Beispielen, wie vergangene und aktuelle Forschungsergebnisse mit in die Lehre integriert werden können. Darüber hinaus stellen wir einen Datensatz vor, der in den Vorlesungen aus Gummibärentüten erstellt wurde. Der Datensatz ist in zwei verschiedenen Versionen auf Anfrage bei den Autoren erhältlich.

### Schlüsselwörter

Forschungsbasierte Lehre; Learning based on enquiry; Academic Storytelling; erfahrungsbasiertes Lernen; Kompetenz (HQR)

## Teaching statistics vividly through storytelling and research-based learning

### Abstract

This article poses the question of how statistics can be taught in such a way that students will see the topic statistics as an important tool of research but also as a means of communication between professional groups. We use examples with elements of academic storytelling and research-based teaching from bioinformatics and biostatistics. We provide suggestions for offering statistics that are sensuously appealing, cognitively demanding, practically comprehensible and context-sensitive. Learning statistics is more than learning how to handle numbers reliably. It is about an important con-

stitutive element of empirical research that shapes both individual knowledge and collegial cooperation. Using various examples, we show how past and current research outcomes can be integrated into teaching statistics. In addition, we present a data set that was created during the lectures by using gummy bear sweets. The data set is available in two different versions on request from the authors.

### Keywords

Research-based teaching; learning based on enquiry; academic storytelling; experience-based learning; competence (HQR)

## 1 Einleitung

### 1.1 Ausgangsproblem

Forschung ist sozial komplex. Verschiedene, oft sehr unterschiedliche, Partner/-innen versuchen gemeinsam, neue wissenschaftliche Erkenntnisse zu gewinnen. Diese Komplexität kann ein Lernanreiz sein. Dieser Artikel beleuchtet deshalb die wissenschaftliche Ausbildung einer zukünftigen Statistikerin, die in enger Zusammenarbeit mit anderen Disziplinen neue Erkenntnisse gewinnen will. Zwischen Statistiker/-innen und anderen Forschenden herrscht oft Sprachlosigkeit. Dort setzen wir im Artikel an, sodass Studierende diese Sprachlosigkeit später in ihrem Beruf überwinden können. Wie kann es einer Statistikerin vermittelt werden, dass sie/er ein Teil des wissenschaftlichen Prozesses ist und sich auch als wichtige Akteurin darin sieht? Viel hängt davon ab, wie sie sich selbst sieht. Und wie sie sich selbst sieht, eignet sie sich während der Ausbildung an. Um erfolgreich in dem Beruf der Statistikerin/des Statistikers zu arbeiten, ist die Fähigkeit zur Methodenanwendung, wie bspw. das Programmieren, basal. Das Kommunizieren und informierte Entscheiden für Dritte fällt jedoch (in der Lehre) meist unter den Tisch. Im Zentrum dieses Artikels stehen Konzepte für eine Lehrveranstaltung um diesen Herausforderungen zu begegnen.

Wir werden in unserem Artikel versuchen, Wege und Konzepte anzubieten, um es Studierenden in Fächern mit Bezug zu Statistik zu ermöglichen, an späteren wissenschaftlichen Projekten besser, freier und aktiver teilnehmen zu können, was zum Vorteil aller am Projekt Beteiligten ist. Damit werden wichtige Impulse für die Entwicklung des Berufs- und Aufgabenverständnisses der Studierenden gesetzt. Somit behandelt dieser Artikel die Lehre der elementaren Fragen der Wissenschaft. Wie funktioniert Wissenschaft und wie funktioniere ich in ihr? Wie erlangen wir gemeinsam neues Wissen?

Die Statistik ist der Maschinenraum der quantitativen Forschung. Durch die Statistik können Forschungshypothesen abgelehnt und neue wissenschaftliche Hypothesen generiert werden. Vielen Studierenden wird der Zugang zum wissenschaftlichen Denken erschwert, da Statistik häufig mathematisch gelehrt, jedoch dabei vergessen wird, den gesamten Kontext der Forschung mit einzubeziehen. In diesem Artikel wird die Frage gestellt, wie Statistik anregend vermittelt werden kann. Dabei geht es nicht nur darum, Statistik lebendig zu erzählen, sondern um die Bewusstmachung des gesamten Kontextes der Zusammenarbeit: Datenerheber/-innen, Datenauswerter/-innen und Programmierer/-innen

arbeiten gemeinsam an einem Projekt und lernen voneinander. Als praktische Beispiele dienen in diesem Artikel Lernsequenzen aus der Biostatistik und Bioinformatik. Sie lassen sich aber auch leicht auf andere Studienfächer übertragen. Teile der präsentierten Methoden und Übungen wurden im Rahmen von *Jugend forscht* oder aber an der FU Berlin im B.Sc. Bioinformatik und dem M.Sc. Biomedical Sciences an der fhg – Zentrum für Gesundheitsberufe Tirol erprobt.

Im Zentrum der angewendeten Konzepte steht das forschungsbasierte Lernen und das Storytelling. Unser Fokus liegt hier auf der unmittelbaren Anwendung der Konzepte in der Lehre. Die Begrifflichkeiten des forschungsbasierten Lernens und des Storytellings setzen wir bei den Leser/-innen voraus und wiederholen nur kurz die Grundlagen. Wir ordnen die Begriffe ein und geben vertiefende Literaturquellen an. Wir nutzen den gewonnenen Raum, um die Lehrkonzepte besser zu beschreiben und direkt anwendbar zu machen. Unser Artikel ist somit relevant für alle Studiengänge, in denen Grundlagen der Statistik gelehrt werden und im späteren Verlauf wissenschaftliche (Abschluss-) Arbeiten verfasst werden. Somit deckt dieser Artikel mindestens alle sozial- und naturwissenschaftlichen Fächer ab, da in diesen Fächern Statistik und wissenschaftliche Grundlagen gelehrt werden müssten.

Im Folgenden werden wir zuerst die Begrifflichkeit und das Berufsbild der auszubildenden Statistikerin erläutert. Wir erklären den Kontext, in dem eine Statistikerin später arbeiten wird um die folgenden Lehrkonzepte besser nachvollziehen zu können. Hier können sich auch andere Disziplinen wie Psychologie, Wirtschaftswissenschaften oder der Bereich der Pflanzenwissenschaften wiederfinden. Nach der Vorstellung der Lehrkonzepte des forschungsbasierten Lernens und des Storytellings stellen wir die Reaktionen der Studierenden dar und geben Hilfestellungen für die eigene Anwendung. Alle Konzepte sind so ausgearbeitet, dass eine Anwendung mit wenig Adaption direkt möglich sein sollte. Eine analytische Auswertung war aus verschiedenen Gründen nicht möglich und abschließend von Seite der Autoren nicht gewollt.

Die Lehrkonzepte sind auf einen langfristigen Lernerfolg ausgelegt. Teilweise können die Studierenden den Sinn der Lehreinheiten erst in der Bachelorarbeit im folgenden Jahr erkennen. Darüber hinaus wurden die Veranstaltungen zentral von der fhg Tirol oder der Freien Universität Berlin evaluiert. Als externer Dozent war es nicht möglich, in die standardisierten Prozesse einzuwirken. Eine Nachverfolgung der Studierenden über eine längere Zeit ist für die Autoren nicht möglich und wäre ein eigenes Projekt.

## 1.2 Begrifflichkeit und Berufsbild der Statistikerin

In unserem Artikel benutzen wir die Berufsbezeichnung der Statistikerin. Wir wollen zuerst diese Bezeichnung näher definieren und die Statistikerin gegen andere Berufsgruppen und Tätigkeitsfelder abgrenzen. Wir verwenden diese Bezeichnung synonym für andere Berufsgruppen, wie Bio-Informatikerin, Bio-Metrikerin, Bio-Statistikerin, Data Scientist oder auch Data Analyst. Abgrenzen wollen wir die Statistikerin von der Mathematikerin oder Stochastikerin, die theoretisch mit Zahlenräumen umgeht, ohne einen direkten Anwendungsbezug in der angewandten Forschung. Wir grenzen uns in diesem Artikel auch von der methodisch, theoretisch forschenden Statistikerin ab, deren Ziel es ist, vorhandene mathematische Modelle zu verbessern und in Software zu implementieren.

Wir nennen daher eine Person eine Statistikerin, i) die mit fremden Daten arbeitet, ii) Daten an einem Rechner auswertet, iii) einen Report mit Entscheidungen über die Daten an Dritte weitergeben muss und iv) nicht mit allen Beteiligten in engem kollegialem Austausch steht. Mit fremden Daten zu arbeiten schließt ein, dass die Statistikerin nicht an der Erhebung der Daten beteiligt war und häufig die Person, die die Daten erhoben hat, nicht mehr zu erreichen ist. Häufig liegt den Daten eine veraltete wissenschaftliche Fragestellung zugrunde. Die Zeitspanne von einem Antrag, z. B. für eine klinische Studie, bis zum Erhalt der Daten kann sich über drei bis fünf Jahre erstrecken. In diesem Zeitraum ändert sich oftmals die personelle Zusammensetzung der Forschungsteams, so dass die ursprüngliche Fragestellung nicht mehr mit den Daten in Einklang zu bringen ist. Die Statistikerin muss daher auch in der Lage sein, die Fragestellung zusammen mit anderen Forscher/-innen weiterzuentwickeln. Dafür ist ein Verständnis der wissenschaftlichen Abläufe unbedingt notwendig. Im späteren Berufsleben werden nicht alle Kriterien immer gänzlich erfüllt.

Im Weiteren werden die Daten mindestens an einem Desktoprechner, wenn nicht gar auf Großrechnern, ausgewertet. Die Auswertung schließt die Aufarbeitung der Daten in ein maschinenlesbares Format sowie das Programmieren passender ausführbarer Skripte mit ein. Die Daten können aus verschiedenen prozessierenden Quellen zugesandt werden, wie etwa Proteinanalysen aus einem physischen Analysegerät, die mit Genanalysen aus einem anderen Analysegerät computergestützt verbunden werden sollen. Es gibt Ländergrenzen übergreifende Kooperationen.

