

Zusammenfassung: *Sich unabhängig und selbstbestimmt eine Meinung bilden zu können ist eine Grundvoraussetzung für mündige Bürger. Im Dickicht des heutigen Informationsdschungels ist das ohne statistische Kenntnisse und eine kritische Einstellung jedoch nur noch bedingt möglich. Im Folgenden wird das Modul Kritisches Hinterfragen datenbasierter Aussagen vorgestellt, das in Ludwigsburg mehrfach in Seminaren mit Lehramtsstudierenden im Fach Mathematik erprobt wurde. Durch eine schrittweise durchgeführte Analyse einer datenbezogenen Schlagzeile, des dazugehörigen Artikels und dessen Datengrundlage soll das Bewerten von Aussagen, die auf einer Datenanalyse basieren, behandelt werden. Dabei werden die wesentlichen Bereiche herausgearbeitet, die für eine evidenzbasierte Bewertung wichtig sind.*

1 Theoretischer Hintergrund

Mündigkeit als Charakteristikum von Bürgern einzufordern ist nicht neu. Bereits Kant forderte von Bürgern sich ihres eigenen Verstandes zu bedienen, um der Unmündigkeit zu entkommen (vgl. Kant 1784). Adorno beschreibt Mündigkeit in Anlehnung an Kant, erweitert um kritische, emanzipatorische oder autonome Facetten und bezeichnet Mündigkeit als wichtige Voraussetzung einer Demokratie (vgl. Adorno 2015). Auch nach Frech & Richter (2013) ist eine demokratische Gesellschaft auf mündige Bürger angewiesen. Aktuelle Sichtweisen beziehen zudem statistische Fähigkeiten für die Informationsverarbeitung mit ein. Gramm (2010, S. 29) beschreibt Mündigkeit als „Orientierungsvermögen in der Unübersichtlichkeit des modernen Informationsdschungels“. Diese befähigt den Bürger, sich die erforderlichen „Informationen zu beschaffen, Fragen zu stellen, zu verstehen“ (Fischer 2012), um eine reflektierte Entscheidung zu treffen. Es gilt also, datenbasierte Argumente und deren Darstellungen zu verstehen, sie zu interpretieren und mögliche Schlussfolgerungen zu hinterfragen (vgl. Engel 2014; Ridgway 2015). Dazu wird aber nicht nur *klassisches* statistisches Wissen gebraucht, wie es in Standardkursen zur Einführung in die Statistik gelehrt wird, sondern beispielsweise auch Wissen um die Datenerhebung, die Operationalisierung der Variablen und Kontextwissen. Darüber hinaus spielen neben Wissenselementen kritisches Reflektieren sowie Überzeugungen und Einstellungen eine wichtige Rolle (vgl. Gal 2002).

In diesem Artikel bedeutet der Begriff *Information* eine neutrale, objektive Tatsachendarstellung ohne Einschätzung oder Bewertung. Im Gegensatz dazu, ist eine *Meinung* eine Äußerung zu einem Sachverhalt welche eine persönliche Bewertung enthält. Schwierigkeiten bereiten Meinungen in Form von datenbasierten Aussagen, beziehungsweise bewertete statistische Informationen. Denn nicht immer ist ersichtlich, dass die statistischen Informationen bereits gedeutet wurden oder wie viel Deutungsspielraum die statistischen Informationen geben. Ein Beispiel für eine statistische Information ist: „Der Durchschnitts-BMI von 2016 ist 24,75; der von 1975 ist 22,08“. Eine datenbasierte Aussage könnte so lauten: „Der Durchschnitts-BMI von 2016 ist deutlich größer als von 1975, da er in dieser kurzen Zeitspanne immerhin von 22,08 auf 24,75 gestiegen ist.“ Dass der Unterschied deutlich ist oder die knapp 40 Jahre eine kurze Zeitspanne bedeuten, sind bereits Deutungen.

Der folgende Beitrag stellt das Modul *Kritisches Hinterfragen datenbasierter Aussagen* (abgekürzt: *CQS*, *critical questioning of data-based statements*) vor, das in Ludwigsburg mehrfach in Seminaren mit angehenden Mathematiklehrer*innen eingesetzt wurde. Das Modul orientiert sich an einem im Folgenden dargestellten Schema (Abb. 1), lebt aber zugleich von der Aktualität der Themen sowie der Verfügbarkeit dazu gehöriger Daten. Hinweise zur Umsetzung des Moduls in der Schule mit Bezug zum schulischen Bildungsauftrag schließen den Beitrag ab.