Des Weiteren müssen die passenden Methoden für die Auswertung der Daten gewählt und angewendet werden. Die Ergebnisse der Analyse müssen für Dritte, die nicht an der Datenerhebung unmittelbar beteiligt sein müssen, aufgearbeitet und ihnen präsentiert werden. Hierfür müssen die Ausgaben der statistischen Auswertung in andere Formate, wie Word oder Excel, übertragen und Texte verfasst werden, die die Entscheidung für die Wahl einer bestimmten Methode erläutern. Die neueste Entwicklung ist hier R Markdown. Sie erlaubt es, statistischen Computercode mit erklärenden Texten zu verbinden (Xie, 2018). Häufig ist dieser Austausch mit Dritten ein iterativer Prozess, bei dem Personen, die der DatenerheberInnen nahestehen, bei der Aufarbeitung der Daten helfen. Abschließend wird ein Bericht an Entscheider/-innen über weitere Mittel für Experimente weitergeleitet. Daher beinhaltet die Arbeit der Statistikerin eine komplexe soziale Komponente: Die Statistikerin steht selten in direktem kollegialem Austausch mit den Datenerheber/-innen und den endgültigen Entscheider/-innen. Diese beiden Gruppen sind meist in anderen wissenschaftlichen Institutionen lokalisiert. So obliegt es häufig der Statistikerin, den unterschiedlichen wissenschaftlichen Hierarchiegraden, wie z. B. Doktorand/-innen, PostDocs, Oberärzt/-innen oder Professor/-innen, die Ergebnisse der statistischen Analysen entsprechend aufzuarbeiten und zu präsentieren. Die Statistikerin muss dabei förmlich in fremden Köpfen denken, um die Ergebnisse ihrer Analysen Fachfremden in einer ihnen gemäßen Sprache zugänglich zu machen. Wir wollen dies am Begriff *Korrelation* verdeutlichen. Für eine Statistikerin beschreibt *Korrelation* ein enges methodisches Feld, um die Abhängigkeit zweier Variablen in einem Zahlenraum von -1 bis +1 abzubilden. Im Sprachraum der Medizin beschreibt *Korrelation* allgemein einen Vergleich zweier Variablen mit allen möglichen sta-

tistischen Methoden, vor allem aber auch Gruppenvergleiche, die per se mit einer klassischen Korrelation statistisch/methodisch nicht durchführbar und auch nicht gewollt sind (Mukaka, 2012).

Wir verwenden im Folgenden die Bezeichnung Statistikerin für die Beschreibung des im obigen Abschnitt erläuterten Berufsbildes und des Ziels der entsprechenden universitären Ausbildung.

### 1.3 Forschendes Lernen und Storytelling

Forschungsbasiertes Lernen und auch zum Teil Storytelling gelten als Königsdisziplin explorativer Lehr- und Lernformen (Decker und Mucha, 2018). Beide Formen der Lehre sind nicht neu und theoretisch ausgiebig beschrieben. Wir wollen uns daher in diesem Artikel auf die Anwendung des forschungsbasierten Lernens und Storytelling in Nebenfächern der Statistik konzentrieren. In diesen Nebenfächern kann die intrinsische Motivation der Studierenden an Statistik niedriger sein, da das Fach aus der Sicht der Studierenden nicht direkt mit dem Hauptfach zu tun hat. Diese Verbindung von Haupt- und Nebenfach wollen wir durch forschendes Lernen und Storytelling herstellen.

Decker und Mucha (2018) zeigen die emotionalen Herausforderungen der Integration von Lernen durch Forschung am Beispiel eines konsekutiven Masterstudiengangs im Studiengang *International Business*. Um Studierende mit einer weniger ausgeprägten Ungewissheitstoleranz nicht zu verlieren, haben wir die forschungsbasierten Konzepte in kleinere Einheiten aufgeteilt. Da es sich beim forschungsbasierten Lernen – neben der Auseinandersetzung mit der Sache – auch um eine Auseinandersetzung mit sich selbst handelt, haben wir versucht, die Einheiten möglichst kleinschrittig zu halten und somit kein Anleitungsvakuum bei den Studierenden aufkommen zu lassen. In unserem Artikel dienen Huber (1970) und Huber (2014) als Grundlage für die Theorie des forschungsbasierten Lernens. Wir empfehlen die entsprechende Literatur als erweiterte Grundlage. Decker und Mucha (2018) fassen Huber (2014) entsprechend der begrifflichen Abgrenzung noch einmal zusammen und unterscheiden: Lernen *über* Forschung (Forschungsbasiertes Lernen), Lernen *für* Forschung (Forschungsorientiertes Lernen) und Lernen *durch* Forschung (Forschendes Lernen). Die in diesem Artikel vorgestellten Konzepte basieren weniger auf Lernen *über* Forschung, um Forschung als Beitrag zur Lösung von offenen Fragen zu erkennen, sondern mehr auf Lernen *für* Forschung, um die Grundlagen für die Durchführung von Forschungsaufträgen von Dritten zu ermöglichen. In Ansätzen lässt sich durch die Arbeit am Gummibären Datensatz (Abschnitt 2.1) oder dem Theaterstück (Abschnitt 2.3) auch Lernen *durch* Forschung zusammen mit den Studierenden erleben.

Neben dem Konzept des forschenden Lernens wollen wir auch die Kraft der Geschichte (englisch Storytelling) nutzen, um gelehrtes Wissen in den Köpfen der Studierenden zu verankern. Storytelling nutzt verschiedene Aspekte des menschlichen Lernens, wie zum Beispiel emotionale Reize, um Wissen längerfristig in Erinnerung zu halten. Wir können hier als vertiefende Literatur Wolfe (2006), veröffentlicht in Johnson und Taylor (2011), empfehlen, die spannende Grundlagen des humanen Lernens näher betrachten. Wolfe (2006) diskutiert insbesondere die emotionalen Grundlagen und das damit verbundene langfristige Lernen. Zazkis und Liljedahl (2019) beschreiben in ihrem Buch detailliert das Storytel-

ling für mathematische Fächer und gehen dabei auf die verschiedenen Aspekte des Storytellings näher ein: Plot, Conflict, Images, Human meaning, The sense of wonder, Humor, Patterns und Summary. Wir nutzen diese in dem Theaterstück über Robert Koch (Abschnitt 2.3) sowie in unserem Schifffahrtbeispiel um James Lind (Abschnitt 2.4). Hansson et al. (2019) beschreibt durch den Besuch von Rutherford in einer Mittelschule ein ähnliches Setting wie unser Theaterstück. Schlussendlich lässt sich der Bogen des Storytellings bis zur Veröffentlichung von Artikeln, wie dieser einer ist, spannen. Schimel (2012) beschreibt die Notwendigkeit einer Geschichte, die sich spannend durch Artikel sowie Anträge ziehen muss, um zitiert und beachtet zu werden. Hier schließt sich der Kreis für uns Autoren als Lehrende und Publizierende.

## 2 Konzept – Lernen durch Research Stories

Im Fokus der Planung einer Lehrveranstaltung im Rahmen der Ausbildung zur Statistikerin sollte das Vermitteln des Gesamtbildes stehen. Es sollte darauf geachtet werden, sich nicht in den methodischen Details der statistischen Ausbildung zu verlieren. Im Idealfall begleitet eine Statistikerin ein Forschungsvorhaben von Beginn an. Die eigentliche Fragestellung in einer Studie bzw. einem Experiment muss in kleinteilige statistische Hypothesenpaare zerlegt werden. Neben der Fallzahlplanung und der Diskussion der Effektgrößen spielt auch die Begleitung der Studie im weiteren Verlauf eine wichtige Rolle. Die erhobenen Daten müssen maschinenlesbar abgelegt werden, um auch später von der Statistikerin ausgewertet werden zu können. Schlussendlich ist das Ziel der Forschenden meist nicht eindeutig. Medizinische Doktorand/-innen wollen eine Monographie schreiben, gleichzeitig wollen die Betreuer/-innen aus der Arbeit noch eine Publikation erstellen. Die unterschiedlichen Ziele müssen von einer Statistikerin abgewogen und kommuniziert werden. Dabei kommt einer Statistikerin häufig die – manchmal ungewollte – Rolle der Vermittlerin zu.

Es gibt im Bereich der Statistik eine Vielzahl an Verfahren, um aus Daten Entscheidungen für weitere Experimente oder Handlungsanweisungen zu gewinnen. Erschwert werden diese Entscheidungen durch die Namensgebung der jeweiligen Entwickler/-innen der statistischen Methode. So sind die Verfahren nicht nach Bereich oder Methodik benannt, aus dem diese Verfahren stammen, sondern nach den Namen der Entwickler/-innen: Student t-Test, Fishers exakter Test, Kruskall-Wallis-Test und viele weitere. So sind meist auch erfahrene Statistiker/-innen überfragt, in welchen Bereich denn nun ein Test gehört, wenn dieser Test nicht zu den häufig verwendeten Verfahren zählt.