2 Das CQS-Modul

Das CQS-Modul setzt sich zusammen aus realen und motivierenden journalistischen Produkten (Schlagzeilen und Artikel) und Datensätzen, sowie Aktivitäten für die Entwicklung von statistischem Argumentieren mit geeigneten technischen Hilfsmitteln. Es legt dabei den Fokus auf kritisches Reflektieren gesellschaftlich relevanter statistischer Informationen inkl. Kontextwissen für eine Bewertung datenbasierter Aussagen aus den Medien. Ziele des Moduls sind:

- (1) die Erarbeitung von Kriterien für eine Bewertung von datenbasierten Aussagen,
- (2) eine Verbesserung der Einschätzung der Aussagekraft von statistischen Kennzahlen und datenbasierten Aussagen innerhalb der Medien und
- (3) das Unterscheiden lernen zwischen einer statistischen Information und einer Meinung, welche mit statistischen Informationen unterlegt ist.

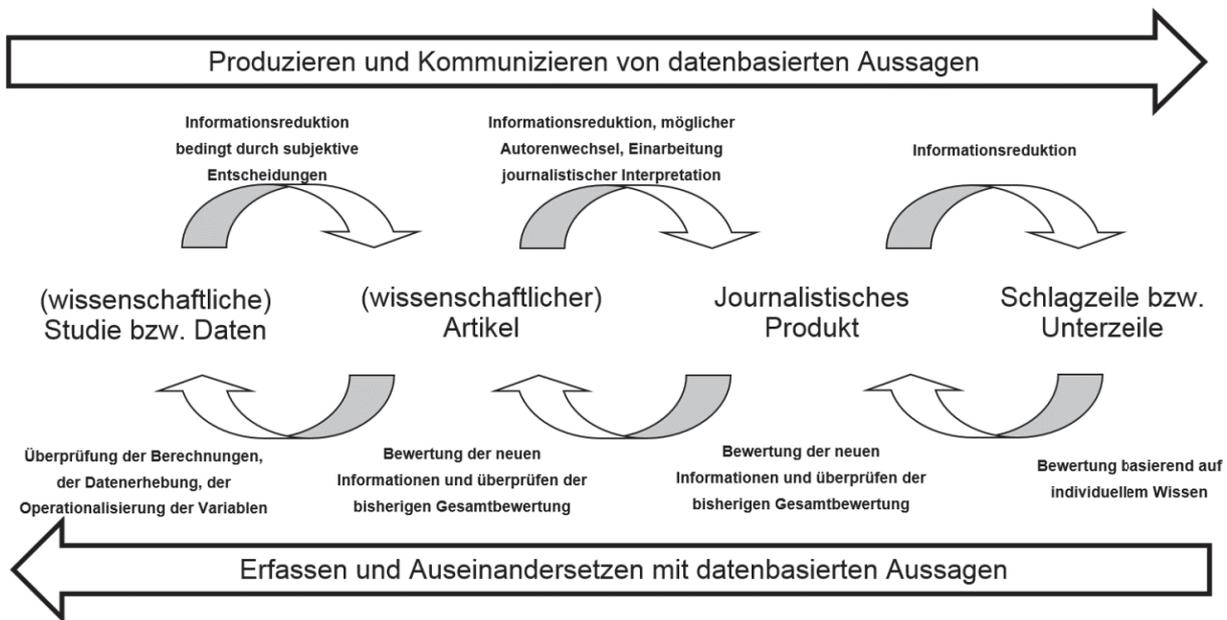


Abb. 1: Rahmen zum kritischen Hinterfragen datenbasierter Aussagen

Das Modul besteht aus zwei Teilen, dem Erfassen und Auseinandersetzen mit datenbasierten Aussagen, sowie dem Produzieren und Kommunizieren datenbasierter Aussagen. Die Grafik (siehe Abb. 1) stellt die relevanten Schritte dar.

Die obere Hälfte der Abbildung zeigt die Schritte, wie aus datenbasierten Informationen einer (wissenschaftlichen) Studie eine Überschrift entsteht, mit der der Leser generell zuerst auf ein Thema aufmerksam wird. Sie verdeutlicht an welchen Stellen bei der Produktion und Kommunikation datenbasierter Aussagen subjektive Entscheidungen einen Einfluss auf die Informationsübermittlung haben. Denn es ist unrealistisch, die Gesamtheit der in wissenschaftlichen Daten bzw. Studien enthaltenen Informationen einer breiten Öffentlichkeit zu präsentieren. Deswegen werden Informationen aus Gründen der Verständlichkeit und der Prägnanz verdichtet oder ganz weggelassen.

Prinzipiell kommt es bei jedem Schritt zu einer Reduzierung der ursprünglichen Informationen, die häufig durch subjektive Entscheidungen beeinflusst werden. Vom (wissenschaftlichen) Artikel zum journalistischen Produkt werden in der Regel vorhandene Informationen gestrichen, fehlende Informationen (jeweils subjektive Entscheidungen) und oftmals Meinungen hinzugefügt. Die Entscheidung, wie die Schlagzeile lautet, ist wiederum subjektiv und orientiert sich meist an der Klientel des Mediums. Entlang des Weges von der (wissenschaftlichen) Studie zu einer Überschrift eines Artikels orientiert sich die Vorgehensweise des zweiten Teils des Moduls.

Die untere Hälfte zeigt, an welchen Stellen beim Erfassen und Auseinandersetzen mit datenbasierten Aussagen subjektive Bewertungen eine Rolle spielen. Üblicherweise erfolgt der erste Kontakt mit gesellschaftlichen Themen durch Schlagzeilen, gefolgt von einem journalistischen Produkt in Zeitungen, dem Internet oder Nachrichtensendungen. Bei Interesse werden die angegebenen Quellen begutachtet, um weitere Informationen zu bekommen. Daran angelehnt ist die Vorgehensweise des ersten Teils des Moduls. Da im Alltag das Erfassen datenbasierter Aussagen häufiger stattfindet, steht dieser Teil auch im Fokus des CQS-Moduls. Während konkrete Inhalte je nach Aktualität, Interesse und Verfügbarkeit von Daten austauschbar sind, liefert das in Abbildung 1 dargestellte Schema einen Rahmen zum kritischen Hinterfragen datenbasierter Aussagen.

Teil 1 des Moduls: Erfassen und Auseinandersetzen mit datenbasierten Informationen

Dieser Teil des Moduls besteht aus drei Arbeitsschritten, die den normalen Ablauf mit datenbasierten Informationen widerspiegeln. Die Studierenden bekommen im ersten Arbeitsschritt eine Schlagzeile mit Fragestellungen, danach den zur Schlagzeile dazugehörigen Artikel mit weiteren Fragen und zum Schluss zusätzliche Hintergrundinformationen in Form von Datensätzen oder Studienergebnissen, die die Autoren der Artikel als Quelle genannt haben. In den in Ludwigsburg durchgeführten Seminaren haben die Studierenden unter anderem folgende Schlagzeile erhalten: „So dick war die Menschheit noch nie. Erstmals leben mehr Fettleibige als Untergewichtige auf der Welt“⁴¹, anhand derer das CQS-Modul exemplarisch vorgestellt wird.

Die Fragen zur Schlagzeile sollen individuell beantwortet werden, um eigenständige Antworten zu erhalten. Zunächst sollen die Studierenden die Aussagekraft der Schlagzeile bewerten. Zudem sollen sie sich auch Gedanken darüber machen, ob sie die Schlagzeile positiv, neutral oder negativ wahrgenommen haben, um auch einen affektiven Zugang zur Thematik herzustellen. Des Weiteren soll überlegt werden, welche weiteren Informationen nötig sind, um Unklarheiten oder Ungenauigkeiten zu beseitigen. Zudem sollen sich die Studierenden überlegen, ob die Schlagzeile schon genug Informationen beinhaltet, anhand derer Entscheidungen getroffen werden sollten.

Die Antworten wurden im Plenum diskutiert, um den Studierenden die Vielfalt der Bewertungen zu verdeutlichen und alle von ihnen beschriebenen weiteren Informationen zu erfassen. In der Diskussion hielt sich die Lehrkraft zurück und beschränkte sich auf die Moderation sowie auf Verständnisfragen, so dass alle Studierenden einen Beitrag leisteten.

Bei der Aufforderung nach Nennung zusätzlicher Informationen wurden durch die Diskussion im Plenum unter anderem folgende Fragen aufgeworfen:

- Ist mit der Menschheit wirklich die gesamte Menschheit gemeint?
- Was bedeutet fettleibig, was untergewichtig?
- Auf welchen Zeitraum beziehen sich die Wörter *nur* und *erstmal*?
- Wie wurden die Daten erhoben?
- Ist eine generelle Zunahme überhaupt problematisch?
- Sind *nur* die Untergewichtigen weniger geworden?
- Wie ist der Prozentuale Anteil der jeweiligen Gruppe?

Auf die Frage, ob man schon anhand der Informationen der Schlagzeile Schlussfolgerungen ziehen kann, gab es mit einer Ausnahme nur Verneinungen, da der Informationsgehalt der Schlagzeile zu gering ist. Nur ein Studierender äußerte die Schlussfolgerung, dass mehr gegen Übergewicht getan werden müsse, um eine weitere Zunahme zu verhindern.

Im zweiten Arbeitsschritt wird der zur Schlagzeile gehörende Artikel ausgeteilt. Die Studierenden sollen (nun in Partnerarbeit) anhand dieses Zusatzmaterials reflektieren, ob sie ihre erste Bewertung verändern (müssen), und falls ja, inwiefern. Die beeindruckenden Zahlen der Studie (1700 ausgewertete Bevölkerungsstudien mit insgesamt 19,2 Millionen Menschen) sorgten bei einigen Studierenden für eine