Es obliegt nun der Lehrperson, die Konzepte hinter den statistischen Verfahren aufzuzeigen. Wie werden statistische Hypothesenpaare gebildet? Was ist die Idee der wissenschaftlichen Entscheidungsfindung nach Karl Popper (Susser, 1986)? Was ist das Falsifikationsprinzip? Sucht die Wissenschaft nach Wahrheit oder ersetzt die Wissenschaft ein schlechtes Modell durch ein weniger schlechtes? Welche Positionen in dem Prozess der wissenschaftlichen Wissensgewinnung nehme ich als Statistikerin ein? Welche Rollen muss ich bedienen, welche Rollen kann ich bedienen und welche Rollen und Fragen kann ich nicht einnehmen? Wie gehe ich mit unerwarteten Ergebnissen um und wie halte ich diese Unsicherheit der Forschung aus? Um diese komplexe Situation auch in größeren Vorlesun-

gen den Studierenden näherbringen zu können, möchten wir in diesem Artikel verschiedene Formen des Story Telling und der forschungsbasierten Lehre vorstellen. Wie kann man den Studierenden der verschiedenen wissenschaftlichen Disziplinen, wie Informatik, Bioinformatik oder Statistik, das statistische Denken näherbringen (Wild und Pfannkuch, 1999)? Vorschläge hierzu bietet der GAISE – *Guidelines for Assessment and Instruction in Statistics Education* – College Report (Carver et al., 2016; Wood et al., 2018). Im GAISE werden die Anforderungen an den Unterricht für Statistik in einer College-Ausbildung wie folgt beschrieben (Carver et al., 2016, S. 3, Übersetzung aus dem Englischen):

1. „Unterrichte statistisches Denken mit dem Fokus auf einen Ermittlungsprozess für Problemlösung und Entscheidungsfindung, besonders im Sinne von multivariablen Problemstellungen;“
2. „Fokussiere auf konzeptionelles Verstehen;“
3. „Integriere echte Daten mit Kontext und Sinn;“
4. „Rege zum aktiven Lernen an;“
5. „Nutze die Technologie, um Konzepte zu erfahren und Daten zu analysieren;“
6. „Nutze die Beurteilungen, um das studentische Lernen zu verbessern und zu evaluieren.“

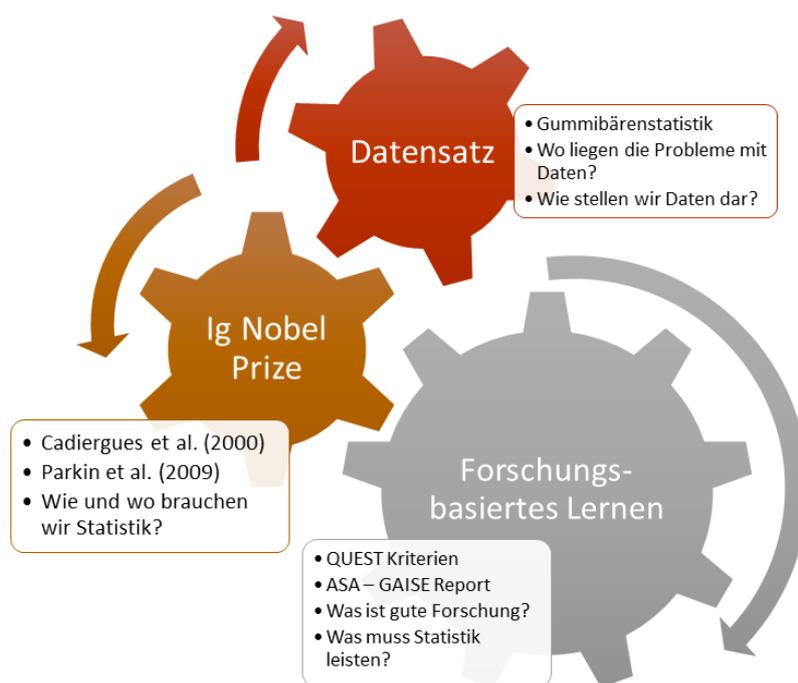


Abbildung 1: Grundsätzliche Verzahnung der einzelnen Ideen im Modul. Der Datensatz zur Gummibärenstatistik wird von den Studierenden (siehe Anhang) selbst erstellt und wird immer wieder in den Softwareübungen genutzt. Der ig Nobel Prize stellt als leicht verständliche wissenschaftliche Referenz die Verbindung zur echten Forschung dar.

In Abbildung 1 stellen wir die Verzahnung der in unserem Artikel vorgestellten Konzepte dar. Zuerst werden Daten benötigt, mit denen analytisch gearbeitet und an denen der Umgang mit Daten erlernt werden kann. Diese Daten werden im späteren Berufsleben von Dritten geliefert. Woher sollen diese Daten genommen werden und wie finden wir möglichst realistische Datensätze, die auch eine persönliche Relevanz für die Studierenden

haben? Im Weiteren wird dargestellt, wie den Studierenden ein besseres Verständnis des Gesamtbildes der Forschung ermöglicht werden kann. Das Publizieren von wissenschaftlichen Erkenntnissen soll hierbei im Zentrum stehen. Abschließend werden die Fragen, was gute Forschung ist und wie die Idee der Forschung Zugang zum sozialen Lebensbereich findet, beantwortet und durch zwei Beispiele verdeutlicht.

Abschließend steht die Frage, was gute Forschung ausmacht, im Fokus der Veranstaltung. Die Studierenden sollen zum selbstständigen Arbeiten und Forschen angeleitet werden.

## 2.1 Woher realistische Daten nehmen?

Ein Großteil der Arbeit von Statistiker/-innen besteht in der Tätigkeit der Bereinigung der Daten (auch *Daten schubsen* genannt). Die statistischen Analysemethoden bzw. die statistischen Analyseprogramme benötigen meist eine einzige mögliche und exakte Art der Datenaufarbeitung, sonst ist das Einlesen der externen Daten nicht möglich. Programmiersprachen sind zum Beispiel nicht in der Lage, Spalten in Datensätzen mit Worten (engl. *character*) und gleichzeitig beinhaltenden Zahlen (engl. *numeric*) sinnvoll zu interpretieren. Um die (soziale) Komplexität der Datenerzeugung zu simulieren, haben wir die Daten von den Studierenden im B.Sc.-Modul *Statistik für Biowissenschaftler I* und dem M.Sc.-Modul *Biomedical Sciences* eigenständig erheben lassen. Die Studierenden bekommen hierbei ein Gefühl dafür, wo die Daten ihren Ursprung haben und wie die Prozesse der Datensammlung und Datenverarbeitung miteinander zusammenhängen. Ein weiterer Punkt stellt die Sensibilisierung für den Umfang des Aufwands der Datenerhebung sowie die hohe Fehleranfälligkeit der Dateneinlese dar, die weitreichende Folgen nach sich ziehen kann (Abbildung 2 der originalen Daten im Anhang). Dazu dienten uns zum einen kleine Gummibären-tüten und allgemeine (bio-) metrische Angaben der Studierenden.

Die Datenerhebung lief wie folgt ab. Jede/-r Studierende erhielt zufällig eine nummerierte Gummibärenpackung und die Information, aus welchem Beutel diese Gummibärenpackung stammte. Normalerweise befinden sich ca. 22 Packungen in einer Großpackung Gummibären. Die Studierenden sollten nun folgende Informationen erfassen: Lfd. Nummer Packung, Lfd. Nummer Beutel, die Anzahl der Farben (Wie viele sind dunkelrot, hellrot, orange, gelb, grün und weiß?), die Anzahl der Gummibären sowie biometrische Daten der Studierenden, wie Lieblingsgeschmack, Geschlecht, Alter, Größe und Hochschulsesemester. Die Angaben waren freiwillig und durch die zufällige Zuordnung der Packungen nicht mehr aufzulösen. Abbildung 2 zeigt einen Screenshot der Liveeingabe der Daten, Abbildung 3 das Ergebnis der Dateneingabe der Studierenden.

Aufgabe war es nun, die Daten in die Statistiksoftware R einzuspeisen und sie nur mit Programmieranweisungen, wie durch reguläre Ausdrücke (engl. *regular expressions*), zu bereinigen. Dabei sollten die ursprünglichen Daten nicht in Excel modifiziert werden, sondern nur in R selbst per Programmierbefehlen. Die Daten lassen sich auf verschiedene Art und Weise bearbeiten. Wir nutzen die Rohdaten, so wie sie die Studierenden erhoben haben, für die Ausbildung von Informatiker/-innen, um ein Verständnis für Hindernisse beim Einlesen von Daten zu fördern: Falsche Kodierung, Wörter in Zahlenspalten, angehängte Leerzeichen, Umlaute oder falsche Zellenbelegung. Ist der Fokus mehr auf dem Auswerten

bzw. Modellieren der Daten als der Bereinigung, so nutzen wir einen schon angepassten Datensatz, in dem die größten Fehler schon behoben wurden.

	A	B	C	D	E	F	G	H	I	J	K
1	Packung	Beutel	Dunkelrot	Hellrot	Orange	Gelb	Grün	Weiß	Anzahl Bären	Anzahl Farben	Lieblingsschm
2	1	1	0	0	5	4	0	0	10	3	Rot
3	1	2	0	3	1	4	1	1	10	5	Yellow
4	1	3	1	2	2	2	1	1	9	6	White
5	1	4	2	0	2	1	2	3	10	5	Weiß
6	1	5	2	1	1	2	2	2	10	6	Weiß
7	1	6	daten ändern du bist in mein feld		4	2	2				
8	1	7	2	4	1	2	0	1	10	5	
9	1	8	2	1	2	3	1	1	10	5	grün
10	1	9	2	1	4	0	4	2	13	5	grün
11		10		1	1	2	1	1			
12	3	11									
13	1	12	2		3	2	2	0	10	5	Rot
14	2	13	2	0	2	4	3	1	12	5	Orange
15	1				3						
16	2	14	1	3	3	2	0	1	10	4	Rot
17	1	15	2	1							

Abbildung 2: Screenshot der Liveeingabe der Daten aus den Gummibärentüten. Zeitweise waren bis zu 32 Studierende gleichzeitig damit beschäftigt, ihre Daten in das Dokument einzugeben. Die Eingabe erfolgte über Handy oder Laptop. Die Spaltennamen waren vorgegeben und die Tabelle wurde vorab einmal dem Plenum erläutert.

	A	B	C	D	E	F	G	H
1	Knast	Zelle	Dunkelrot	Hellrot	Orange	Gelb	Grün	
41	2			3	4	0	1	2
42	2			1	2	4	2	1
43	5	6		2	1	1	3	2
44								
45	5	42*42		1	0	3	3	0
46	animalcruelty							
47	1	13		0	2	4	2	2
48	1	10		2	0	1	5	
49	4				1	3	1	4
50	4				1	3	2	2
51	yo yo yo							
52	4			2	2	2	4	0
53	5	5		1	1	2	5	1
54	1	2		0	1	4	2	2
55	5	20		1	1	3	3	2
56		4			Bioinformatiker	sind besser	als Mediziner	
57		4		1	0	2	4	0
58								
59	2	8		2	1	3	2	1
60	2	9		2	1	2	4	0
61	2	10		2	3	1	1	0
62	2				1	2	3	2
63	3	3		2	3	3	1	0
64	5	8		0	1	5	3	0

Abbildung 3: Screenshot des unteren Endes der Ergebnistabelle nach der Eingabe der Studierenden. Die oberste Zeile ist fixiert. Am Ende haben 74 Studierende 64 Zeilen Daten produziert. Es sind viele Fehlstellen zu beobachten. In den folgenden Übungen müssen diese Daten von den Studierenden verwendet und dafür entsprechend aufgearbeitet werden. Diese Aufarbeitung muss durch die Studierenden selbstständig erfolgen.

## 2.2 Das wissenschaftliche Gesamtbild erläutern durch den ig Nobel Prize

Wie soll den Studierenden nun nähergebracht werden, wohin die Reise am Ende der statistischen Analyse gehen soll? Die Statistikerin hat nicht die Daten erhoben, liefert aber auch einen methodischen Bericht mit Ergebnissen, die Dritte fachlich bewerten und veröffentlichen wollen. Daher muss der Statistikerin bewusst sein, wie eine wissenschaftliche Publikation aufgebaut ist und, überdies, wie die Genese der wissenschaftlichen Hypothese zur statistischen Hypothese und zur abschließenden fachlichen Interpretation vonstattengeht.

Es würde sich nun anbieten, klassische medizinische Beiträge in der Vorlesung als Hausarbeit oder im Rahmen des Flipped Classroom zu bearbeiten (Ulrich, 2016). Hierbei sei jedoch auf die hohen Verständnisschwierigkeiten beim Lesen einer medizinischen wissenschaftlichen Arbeit hingewiesen. Das Problem des Verständnisses ergibt sich aufgrund der Verwendung der englischen Sprache sowie von medizinischen Fachwörtern. Der Kontext einer medizinischen Arbeit erschließt sich nicht sofort, da in der aktuellen Forschung keine Review-Artikel geschrieben werden und die medizinischen bzw. biologischen Zusammenhänge nur kurz angerissen werden. Auch ist das Ziel irreführend. Die Statistikerin soll lernen, der Medizinerin zu helfen, eine solche Arbeit zu verfassen. Sie arbeitet nur in Teilen der Veröffentlichung zu. Daneben beschreiben Pöpel und Morisse (2019) weitere Probleme des Flipped Classroom. So ist die Frustschwelle beim Lesen von englischen medizinischen Aufsätzen sehr niedrig, hier könnten deutsche Texte, wie aus dem Deutschen Ärzteblatt, einen Ausweg anbieten. Studierende brechen hier eher die Vorarbeiten ab, da die Texte auch sehr lang sind. Hier bieten sich Unterstützung durch weitere Medien wie z.B. begleitende Videos an.

Wir umgehen dieses komplexe Problem durch eine andere Auswahl von Beiträgen. In der Vorlesung und Übung werden wissenschaftliche Beiträge behandelt, die mit dem ig Nobel Prize (Abrahams, 2002) ausgezeichnet wurden. Zum einen sind dies reale Veröffentlichungen in anerkannten wissenschaftlichen Zeitschriften. Zum anderen sind die Themen und Fragestellungen originell und leicht zu verstehen. Die biologische Fragestellung steht hierbei nicht im Vordergrund, sondern kann durch einen methodischen Einstieg vereinfacht werden. So nutzen wir den wissenschaftlichen Beitrag von Cadiergues et al. (2000), welcher sich mit dem Gruppenvergleich von Sprunghöhe und Sprungweite von Katzen und Hundeflöhen beschäftigt, für die Vorstellung von Zweistichproben tests. Das Prinzip der klinischen Studien, das CONSORT Statement (Boutron et al., 2015) und die QUEST-Kriterien (Kip und Dirnagl, 2018), erklären wir anhand der Studie von Parkin et al. (2009), die sich in einer klinischen Feldstudie mit der Frage der Abwendung von Sturzschäden in Wintermonaten beschäftigt. Die Arbeit von Unger (1998) kann hervorragend dazu genutzt werden, eine wissenschaftliche Zusammenfassung (engl. abstract) zu veranschaulichen. In seiner Zusammenfassung erläutert er seine Forschung an der Fragestellung, ob das Knacken der Fingerknöchel zu Arthritis in den Händen führt.

Die wissenschaftlichen Beiträge haben mehrere klare Vorteile. Zum einen ist sofort die medizinische bzw. biologische Fragestellung klar. Der Fokus kann auf die statistischen Methoden und die Wahl der Analysewege gelegt werden (d.h. welcher Test in welcher Reihenfolge durchzuführen ist) und driftet nicht in eine biologische Diskussion über die Sinnhaftigkeit der Veröffentlichung ab. Dies ist häufig der Fall, wenn ein/-e Spezialist/-in aus

dem Forschungsfeld, in das der Aufsatz gehört, anwesend ist. Im Weiteren sind die Arbeiten häufig kurz und in einem stilistisch ansprechenden Englisch verfasst. Nach der Lektüre und der Diskussion der Methoden lässt sich auch die Fragestellung beantworten: Warum hat diese Arbeit einen ig Nobel Prize gewonnen? Hier lässt sich nachzeichnen, was eine gute Forschungsarbeit ausmacht, wie eine gute Forschungsarbeit geschrieben sein sollte und wie die Zuarbeit der Statistikerin – die Aufwertung und Abbildung von Daten – in einen größeren Veröffentlichungskontext eingebunden wird. Schlussendlich bietet der ig Nobel Prize nicht nur eine Quelle für Artikel aus den Bereichen Medizin und Biologie, sondern auch aus der gesamten Breite der Wissenschaften, von Wirtschaftslehre bis Mathematik und Ökologie.

### 2.3 Was ist wahr und was ist falsch? Robert Koch vs. Max von Pettenkofer

Eine zentrale Frage der Wissenschaft lautet: Was ist wahr und was ist falsch? Die Idee, auf der Suche nach der Wahrheit und dem, „was die Welt im Innersten zusammenhält“ (Goethe, 1808, S. 35), zu sein, fasziniert seit jeher die Menschen. Studierende haben ebenso diesen Antrieb und möchten die inneren Zusammenhänge erfahren und verstehen. Der Weg dorthin ist jedoch meist kurvig und nicht alle wissenschaftlichen Theorien können sich durchsetzen und als weniger falsch als andere Theorien angesehen werden. Dieser komplexe Weg kann sich über Jahrzehnte und ganze Forscherleben erstrecken. Häufig sind sich erst nachfolgende Generationen einig und dies auch wiederum nur vorläufig, wer auf dem richtigen wissenschaftlichen Pfad gewesen ist. Dieser Teil des Gesamtbildes ist schwierig an einem aktuellen Beispiel zu verdeutlichen, denn den Ausgang einer aktuellen Debatte, egal wie spannend die Debatte sein mag, kennen wir noch nicht.

Inspiziert von Max von Pettenkofer und seinem fruchtbaren Disput mit Robert Koch wurde die Idee entwickelt, anhand dieses wissenschaftlichen Konflikts in Bezug auf die Frage, wodurch oder wie Cholera ausgelöst wird, ein Theaterstück oder aber eine Übung zu generieren. Cholera wird von Bakterien ausgelöst. Dieses Faktum war jedoch zu der Zeit von Koch und von Pettenkofer umstritten. Von Pettenkofer glaubte fest daran, dass die Ursachen falsche Hygiene sowie schlechte Boden- und Wasserbeschaffenheit seien. Von Pettenkofer schluckte 1892 eine Kolonie von Cholerabakterien und bekam auch eine heftige Diarrhöe. Dies stimmte ihn aber nicht um, dass Cholerabakterien Cholera auslösen und zum Tode führen würden. Er wiederholte den Versuch an einigen seiner Studierenden und auch diese bekamen nur leichte Diarrhöe und keiner verstarb, was von Pettenkofer als Bestätigung für seine Theorie ansah (Raschke, 2008). Bei alledem darf nicht vergessen und muss auch ausdrücklich hervorgehoben werden, dass von Pettenkofer nach streng naturwissenschaftlich-experimentellen Regeln arbeitete und als Begründer der experimentellen Hygiene gilt. Von Pettenkofer war ein Positivist und erkannte nur im Experiment sichtbare Ergebnisse an. Vermutlich nur durch den Wettstreit der Ideen mit von Pettenkofer war es Robert Koch überhaupt möglich, seine Theorie über das Cholerabakterium so weit zu entwickeln. Von Pettenkofer erschuf auch die Idee der *atmenden Wand*, da er bei einem Experiment nicht alle Faktoren kontrolliert hatte und übersah, dass der Luftaustausch auch über den Kamin stattfand (Pettenkofer, 1858). Von Pettenkofer postulierte, dass der Luftaustausch in Räumen bei verriegelten Fenstern und Türen durch die Wände erfolgen musste.

Für das Theaterstück in einer Kleingruppe und die Übung in einer größeren Vorlesung haben wir einige dramaturgische Annahmen getroffen: i) die Chronologie wird aufgebrochen, alle Ereignisse finden auf der Cholera-Konferenz 1867 in Weimar statt, ii) Robert Koch und von Pettenkofer führen die Diskussion mit Teilwissen von heute, daher teilten wir nur einen Kurzlebenslauf von Robert Koch und von Pettenkofer aus, iii) Ereignisse wurden in ihrer Länge dramaturgisch stark gekürzt und iv) Gastwissenschaftler aus der jüngeren Vergangenheit, wie Feynman oder Einstein, waren auf der Konferenz zugegen (Feynman, 1974). Es empfiehlt sich, mit den Studierenden auf großen Bögen ein Screenplay zu entwerfen. Dies kann auch in Gruppenarbeit um jeweils einen der Hauptdarsteller geschehen. Weiterhin wäre es von Vorteil, wenn die Konferenz wiederholt geprobt würde und sich so weiterentwickeln könnte. Durch das wiederholte Spielen der Rollen finden sich die Studierenden besser in ihren Robert Koch oder Max von Pettenkofer ein. Tabelle 1 zeigt ein beispielhaftes Drehbuch. Dieses Drehbuch kann auch als Übung verwendet werden, indem die Studierenden das Streitgespräch selbst in kurzen Sätzen niederschreiben müssen. Für Studierende niedrigerer Semester kann diese Übung mit Fragen in noch kleinere Abschnitte zergliedert werden. Ein Seminar von ca. 20 Personen lässt sich auch in vier Aufführungen aufteilen, die dann nacheinander gezeigt werden. Aus Erfahrung ist jede Aufführung so anders, wie die Studierenden es sind und bereichert den Unterricht mit neuen Ideen und Blickwinkeln.