teils deutlich bessere Bewertung der Aussagekraft des Artikels. Allerdings gab es auch kritische Anmerkungen, dass man ja gar nicht wisse, wie die Daten erhoben wurden. „Eine solch hohe durchschnittliche Gewichtszunahme über eine solch lange Zeit (von 1975 bis 2014 zehn Jahre ein Plus von 1,5 kg) und die daraus resultierende große Anzahl an Fettleibigen sind erschreckend“ war eine Aussage, die mit einer negativen Konnotation und auf Grund der Informationen einer hohen Aussagekraft einherging. Diejenigen Studierenden, die immer noch bei der Einschätzung einer geringen Aussagekraft blieben, begründeten dies vor allem mit der Verwendung des BMI, „der ja anfällig ist, da er nicht zwischen Muskelmasse und Fett unterscheidet und Muskelmasse ja schwerer ist“ oder „die Einteilung der Grenzen von Untergewicht, Normalgewicht und Fettleibigkeit ist willkürlich und zudem sollte auch das Alter berücksichtigt werden.“

Nach der Lektüre des Artikels konnten zudem einige Fragen geklärt werden: Wie die Begriffe *unter- und übergewichtig* zu verstehen sind; als Bezugspunkt ist wirklich die gesamte Menschheit gemeint; das „noch nie“ bezieht sich auf den Zeitraum von 1975–2016. Zusätzliche Informationen zum BMI, der mit Fettleibigkeit verbundenen Krankheiten, aber auch das immer noch vorhandene Problem von (starkem) Untergewicht in der Welt wurde im Artikel thematisiert. Es haben sich durch die Lektüre des Artikels auch neue Fragen aufgetan:

- Welche Unterschiede gibt es zwischen Männern und Frauen?
- Gibt es Regionen oder Länder die besonders betroffen sind?
- Gibt es Länder bei denen der durchschnittliche BMI nicht (immer) ansteigt?
- Gibt es Alternativen zum BMI, die genauer/besser sind?

Einzelne weitere Studierende veränderten ihre Schlussfolgerungen dahingehend, dass mehr gegen Fettleibigkeit getan werden müsse. Andere wiederum blieben bei ihrer Aussage und argumentierten weiterhin, dass der BMI keine gute/genauere Variable sei und deshalb keine Schlussfolgerungen gezogen werden dürften. Die letzte Aufgabe zum Artikel besteht darin, eine alternative Schlagzeile für den Artikel zu formulieren, welche im zweiten Teil des Moduls verwendet wird.

Für den dritten Arbeitsschritt benötigen die Studierenden einen PC mit Internetzugang, um Online mit verschiedenen Softwaretools und Datensätzen arbeiten zu können. Die zusätzlichen Informationen bestehen aus der Studie *Trends in adult body-mass index*

(Di Cesare et al. 2016) und verschiedenen Datensätzen, die von der Webseite www.ncdrisk.org heruntergeladen oder mit verschiedenen Visualisierungsmöglichkeiten direkt auf der Webseite angeschaut bzw. analysiert werden können. Der downloadbare Datensatz beinhaltet 200 Länder mit den Variablen Region, Geschlecht, Jahr, Durchschnitts-BMI, sowie die jeweiligen Anteile der Bevölkerung aufgeteilt nach den verschiedenen BMI-Stufen, die von der WHO verwendet werden². Auf der Webseite finden sich zudem dynamische Visualisierungstools, die zusätzlich die absoluten Zahlen der jeweiligen BMI-Stufen beinhalten (zum Beispiel Abb. 2).

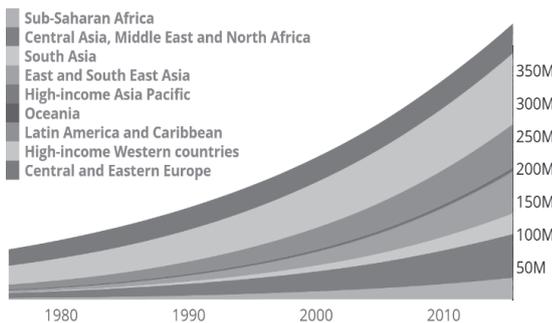
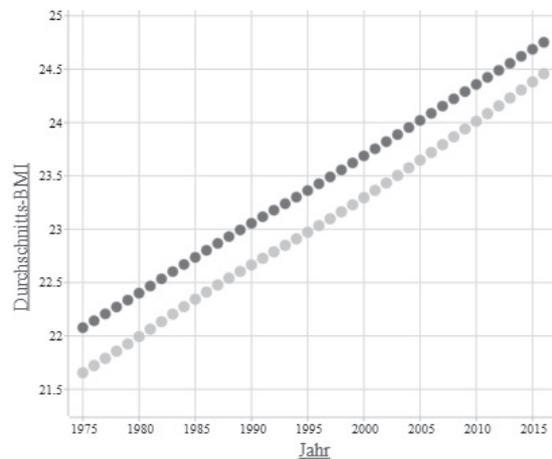


Abb. 2: Anzahl fettleibiger Frauen (BMI > 30) in Millionen weltweit von 1975 bis 2016 getrennt nach Regionen. Quelle: <http://ncdrisk.org/obesity-population-stacked.html>

Die Studierenden arbeiteten nun zu zweit an den PCs. Da viele interessante Fragenstellungen durch die Studierenden aufkamen, waren sie prinzipiell frei bezüglich ihrer Datenanalyse, sie sollten sich an ihren noch nicht geklärten Fragen orientieren, die während der vorherigen Arbeitsschritten aufgetreten sind. Die Lehrkraft begleitete die Teams, gab Hilfestellungen und Tipps zur weiteren Analyse und beantwortete Fragen bei Schwierigkeiten mit der Software und der Datenanalyse sowie bei den Interpretationen der Ergebnisse. Ein Team suchte nach Alternativen zum BMI, da er, obwohl oft verwendet, nur einen möglichen Hinweis auf bspw. Übergewicht oder Fettleibigkeit gibt. Individuell kann der berechnete Wert ein unpassendes Bild wiedergeben, da er nur Gewicht und Größe verwendet. Die unterschiedlichen Körperbestandteile (wie z. B. Fett oder Muskeln) werden dadurch beispielsweise nicht berücksichtigt. Zudem sind große Menschen durch den BMI etwas benachteiligt und die BMI-Stufengrenzen sind willkürlich festgelegt³. Die Studierenden fanden zwar einige Alternativen, allerdings hatten sie entweder ähnliche Schwächen (Broca-Index, Ponderal-Index, BAI) oder sie ergeben zwar besser interpretierbare Werte, sind aber komplexer zu berechnen (ABSI) oder aufwendiger zu ermitteln (Körperfettanteil).

Die Datenanalyse konnte entweder mit den verschiedenen Visualisierungen durchgeführt werden, die auf der Webseite bereitgestellt wurden (siehe z. B. Abb. 2) oder mit der Software CODAP⁴ (siehe Abb. 3 und Abb. 4). Abbildung 2 zeigt zum Beispiel ein gestapeltes Flächendiagramm, welches die Entwicklung der Anzahl der weiblichen Fettleibigen verschiedener Regionen vom Jahr 1975 bis 2016 zeigt. Man erkennt eine sehr deutliche Zunahme der absoluten Zahlen (von 69,3 Mio. auf 390 Mio.) in den meisten Regionen. Die Anzahl der Frauen weltweit hat sich in diesem Zeitraum zwar auch erhöht (von 2,03 Mrd. auf 3,65 Mrd.), jedoch nicht so stark wie der Anteil der fettleibigen Frauen. Durch Anklicken der Region werden die jeweiligen Untergruppen sichtbar, bis hin zu den einzelnen Ländern.



Geschlecht
 männlich weiblich

Abb. 3: Darstellung der Entwicklung des weltweiten Durchschnitts-BMI getrennt nach Geschlechtern von 1975 bis 2014 mit CODAP

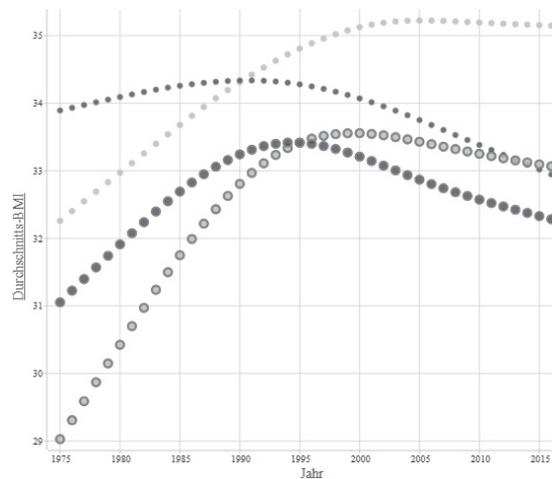


Abb. 4: Darstellung der Entwicklung des Durchschnitts-BMI der Länder American Samoa (hell) und Nauru (dunkel) getrennt nach Geschlechtern von 1975 bis 2016 mit CODAP, die Punkte der Männer sind größer.

Mit Hilfe anderer dynamischer Visualisierungen der Webseite können die Studierenden beispielsweise die Entwicklung des durchschnittlichen BMI getrennt nach Geschlechtern der einzelnen Länder von 1975 bis 2016 anhand Ranglisten erforschen. Anstatt des Durchschnitts-BMI können auch beispielsweise die prozentualen Anteile der weltweiten Untergewichtigen oder Fettleibigen der jeweiligen Länder erkundet werden. Allerdings kann man anhand der verschiedenen Visualisierungen auf der Webseite nicht direkt erkennen, wie sich der Wert des jeweiligen Landes jährlich verändert hat. Dazu wurden Teile des Datensatzes heruntergeladen und in CODAP importiert. Im Folgenden werden exemplarisch zwei Bearbeitungen der Studierenden dargestellt.