Tabelle 1: Beispielhaftes Drehbuch der Cholera-Konferenz 1867 in Weimar

<b>Dramaturgische Handlung</b>	
1.	Auftritt von Robert Koch und Darstellung seiner Position
2.	Auftritt von Max von Pettenkofer und Darstellung seiner Position
3.	Argumente von Robert Koch und Wiederholung seiner Verdienste an der Wissenschaft
4.	Argumente von Max von Pettenkofer und seiner Idee des Positivismus sowie seiner experimentellen Erfolge. <i>Optional: Die Atmende Wand und weiterer Unterplot.</i>
5.	Auftritt von Feynman und Darstellung der Cargo Cult Science sowie eventueller Parallelen zur (aktuellen) Forschung
6.	Studierende/r aus dem Plenum stellt die Erkenntnisse von Koch und von Pettenkofer gegenüber: Koch behauptet, Cholera werde durch ein Bakterium ausgelöst, laut von Pettenkofer jedoch durch schlechte Hygiene. Versuch der Vermittlung, dass diese beiden Sachen eventuell zusammenhängen.
7.	Widerworte von Koch und von Pettenkofer zu den Thesen der/des Studierenden
8.	Von Pettenkofer testet die Cholera an sich selbst und krümmt sich vor Schmerzen. Beweis, dass es nicht Cholera ist, welche Menschen tötet.
9.	Robert Koch widerspricht. Von Pettenkofer ruft drei Studierende, die ebenfalls Cholera trinken. Wieder stirbt niemand.
10.	Studierende/r aus dem Plenum trägt ein Resümee vor

## 2.4 Ein Schiff kann nur 111 Tage zur See fahren

„Ohne ausreichende Zufuhr von Vitamin C stellen sich nach 68 Tagen die ersten Symptome ein; die ersten Toten sind nach 84 Tagen zu beklagen; nach 111 Tagen rafft die Skorbut eine

ganze Schiffsbesatzung dahin.“ Europa und das Meer – Deutsches Historisches Museum, Berlin aus Blume et al. (2018, S. 77)

Überlebenszeitanalysen oder Ereigniszeitanalysen stellen Studierende immer wieder vor größere Probleme. Der methodische Hintergrund ist sehr komplex und die Fragestellung den Studierenden aus ihrer Alltagserfahrung vollkommen unbekannt. Hinzu kommt, dass die Ergebnisse einer Überlebenszeitanalyse in einem Stufendiagramm dargestellt werden, das speziell für diesen Analysetyp in der Form angewendet wird (Kaplan-Meier-Kurve). Ausgang für unsere Geschichte um die Überlebenszeitanalyse ist das obige Zitat aus der Ausstellung „Europa und das Meer“ des Deutschen Historischen Museums Berlin. Das Beispiel eignet sich besonders, um das statistische Denken nach Wild und Pfannkuch (1999) zu trainieren.

James Lind war ein Pionier der Bordhygiene und entdeckte die Therapie von Skorbut durch Zitronensaft. Lind führte erste experimentelle Versuche durch, um die Ursache von Skorbut zu entdecken. Er teilte zwölf an Skorbut erkrankte Matrosen in sechs Gruppen ein und behandelte jede dieser Gruppen anders, z. B. mit Schwefelsäure, Essig oder Apfelsinen und Zitronen. An diesem Beispiel kann man sehr schön die Gruppeneinteilung, Fallzahl sowie Konzepte wie *doppelblind* und Randomisierung in einer klinischen Studie darstellen und erarbeiten. Ethische Fragen lassen sich an dem konstruierten Beispiel erörtern, dass James Cook sofort, nachdem er von den Ergebnissen von Linds Versuch erfahren hatte, mit seinen Schiffen aufbrach, ohne den Studienabschluss abzuwarten, und allen Matrosen Zitronen gab. Hier lassen sich die Regularien und Beschränkungen einer klinischen Studie, wie Interimsanalyse oder Studienabbruch, diskutieren.

Die Studierenden sollen die sich ergebende Überlebenszeitanalyse graphisch darstellen und die Limitierungen der Studie erarbeiten. Beispielhaft lässt sich die Überlebenszeitkurve mit 100 Matrosen gleich 100 % Überleben starten. Nach 68 bis 84 Tagen ergeben sich die ersten drastischen Stufen, die dann zu 0 % am 111 Tag verlaufen. Hier lässt sich dann auch eine Vergleichskurve unter der Gabe von Vitamin C mit auftragen und diskutieren, ob diese Kurve immer bei 100 % über die 111 Tage läuft, da schließlich auch anderweitig Todesfälle auftreten können, die nichts mit Skorbut zu tun haben. Die natürliche Variabilität eines Experiments lässt sich anhand eines Beispiels noch einmal sehr gut verdeutlichen.

### 3 Einsatz der vorgestellten Konzepte

Die in dem obigen Abschnitt beschriebenen Konzepte wurden in drei verschiedenen Veranstaltungen angewendet. Diese unterschieden sich sowohl hinsichtlich der Anzahl der Teilnehmenden als auch bezüglich ihres Wissensstandes sowie der Dauer der Veranstaltung.

Tabelle 2: Kurzübersicht über die Veranstaltungen, in denen die vorgestellten Konzepte angewendet wurden

<b>Universität</b>	Freie Universität Berlin	fhg – Zentrum für Gesundheitsberufe Tirol	Stiftung Jugend forscht e.V.
<b>Studiengang</b>	B.Sc. Bioinformatik	M.Sc. Biomedical Sciences	World Health Summit 2018 Berlin
<b>Semester</b>	3.	1.	1.-6.
<b>Veranstaltung</b>	Statistik für Biowissenschaftler I	Biomedizinische Analytik	Workshop Perspektiv-Forum
<b>Dauer</b>	Semesterbegleitend (~ 28 UE)	Semesterbegleitende Blockkurse (~20 UE)	Zweitägiger Blockkurs (~16 UE)

### 3.1 Bachelorstudiengang 3. Semester Bioinformatik an der FU Berlin

Die vorgestellten Konzepte wurden ursprünglich für das 3. Semester im Bachelorstudiengang Bioinformatik an der Freien Universität Berlin für die Vorlesung *Statistik für Biowissenschaftler I* konzipiert. Die Veranstaltung setzte sich aus Teilnehmenden der Studiengänge M.Sc. Informatik (10 % der Studierenden), B.Sc. Bioinformatik (70 % der Studierenden) und M.Sc. Statistik (20 % der Studierenden) zusammen. Insgesamt nahmen 100 Studierende an der Vorlesung und den Übungen teil. Die Anzahl der anwesenden Studierenden variierte stark im Laufe der Vorlesungszeit und belief sich im Mittel auf etwa 50 Studierende (100-120 angemeldete Studierende), größtenteils Studierende des B.Sc. Bioinformatik. Die Teilnahme war freiwillig. Die Veranstaltung wurde abschließend evaluiert. Folgende Konzepte wurden angewandt: i) Erstellung und Auswertung des Gummibären Datensatzes, ii) Lesen und Nachrechnen von ig Nobel Prize Veröffentlichungen, iii) Diskussion der Forschung von Robert Koch versus Max von Pettenkofer und iv) Lernen von statistischem Denken anhand des Skorbut-Beispiels.

Es bot sich nun an, die Veranstaltungen *Deskriptive Statistik* sowie *Statistisches Testen I & II* thematisch enger miteinander zu verbinden und neu zu gestalten. Als Grundlage dafür wurden die Veranstaltungen *Einführung in klinische Studien* und *Evidenzbasierte Medizin* überarbeitet und angepasst. Dabei wurden den Studierenden die Inhalte der Vorlesung in neuer Form präsentiert. Im Mittelpunkt der Vorlesung und Übung zu den klinischen Studien standen die Beispiele der Konkurrenz von Robert Koch und von Pettenkofer sowie dem Überleben von Matrosen auf See. Dies fand jedoch im Rahmen einer schriftlichen Übung und der Diskussion im Plenum statt. Die Grundlagen der evidenzbasierten Medizin wurden an dem Beispiel der mit dem ig Nobel Prize ausgezeichneten Studie von Parkin et al. (2009) „Preventing Winter Falls: A Randomised Controlled Trial of a Novel Intervention“ in der Verbindung mit dem CONSORT Statement (Boutron et al., 2015) und den QUEST-Kriterien (Kip und Dirnagl, 2018) erarbeitet. In der Vorlesung zur deskriptiven Statistik wurde der Gummibären Datensatz als Rohdaten erstellt und mögliche Probleme beim Einlesen der Daten im Anschluss im Plenum besprochen. Die Studierenden mussten für die Übung deskriptive Maßzahlen aus den Daten berechnen. Diese Berechnungen waren ohne eine vorherige Bereinigung der Daten nicht möglich (siehe dazu auch Abbildung 3). In dem folgenden Block *Statistisches Testen I & II* wurde als Grundlage die mit dem ig Nobel Prize prämierte Arbeit von Cadiergues et al. (2000), „A Comparison of Jump Performances of

the Dog Flea, *Ctenocephalides canis* (Curtis, 1826) and the Cat Flea, *Ctenocephalides felis felis* (Bouché, 1835)“, genutzt.