Betrachtet man die Veränderung des durchschnittlichen BMI über die Jahre, auch nach Geschlechtern getrennt (Abb. 3), erkennt man einen kontinuierlichen, nahezu linearen Anstieg des durchschnittlichen BMI bei beiden Geschlechtern. Auf Basis dieser Grafik stimmt die Schlagzeile des Artikels.

Eine Äußerung dazu war: „Allerdings ist die Menschheit eine sehr heterogene Gruppe, vor allem was die Ernährung angeht. Vielleicht sorgt eine große Untergruppe mit sehr starkem Anstieg dafür, dass der weltweite Durchschnitts-BMI kontinuierlich ansteigt, es aber nicht in allen Ländern so aussieht.“ Die Untersuchungen der Studierenden widerlegten diese Vermutung, da der Durchschnitts-BMI in fast allen Ländern zunahm, nicht nur in einigen bevölkerungsreichen Ländern.

Eine andere aufgekommene Frage war, ob es Länder gibt, bei denen der durchschnittliche BMI nicht (immer) ansteigt? Diese Frage konnte beispielsweise anhand der Länder American Samoa und Nauru beantwortet werden (siehe Abb. 4).

Durch die Ergebnisse ihrer vielfältigen und unterschiedlichen Datenanalysen kamen die Studierenden zu folgenden Einschätzungen:

- Die Menschheit wird immer *dicker* (steigender Durchschnitts-BMI). Auf Grund der Entwicklung von 1975 bis 2016 kann nicht davon ausgegangen werden, dass die Menschheit zu einem früheren Zeitpunkt einmal *dicker* war.
- Generell sind Frauen stärker von Fettleibigkeit (BMI > 30) betroffen als Männer und haben auch einen höheren Durchschnitts-BMI gemäß der WHO Definition.
- In einigen Ländern sind bei beiden Geschlechtern die Anteile der Bevölkerung, die als fettleibig eingestuft werden, sehr extrem. Vor allem bei den Inselstaaten im Pazifik liegen die Anteile

der fettleibigen Bevölkerung oft über 50 % zum Teil sogar über 60 %.

- Bei einigen Ländern entwickelt sich der Durchschnitts-BMI entgegen dem Trend und nimmt ab, wobei dies ab 1975 betrachtet nur bei Frauen zutrifft (Nauru, Bahrain, Singapur, Japan, Spanien, Litauen, Malta oder Tschechien).
- In allen Ländern hat sich der Durchschnitts-BMI der Männer von 1975 bis 2016 erhöht. Betrachtet man die Daten der einzelnen Länder detaillierter (jährliche Veränderung), erkennt man, dass in einigen Ländern der Durchschnitts-BMI seit einigen Jahren auch bei den Männern sinkt (beispielsweise in Nauru ab 1995, in Bahrain ab 1999 und in American Samoa ab 2000)
- Die Anzahl der Untergewichtigen steigt auch. Waren 1975 306 Mio. Menschen untergewichtig, waren es 2016 418 Mio. Menschen. In Bezug zur gestiegenen Weltbevölkerung (von 4,06 Mrd. auf 7,43 Mrd.) ist der Anteil jedoch von ca. 7,5 % auf 5,6 % gesunken.
- Bezüglich der Qualität der Daten gab es vereinzelt Bedenken, da insgesamt 1700 unterschiedliche Studien zusammengefasst wurden und deswegen möglicherweise unterschiedliche Erhebungsmethoden zum Einsatz kamen.
- Darüber, den BMI als Variable für Fettleibigkeit zu verwenden, herrschte Uneinigkeit, da der BMI diverse Nachteile hat. Jedoch ist er auf Grund seiner leichten Erhebungsmethode vor allem bei der Anwendung auf Populationen ein praktikables Maß.

Bei der anschließenden Diskussion der Ergebnisse der Studierenden sind sie aufgefordert ihre ursprüngliche Bewertung der Schlagzeile zu überprüfen und anhand ihrer erarbeiteten Ergebnisse gegebenenfalls anzupassen. Die Lehrkraft kommentiert die Äußerungen der Studierenden bezüglich ihrer Plausibilität und korrigiert nur bei falschen Begründungen oder Formulierungen, zum Beispiel, wenn Zusammenhänge kausal ausgedrückt werden, dies auf Grund der Datenlage aber nicht zweifelsfrei möglich ist.

Die meisten Studierenden waren nach den Analysen und mit den Informationen, die sie innerhalb des Moduls erhalten hatten, der Meinung, dass es weltweit einen deutlichen Trend zu mehr Übergewicht und Fettleibigkeit gibt, dass der durchschnittliche BMI seit 1975 kontinuierlich ansteigt, und dieser Trend aufgehalten werden sollte. Zumindest in einigen wenigen Ländern geht der Trend des Durchschnitts-BMI nach unten, die Daten und die Kontextinformationen bieten aber keine Erklärung für das wie. Zudem be-

schrieben die meisten Studierenden ebenfalls, dass sie die übermittelte Konnotation als negativ empfunden haben, obwohl der Rückgang von untergewichtigen Menschen als positiv gewertet wurde. Obwohl Untergewichtige ebenfalls in der Schlagzeile vorkamen, beschäftigte sich nur ein Team mit deren Analyse. Wenige Studierende sahen die Verwendung des BMI als ein großes Manko an, so dass sie, trotz der deutlichen Ergebnisse der Datenanalyse, der Studie keine große Aussagekraft zuschreiben wollten. Die Argumente der Studierenden gegen den BMI bezogen sich auf die möglichen Fehldeutungen, wenn der BMI einen Wert liefert, der auf Grund der Eigenschaften des Individuums unpassend ist. Ihre Meinung dazu änderte sich auch nicht, als argumentiert wurde, dass unpassende Ergebnisse der Individuen innerhalb der Population statistisch nicht von Bedeutung sind und der BMI deswegen trotz der individuellen Nachteile bei Populationen angewendet werden kann.

Nach der Beendigung der Diskussion sollen die Studierenden anhand der gestellten Fragen bzw. den genannten *fehlenden/weiteren Informationen* die Bereiche herausarbeiten, die für eine evidenzbasierte Bewertung einer datenbasierten Aussage wichtig sind. Dazu gehören:

- **Datenerfassung:** Verschiedene Datenerfassungen (bspw. Befragung, Beobachtungsstudie oder Experiment) haben mit spezifischen Problemen zu kämpfen. Neben der Stichprobengröße können beispielsweise auch der Ort, die Zeit und die Anzahl der Durchführungen Einfluss auf die Ergebnisse haben und dementsprechend mögliche Schlussfolgerungen auf Populationen erschweren.
- **Operationalisierung der Variablen:** Die Operationalisierung der Variablen beeinflusst die Messbarkeit und hat somit einen entscheidenden Einfluss auf die Genauigkeit der Ergebnisse und dementsprechend auf mögliche Schlussfolgerungen (bspw. die Messung der Körperlänge ist einfacher und eindeutiger als die Messung des IQ). Die Skalierung der Variablen hat wiederum Einfluss auf mögliche statistische Berechnungen.
- **statistische Berechnungen:** Die Berechnungen der einzelnen statistischen Parameter sind zwar klar definiert, welcher aber verwendet wird und ob alle Werte in die Berechnung einfließen ist variabel.
- **Repräsentation:** Die Art und Weise sowie der Ort der Repräsentation können die Interpretation einer Aussage beeinflussen. Grafische als auch numerische Darstellungen können für Manipulationen anfällig sein.

- **Kontextwissen:** Das Kontextwissen hat eine große Bedeutung bei der Übertragung der statistischen Ergebnisse auf die Realsituation. Je mehr zu der jeweiligen Thematik an Kontextwissen bekannt ist, desto besser können datenbasierte Aussagen auf ihre Plausibilität eingeschätzt und eingeordnet werden.

Je nach der zur Verfügung stehenden Zeit und dem Wissensstand der Studierenden/Schüler*innen kann dies mit geringer oder mit viel Unterstützung seitens der Lehrkraft ablaufen. Im ersten der vier Seminar-Durchgänge bekam dazu die ganze Gruppe die von ihnen geäußerten *fehlenden Informationen* einzeln auf Karteikarten ausgehändigt. Sie sollten diese, erst einmal nach Gemeinsamkeiten (ohne weitere Hinweise) gruppieren und soweit möglich mit einem Oberbegriff versehen. Unpassende Zuordnungen wurden im Plenum aufgegriffen und nachgefragt warum beispielsweise „was ist mit dick gemeint“ dem Kontextwissen zugeordnet wurde. So wurde versucht, die Gruppierungen im Hinblick auf die fünf Bereiche zu verbessern. Nicht immer ist eine *fehlende Information* trennscharf zuzuordnen beispielsweise „wie aussagekräftig der BMI ist“, kann dem Kontextwissen zugeordnet werden, aber auch der Operationalisierung der Variable. Dies bietet wiederum Anlässe zur Diskussion und verdeutlicht auch Zusammenhänge zwischen den Bereichen.

Zur Unterstützung, da die Gruppe zwar die Bereiche *Datenerfassung* und *statistische Berechnungen* gefunden hatte, aber dann nicht weiterkam, wurden die fehlenden Bereiche dann angegeben, um die übrigen *Informationen* dann zuzuordnen. Im Anschluss wurden die Bereiche von der Lehrkraft näher erläutert, um deren Bedeutung für eine evidenzbasierte Bewertung zu verdeutlichen. In den späteren Durchgängen des Moduls wurden die Bereiche dann zu Beginn beispielhaft durch zwei bis drei Beispiele vorgegeben, um mehr Zeit für die Diskussionen zur Verfügung zu haben.