Die Studierenden haben das Experiment für die Sprungweite und Höhe visualisieren und die statistischen Tests nachgerechnet. Der Endpunkt für die Sprungweite war in Zentimetern angegeben, der Endpunkt für die Sprunghöhe in Prozent. Überdies ergab sich bei der Sprunghöhe ein Stufendiagramm, welches aus der Überlebenszeitanalyse bekannt war. Das experimentelle Design musste daher im ersten Schritt verstanden werden, bevor weitere Analysen durchgeführt werden können. Wir legten hier großen Wert auf die schematische Visualisierung. Die schematische Visualisierung führte zu einem besseren eigenen Verständnis des Experiments. Dieses Verständnis ist im späteren Berufsleben wichtig, um die beratende Tätigkeit in statistischen Fragen auszuführen. Beiden Seiten wurde noch einmal klar, was die Problemstellung ist (Segel und Heer, 2010). In dem Aufsatz von Cadiergues et al. (2000) werden der *t*-Test und der Chi-Quadrat-Test auf die Daten angewendet. Beide statistische Tests werden in der Vorlesung vorgestellt und die Studierenden müssen die Voraussetzungen an die Tests, wie normalverteilte Daten und Aussagen über die Varianzen, in der Arbeit von Cadiergues et al. (2000) wiederfinden. Können die Studierenden die Ergebnisse von Cadiergues et al. reproduzieren, ist auch auf die Reproduzierbarkeit der Forschung einzugehen.

### **3.2 Berufsbegleitender Masterstudiengang 1. Semester Biomedical Sciences an der fhg – Zentrum für Gesundheitsberufe Tirol**

An drei voneinander getrennten Tagen wurden Teile der Konzepte für Studierende im ersten Semester des berufsbegleitenden Masterstudiengangs Biomedical Sciences an der fhg – Zentrum für Gesundheitsberufe Tirol angewendet. Die Studierenden hatten bereits eine Berufsausbildung im Gesundheitswesen abgeschlossen und bildeten sich über den Masterstudiengang auf eine Promotion vor oder aber für einen beruflichen Aufstieg weiter. Den Kurs besuchten 20 Studierende, wobei der Frauenanteil bei 90 % lag. Die Teilnahme war verpflichtend. Die Veranstaltung wurde abschließend evaluiert. Folgende Konzepte wurden angewandt: i) Erstellung und Auswertung des Gummibären Datensatz, ii) Lesen und Nachrechnen von ig Nobel Prize Veröffentlichungen und iii) Lernen von statistischem Denken anhand des Skorbut-Beispiels. Teile aus der Vorlesung *Statistik für Biowissenschaftler I* wurden verwendet, allerdings wurde die Stoffmenge reduziert und auf die eintägigen Seminare angepasst. So war es möglich, die Inhalte in einem gestrafften Raster aus Vorlesung und Übung (alle 30 Minuten eine Übung von ca. 15 Minuten) zu vermitteln. Teile der Übungen bestanden auch aus praktischen Aufgaben des Programmierens in R. Am Ende der Veranstaltung wurde der Lernerfolg mit einer Klausur überprüft.

### **3.3 Zweitägiger Workshop PerspektivForum Grand Challenges der Stiftung Jugend forscht e.V.**

Abschließend wurden die Konzepte im Rahmen des zweitägigen Workshops „Cargo cult science! Why most published research findings are false“ des PerspektivForums *Grand Challenges* der Stiftung Jugend forscht e.V. auf dem World Health Summit 2018 in Berlin ange-

boten. Es wurden zehn Studierende, die an *Jugend forscht* teilgenommen hatten, eingeladen, diesen Workshop zu besuchen. Die Studierenden verschiedener Fachrichtungen, von Wirtschaftswissenschaften bis Medizin und Ökologie, mussten sich vorab aus sechs Workshops einen auswählen. Eine Evaluierung fand nicht statt. Im Rahmen des PerspektivForums konnte die Idee der Konkurrenz von Robert Koch und Max von Pettenkofer als Theaterstück ausgebaut und dargestellt werden. Das Theaterstück wurde abschließend auf dem World Health Summit 2018 in Berlin aufgeführt. Die geringe Studierendenzahl ermöglichte es, sich ganz auf das Theaterstück als zentrales Element zu konzentrieren und darüber hinaus alle weiteren Aspekte im Prozess zu entwickeln.

## 4 Diskussion und Lehrerfahrungen

### 4.1 Feedback der Studierenden

Die vorgestellten Methoden wurden im Bachelorstudiengang Bioinformatik an der Freien Universität Berlin (FU Berlin) in der Vorlesung *Statistik für Biowissenschaften I* für das 3. Semester evaluiert. Die Veranstaltung wird von 80 bis 100 Studierenden besucht. An der Evaluierung haben sich 35 Studierende beteiligt, die zu diesem Zeitpunkt die Vorlesung besucht haben. Das Leistungsniveau und das Vorwissen der Studierenden waren sehr heterogen. Die Vorlesung wird auch von Studierenden des Masterstudiengangs Statistik der Humboldt-Universität zu Berlin sowie von Informatiker/-innen der FU Berlin besucht. Der Studiengang Bioinformatik hat keinen NC und somit ergibt sich ein starkes Leistungsgefälle in beide Richtungen. Dieses Gefälle wird noch dadurch verstärkt, dass im 1. Semester keine benoteten Prüfungen stattfinden. Da die Veranstaltung im 3. Semester des Bachelorstudiengangs stattfindet, wurde also vorher noch nicht stark nach Leistung selektiert.

Für den B.Sc. Bioinformatik wurde die Evaluation in der Mitte des Semesters zentral durch die Freie Universität Berlin mit standardisierten Formularen durchgeführt. Die Formulare wurden in der Veranstaltung ausgeteilt und anschließend in schriftlicher Form an die FU Berlin zurückgesendet. Die Evaluierung des M.Sc. Biomedical Sciences wurde zentral von der fhg Tirol mit EvaSys am Ende der Veranstaltung über ein digitales Formular durchgeführt. In beiden Fällen war es den Autoren als externe Dozenten nicht möglich, in die Prozesse involviert zu werden. Der Fokus der Evaluierung lag auf Zufriedenheit und Lernergebnissen. Die Studierenden hatten die Möglichkeit, Freitextantworten zu geben. Der Workshop auf dem PerspektivForum wurde nicht systematisch evaluiert. Unsere hier vorgestellten Konzepte lassen nicht sinnvoll mit einem standardisierten Formular evaluieren. Unser Lernziel geht weit über die eigentliche Veranstaltung hinaus und mag erst in zukünftigen Projekten den Studierenden bewusstwerden. Im Idealfall erfahren die Studierende gar keine Sprachlosigkeit in den zukünftigen Projekten.

Abbildung 4 und Tabelle 3 zeigen Auszüge aus der Evaluation der beiden Lehrveranstaltungen für den B.Sc. Bioinformatik (Abb. 4.A) und den berufsbegleitenden M.Sc. Biomedical Sciences (Abb. 4.B). Abbildung 4 greift jeweils die in der Evaluation vorkommende Frage nach der Anwendbarkeit im späteren Studium bzw. Berufsweg auf. Man beachte die gedrehte Bewertungsskala. Die Bewertung der Bachelorstudierenden ist eher uniform, wo-

hingegen sich die Masterstudierenden sehr im positiven Bereich konzentrieren. Die Unterschiede können nochmal aus der Tabelle 3 extrahiert werden, in der die Freitextantworten gesammelt sind. Die Masterstudierenden haben neben technischen Unzulänglichkeiten und einer zu kurzen Kursdauer nur wenige kritische Aspekte genannt und waren sehr von der Art des Seminars begeistert. Beide Gruppen schrieben am Ende der Veranstaltung eine Klausur. Die Bachelorstudierenden waren jedoch sehr viel mehr auf die Klausur fixiert und empfanden die Vorlesung sowie die Übung teilweise als zu unstrukturiert und nicht klar genug auf mögliche Klausurfragen fokussiert. Die Bachelorstudierenden zeigten somit eine starke Fokussierung auf die Klausur und nahmen die Methoden kritischer auf, als die Masterstudierenden mit dem berufsbezogenen Hintergrund (Tabelle 3).

Tabelle 3: Schriftliche Evaluation des Moduls *Statistik für Biowissenschaftler I* an der FU Berlin in Kooperation mit der Charité – Universitätsmedizin Berlin sowie *Medizinische Biometrie und Statistik* an der fhg – Zentrum für Gesundheitsberufe Tirol.

Was hat Ihnen gut gefallen?	Was könnte verbessert werden?
Doch noch mehr Fallbeispiele und noch mehr auf Gummibären Daten zurückkommen (B.Sc.)	Sehr unterschiedliches Niveau zwischen Vorlesungen und zwischen Vorlesungen und Übungen (B.Sc.)
Viele Bezüge zu wissenschaftlichen Partnerschaften (M.Sc.)	Einige Begriffe, Formeln etc. nicht in Vorlesung besprochen, aber in Übungen gefragt (B.Sc.)
Bezug zu realen Anwendungen (M.Sc.)	Viele vage/nicht eindeutige Fragestellungen in der Übung (B.Sc.)
Humor und Anwendungsbezug (M.Sc.)	Nicht sehr informatives Feedback zu Fragen zur Klausurrelevanz (B.Sc.)
Die Art und Weise, wie die Vorlesung gestaltet wird (M.Sc.)	Bezug zu Vorlesungsfolien teilweise sehr unterschiedlich (B.Sc.)
Mir helfen persönlich Beispiele mehr als mathematische Definitionen. (B.Sc.)	Kaum können Fragen zur Klausurrelevanz beantwortet werden. (B.Sc.)
Aktives Umsetzen des Gelernten (M.Sc.)	Vor allem in den ersten Vorlesungen unsympathisch, viel heißer Brei, bevor es zu den Fakten kommt. Was ist wichtig, was ist unwichtig? (B.Sc.)
Praktisches Arbeiten (M.Sc.)	Manchmal zu viel „Storytelling“ über Mediziner. Mehr Zeit für mathematische Prinzipien verwenden. (B.Sc.)
Viel selbst erarbeiten (B.Sc.)	Die Aufgaben der Übung waren teilweise nicht nur mit der Vorlesung lösbar und das Modul war zu chaotisch geregelt. (B.Sc.)

## 4.2 Diskussion der Konzepte

Unsere Evaluation kann nicht den strengen Anforderungen einer evidenzbasierten Analyse folgen. Wir haben deshalb bewusst auf eine evidenzbasierten Analyse verzichtet und bilden hier die Meinungen der Studierenden sowie einen für uns passenden Auszug aus den

beiden Evaluierungen wider (Abbildung 4 und Tabelle 3). Wie Decker und Mucha (2018) herausarbeiten, kann forschungsbasiertes Lernen bei einigen Studierenden Gefühle der Unsicherheit und Leistungsängstlichkeit auslösen. Dies würde sich in einer schlechten Evaluation mit negativen Ergebnissen für den Dozenten spiegeln. Da wir diese Faktoren nicht erfassen können, würde dies die Auswertung verzerren. Studierende mit einer niedrigen Ungewissheitstoleranz würden die Veranstaltung als weniger sinnvoll empfinden und schlechter bewerten. Wir sehen in den angebotenen Konzepten aber die Chance einer langfristigen Wirkung, die sich vielleicht erst in den folgenden Semestern oder bei einer Abschlussarbeit entfaltet.

Einigen Bachelorstudierenden war nicht bei allen Fragestellungen klar, was als Lösung der gestellten Aufgaben erwartet wurde. Auch die Aussage, dass dies in der echten Forschung eben auch häufig das Problem sei und eine Statistikerin die Methoden und Möglichkeiten abwägen müsse, wurde im Kontext einer zu bestehenden Klausur als unbefriedigend empfunden. Der Mangel, dass Aufgaben nicht eindeutig formuliert wurden, wird von den Studierenden des Bachelorstudiengangs kritisch, von jenen des Masters im Gegensatz als anregend angesehen. Durch die Evaluationsergebnisse konnten neue Erkenntnisse zur Verbesserung des Lehrkonzepts gewonnen werden. Ein Aspekt ist die Forderung nach vermehrter Arbeit mit Beispielen in den Aufgaben, um die konzeptionelle Idee und die Lernziele besser zu kommunizieren. Ferner mussten die Übungsleiter/-innen umgeschult werden, um an die neue Ausrichtung der Aufgabenstellung anzuknüpfen. Diese waren zuvor rein mathematisch zu lösen. Hier mussten auch Widerstände überwunden werden und nicht jede Übungsleiterin konnte anfänglich vollständig überzeugt werden. Wir erhoffen uns in den nächsten Semestern, die Übungsleiter/-innen besser weiterzubilden, sodass die hier vorgestellten Konzepte noch besser umgesetzt werden können.

Die Übernahme von Aspekten der hier vorgestellten Konzepte in die zu schreibende Klausur für Bachelorstudierende gestaltet sich schwierig, da bis zu 120 Personen an einer schriftlichen Klausur teilnehmen und diese mit sehr begrenzten Personal zeitnah korrigiert werden muss. Die Klausur als Prüfungsart ist vom Modulhandbuch vorgegeben. Die Masterstudierenden erhielten eine online zu absolvierende Klausur. Hier war es begrenzt möglich, entsprechende Fragen zu stellen, die mehr auf die hier vorgestellten Konzepte eingingen. Die Anzahl der Studierenden im Master war klein genug, um längere Texte zu korrigieren.

Obwohl die Darstellung der kritischen Auseinandersetzung in Form eines Theaterstücks der Diskussion von Robert Koch und Max von Pettenkofer bei den Studierenden des PerspektivForums positive Rückmeldung erhielt, wurde auf eine Darstellung des Workshops durch die Pressestelle von *Jugend forscht* in Bild und Text verzichtet. Es kann nur vermutet werden, dass die Form und der Aufbau des Workshops nicht zu den anderen angebotenen Workshops auf dem PerspektivForum passte und aus diesem Grund nicht in der Außendarstellung präsentiert wurde.

In den folgenden Semestern wollen wir die Richtlinien aus dem GAISE Report der American Statistical Association (ASA) weiter anwenden. Insbesondere ist die ASA eine gute Quelle für aktuelle statistische (Lehr-)Entwicklungen, wie dem problematischen Nutzen des  $p$ -Wertes und dessen Diskussion (Wasserstein und Lazar, 2016). Wir hoffen, durch un-

sere erste Ausbildung im statistischen Denken die Grundlage dafür gelegt zu haben, konzeptionell fragwürdige Entwicklungen in der Datenanalyse in folgenden Semestern besser besprechen zu können. Weitere statistische Lehrveranstaltungen sollen mehr auf das konzeptionelle Denken fokussiert werden. Delmas et al. (2007) schlägt hierfür den CAOS (engl. Comprehensive Assessment of Outcomes in Statistics) Test vor, welcher in die Überlegungen des Konzepts zum folgenden Curriculum einbezogen wird.

### 4.3 Tipps für den Einsatz in der eigenen Lehre und Limitationen

Die Anzahl der Studierenden ist ein stark limitierender Faktor. Bei der Erstellung des Gummibärendatensatzes ist es von Vorteil, wenn die Gruppe zuvor gut organisiert wurde, da viele Datenpunkte anfallen. Die Studierenden waren sehr aufgeregt, als sie die Gummibärentüten ausgepackt sahen. Sie konnten ahnen, dass der Veranstaltungsablauf nicht dem Üblichen entsprechen werde. Es ist daher wichtig, die einzelnen Schritte in Ruhe mehrfach zu erklären. Die Fragen, die sich hierbei stellen, sind: Was sind die folgenden einzelnen Schritte? Wie funktioniert Google Spreadsheet? Welche Stuhlreihe gibt in welcher Reihenfolge die Daten ein? Wann erhält man die Daten und vor allem eine sehr wichtige Information: Wann dürfen die Studierenden die Gummibären essen? Funktionierendes WLAN ist für die Erstellung der Gummibärendatensätze unbedingt notwendig. In Einrichtungen, in denen EDUROAM zur Verfügung steht, ist meist auch die Bandbreite groß genug, um Google Spreadsheet starten und auch Daten eintragen zu können. Bei einer kleineren Gruppe von etwa 30 Studierenden können darüber hinaus alle Daten zentral an einem Rechner eingegeben werden. Dies ist mit ein wenig Organisation – ein/-e Studierende/-r tippt, eine/-r diktiert – ohne viel Aufwand umsetzbar. Um die Limitation zu umgehen, stellen wir den Gummibärendatensatz in den zwei Versionen (original und gereinigt) zur Verfügung (per E-Mail-Anfrage an den Erstautor).

Am Skorbut-Beispiel und dem Streitgespräch zwischen Koch und von Pettenkofer zeigte sich die Limitierung durch die Gruppengröße. In einer Kleingruppe von ca. acht bis 10 Studierenden lässt sich das Theaterstück mit der notwendigen Zeit von ca. einem Tag (8 UE) hervorragend erarbeiten. Dies ist natürlich mit einer sehr viel größeren Gruppe mit weniger Zeit kaum möglich. Doch lassen sich hier Punkte verschriftlichen und in kleineren Gruppen diskutieren. Wir werden versuchen, in einer größeren Veranstaltung mehrere Gruppen zu bilden. Jede Gruppe erhält dann den Auftrag, das Stück zu erarbeiten. Eine Erprobung des Workshops ist bereits geplant.

Einige Bachelorstudierende waren nicht einverstanden mit unserer Definition des Berufsbildes der Statistikerin bzw. hatten sich mit dem Berufsbild kaum auseinandergesetzt. Dies führt bedingt zu einem Konflikt in den Konzepten, da die Teilhabe an der Vorlesung und der Übung von manchen Studierenden als Hürde gesehen wird (Zitat Studierende: „viel heißer Brei“). Hier wirkte auch limitierend, dass sich der Großteil der Bachelorstudierenden, anders als die Masterstudierenden, zwischen den Veranstaltungen nicht mehr mit dem Stoff beschäftigten und sich dies erst zum Ende der Vorlesungszeit, angesichts der nahenden Klausur, einstellte. Wir versuchen, dem mit einer besseren Vorlesungs- und Übungsstruktur entgegenzuwirken, die eine aktive Teilnahme in den Übungen vorsieht.

Wir haben uns in den Beispielen auf das weite Feld der Humanwissenschaften konzentriert. Die Konzepte lassen sich aber leicht auch auf andere Felder erweitern, wenn die

Beispiele als Inspiration dienen. Insbesondere der ig Nobel Prize lässt sich auf *alle* Bereiche der Wissenschaft anwenden. Hier bedarf es nur ein wenig mehr Recherche nach der passenden Veröffentlichung. Soll es in der Veranstaltung um die Erfahrung von wissenschaftlichen Denkmustern gehen, wie zum Beispiel eine Veranstaltung zu guter, wissenschaftlicher Praxis, so lassen sich die Konzepte auch mit allen Studierenden aller Fachrichtungen anwenden. Meistens ist es sogar spannender, mit fachfremden aber leichter zugänglichem Material zu arbeiten. Der Fokus kann so leicht wieder auf die wissenschaftliche Metaebene geholt werden und verliert sich nicht im Detail. Die Theaterstücke lassen sich auf jedes wissenschaftliche Feld erweitern. So bieten sich im wirtschaftswissenschaftlichen Bereich sicherlich ein Streitgespräch zwischen den Ideen des österreichischen Wirtschaftswissenschaftler Friedrich August von Hayek und seines britischen Kollegen John Maynard Keynes an. Beide haben die Volkswirtschaftslehre ein ganzes Jahrhundert lang geprägt.

Um die hier vorgestellten Konzepte auch nachhaltig an andere Lehrende der Statistik in verschiedenen Hochschulen weiterzugeben, werden die Inhalte in einem eintägigen Workshop am Berliner Zentrum für Hochschullehre von den Autoren weitergegeben. Somit ist sichergestellt, dass die Erkenntnisse aus diesem Artikel vertieft weitergegeben werden können.

## 5 Fazit

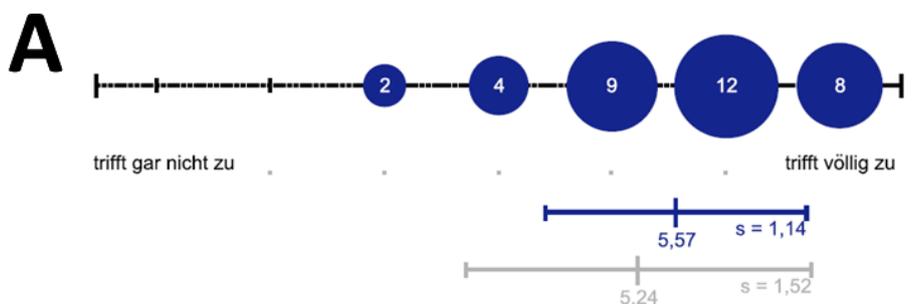
Was ist gute Hochschullehre? Ulrich (2016, Seite 16) benennt „das Ziel guter Hochschullehre besteht letztendlich in der Ausbildung mündiger, kompetenter und wertgefestigter (Staats-)Bürger“. Wir versuchten in diesem Artikel einen kleinen Beitrag zu diesem humboldtschen Bildungsideal zu erbringen. Durch die Reduzierung der Sprachlosigkeit zwischen den wissenschaftlichen Disziplinen wollen wir eine Unterstützung zur besseren Forschung leisten. Forschungsbasiertes Lernen und Storytelling bietet hier die Möglichkeit, langfristige Lernerfolge zu erzielen und Selbstbewusstsein in ihr eigenes Forschen in zukünftigen Wissenschaftler/-innen zu wecken. Daher werden in unserer Arbeit die Inhalte zum Thema Statistik aus der Isolation des Curriculums in die Gesellschaft komplexer arbeitsweltnaher Geschichten gebracht. Das Ergebnis ist eine soziale Reaktionskette durch Irritation und Erkenntnis bei den Studierenden. Sie sind in der Lehrveranstaltung bereits in der komplexen Arbeitswelt, auf die wir sie vorbereiten wollen. Das ist unsere Aufgabe und Verantwortung.

## Anhang

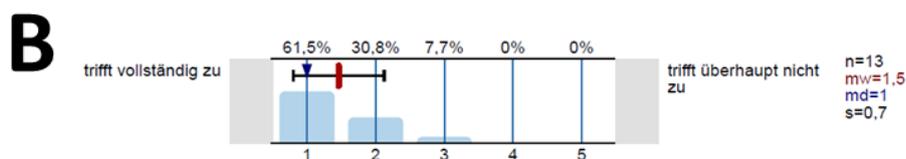
### Gummibären Datensatz

Im Folgenden findet sich der Anhang mit den Abbildungen zu der Datenerhebung mit Google Spreadsheet. Die/der Lehrende muss hierfür einen Account bei Google besitzen, die Studierenden können anonym ohne Anmeldung die Daten durch einen vorher veröffentlichten Link eintragen. Es empfiehlt sich, über einen Kurz-URL-Dienst wie bitly den Link zu verkürzen und so die Eingabe zu erleichtern. Google Spreadsheet lässt sich über jedes Smartphone auch ohne die passende App aufrufen und bedienen. Der Datensatz ist im Original sowie in der gereinigten Version vom Erstautor Jochen Kruppa ([jochen.kruppa@charite.de](mailto:jochen.kruppa@charite.de)) zu erhalten.

### Auszug aus der Evaluation



...verdeutlicht den Anwendungsbezug der Lerninhalte/des Stoffs.



Die/der Lehrende macht Bezüge zwischen dem Thema der Lehrveranstaltung und der beruflichen Praxis deutlich.

Abbildung 4: Ausschnitt aus der Evaluation von 35 Studierenden des Moduls *Statistik für Biowissenschaftler I* an der FU Berlin in Kooperation mit der Charité – Universitätsmedizin Berlin (A) und 20 Studierenden des Moduls *Medizinische Biometrie und Statistik* an der fhg – Zentrum für Gesundheitsberufe Tirol (B). Ausgewählt wurde die passende Frage zum Anwendungsbezug. Die gedrehten Bewertungsskalen sind zu berücksichtigen.

## Literatur

- Abrahams, M. (2002). *Ig Nobel Prizes: The Annals of Improbable Research*. Orion. Online unter <https://www.improbable.com/ig/winners/> [06.08.2019]
- Blume, D., Brennecke, C., Brey Mayer, U. & Eisentraut, T. (Hrsg.) (2018). *Europa und das Meer. Katalog zur Ausstellung im Deutschen Historischen Museum*. München: Hirmer Verlag.
- Boutron, I., Altman, D. G., Moher, D., Schulz, K. F. & Ravaud, P. (2017). CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Annals of internal medicine*, 167(1), 40–47.
- Cadiergues, M. C., Joubert, C. & Franc, M. (2000). A comparison of jump performances of the dog flea, *Ctenocephalides canis* (Curtis, 1826) and the cat flea, *Ctenocephalides felis felis* (Bouché, 1835). *Veterinary parasitology*, 92(3), 239–241.
- Carver, R., Everson, M., Gabrosek, J. et al. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2016*. Alexandria: American Statistical Association.
- Decker, C. & Mucha, A. (2018). Forschendes Lernen lernen. Zu den didaktischen und emotionalen Herausforderungen der Integration von Lernen über, für und durch Forschung. *die hochschullehre*, 4, 143–160.
- Delmas, G., Joan, G., Ooms, A. & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28–58
- Feynman, R. P. (1974). Cargo cult science. *Engineering and Science*, 37(7), 10–13.
- Goethe, J. W. (1808). *Faust: Der Tragödie Erster Teil, Nacht*. Tübingen.
- Hansson, L., Arvidsson, Å., Heering, P. & Pendrill, A. M. (2019). Rutherford visits middle school: a case study on how teachers direct attention to the nature of science through a storytelling approach. *Physics Education*, 54(4), 1–11.
- Huber, L. (1970). Forschendes Lernen: Bericht und Diskussion über ein hochschuldidaktisches Prinzip. *Neue Sammlung*, 10(3), 227–244.
- Huber, L. (2014). Forschungsbasiertes, Forschungsorientiertes, Forschendes Lernen: Alles dasselbe? Ein Plädoyer für eine Verständigung über Begriffe und Unterscheidungen im Feld forschungsnahen Lehrens und Lernens. *Das Hochschulwesen*, 62(1), 2.
- Johnson, S. & Taylor, K. (Hrsg.). (2011). *The Neuroscience of Adult Learning: New Directions for Adult and Continuing Education, Number 110* (Vol. 81). San Francisco: John Wiley & Sons.
- Kip, M. & Dirnagl, U. (2018). *Mechanisms of robust, innovative and translational research (MERIT) – QUEST criteria*. Berlin: QUESTCenter, BIH-Berlin. Online unter <https://osf.io/n4hcm/> [06.08.2019]
- Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69–71.
- Parkin, L., Williams, S. M. & Priest, P. (2009). Preventing winter falls: a randomised controlled trial of a novel intervention. *New Zealand Medical Journal*, 122(1298), 31–38.

- Pöpel, N. & Morisse, K. (2019). Inverted Classroom: Wer profitiert – wer verliert? Die Rolle der Selbstregulationskompetenzen beim Lernen im umgedrehten MINT-Klassenraum. *die hochschullehre*, 5, 55–74.
- Pettenkofer, M. (1858). *Über den Luftwechsel in Wohngebäuden*. München: Cottaesche Buchhandlung.
- Raschke, G. (2008). *Die Cholera Theorie Max von Pettenkofers im Kreuzfeuer der Kritik*. Die Cholera Diskussion und ihre Teilnehmer. Dissertation, Technische Universität München.
- Schimmel, J. (2012). *Writing science: How to write papers that get cited and proposals that get funded*. New York: OUP USA.
- Segel, E. & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics*, 16(6), 1139–1148.
- Susser, M. (1986). The logic of Sir Karl Popper and the practice of epidemiology. *American Journal of Epidemiology*, 124(5), 711–718.
- Ulrich, I. (2016). *Gute Lehre in der Hochschule. Praxistipps zur Planung und Gestaltung von Lehrveranstaltungen*. Wiesbaden: Springer.
- Unger, D. L. (1998). Does knuckle cracking lead to arthritis of the fingers? *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 41(5), 949–950.
- Wasserstein, R. L. & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129–133.
- Wild, C. J. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International statistical review*, 67(3), 223–248.
- Wolfe, P. (2006). The role of meaning and emotion in learning. In S. Johnson & K. Taylor (Hrsg.), *The Neuroscience of Adult Learning: New Directions for Adult and Continuing Education* (Number 110, 81, S. 35–43). San Francisco: John Wiley & Sons.
- Wood, B. L., Mocko, M., Everson, M. et al. (2018). Updated guidelines, updated curriculum: The GAISE College Report and introductory statistics for the modern student. *CHANCE*, 31(2), 53–59.
- Xie, Y., Allaire, J. J. & Grolemond, G. (2018). *R markdown: The definitive guide*. Boca Raton: CRC Press.
- Zazkis, R. & Liljedahl, P. (2019). *Teaching mathematics as storytelling*. Rotterdam: Brill Sense.

## Autoren

Dr. Jochen Kruppa, Charité – Universitätsmedizin Berlin, Institut für Biometrie und Klinische Epidemiologie, Berlin, Deutschland; Email: <mailto:autorin@universität.de> [jochen.kruppa@charite.de](mailto:jochen.kruppa@charite.de)

Dr. Björn Kiehne, Technische Universität Berlin, Berliner Zentrum für Hochschullehre, Berlin, Deutschland; Email: [Bjoern.Kiehne@tu-berlin.de](mailto:Bjoern.Kiehne@tu-berlin.de)



**Zitiervorschlag:** Kruppa, J. & Kiehne, B. (2019). Statistik lebendig lehren durch Storytelling und forschungsbasiertes Lernen. *die hochschullehre*, 5, 501–524. Online unter [www.hochschullehre.org](http://www.hochschullehre.org)