Teil 2 des Moduls: Produzieren und Kommunizieren datenbasierter Informationen

In diesem Teil des Moduls sollen die Studierenden anhand von bereits zusammengefassten Ergebnissen einer Studie einen Artikel inklusive Schlagzeile verfassen und diese dann untereinander und mit realen Artikeln vergleichen. Dadurch sollen die Studierenden erfahren, welchen Einfluss das Vorwissen und subjektive Entscheidungen auf die Veränderung und Auswahl von kommunizierten Informationen haben können. Zudem sollen sie durch die Diskussionen im Plenum herausfinden, wie zwischen einer statistischen Information

und einer Meinung, welche mit statistischen Informationen unterlegt ist, unterschieden werden kann.

Um einen solchen möglichen Einfluss aufzeigen zu können, sollen die Studierenden anhand einer Studie über Studierende von *Ernst and Young*⁵ (EY) einen kurzen Artikel (ca. eine halbe Seite) mit einer passenden Schlagzeile schreiben. Dies eignet sich als Arbeitsauftrag zwischen zwei Sitzungen. Für diese Studie wurden von EY 4300 Studierende online befragt, welche Branchen sie attraktiv finden und welche Faktoren ihnen bei der Berufswahl wichtig sind. Diese Studie wurde deshalb ausgewählt, weil es neben den Studienergebnissen und einer Pressemitteilung von EY zusätzlich noch zahlreiche journalistische Produkte gibt, anhand derer eine Problematik eines Vergleichs verschiedener Quellen deutlich gemacht werden kann.

Bei der Diskussion über die Inhalte der geschriebenen Artikel (aus Zeitgründen wurden dazu drei möglichst verschiedene ausgewählt, dies aber auch den Studierenden mitgeteilt) zeigte sich, dass es auf Grund des individuellen Hintergrundes (beispielsweise persönliche Auswahl, Wissensgrundlage oder Einstellungen) der Studierenden große Unterschiede zwischen einzelnen Artikeln gab. Die Unterschiede waren dabei vor allem inhaltlicher Natur, der Fokus lag beispielsweise auf den wichtigen Faktoren bei der Berufswahl oder auf Unterschieden zwischen den Geschlechtern, aber auch bezüglich der möglichen Branchenauswahl der Studierenden. Dennoch versuchten die Studierenden möglichst viele der Studienergebnisse in ihren Artikel zu integrieren.

Im Anschluss wurden dann die Schlagzeilen der Studierenden (Beispielsweise: „Scheuklappengesellschaft! Engagement für die Mitmenschen stirbt aus!“, „Männer denken an Geld, Frauen an die Sicherheit“, „Für Top-Studenten ist der Öffentliche Dienst weniger attraktiv“) mit der Originalschlagzeile („Gerne zu Vater Staat – Studenten zieht es in den öffentlichen Dienst“) der Pressemitteilung von EY verglichen. Die vorhandene Vielfalt sorgte nur bedingt für Überraschung, auch nicht, dass manche Schlagzeilen negativ aufgefasst wurden. Denn trotz der gleichen Studie als Informationsgrundlage entstehen durch subjektives Entscheiden und individuelle Dispositionen teilweise sehr unterschiedliche Schlagzeilen.

Das Hinzuziehen mehrerer Artikel diverser Zeitungen zu dieser Studentenstudie (z. B. Manager Magazin, Spiegel Online, Süddeutsche.de, Wirtschafts-Woche) zeigte jedoch ein anderes Bild. Sechs von sieben Artikeln erwähnten nahezu den gleichen Inhalt wie die Pressemitteilung von EY. Weitere Ergebnisse der Studie waren in diesen Artikeln jedoch nicht vorhanden.

Die Schlagzeilen waren ebenfalls alle sehr ähnlich, sie bezogen sich alle darauf, dass die Studenten in den Öffentlichen Dienst wollen, nur mit unterschiedlichen Konnotationen. Die Artikel bezogen sich dem Anschein nach auf die Pressemitteilung von EY und nicht direkt auf die Ergebnisse der Studentenstudie.

Eine in dem Modul mehrfach beobachtete Ansicht von Studierenden war, dass die Aussagekraft und Glaubwürdigkeit eines Artikels oder der Schlagzeile steigt, wenn in anderen Quellen ähnliches geäußert wird. Mit dem Vergleich der Artikel und deren Schlagzeilen soll den Studierenden bewusstwerden, dass das alleinige Vorhandensein weiterer Quellen, die sich ähnlich äußern, kein Kriterium für deren Aussagekraft und Glaubwürdigkeit ist, wenn sie sich beispielsweise auf die gleiche Quelle beziehen. Eine Überprüfung weiterer Quellen kann allerdings aufdecken, was sich durch den Autorenwechsel in Bezug auf die Aussage und Konnotation verändern kann.

Zum Abschluss des zweiten Teils des Moduls wurden die alternativen Schlagzeilen der Studierenden im Plenum verglichen, welche sie zum Artikel „*So dick war die Menschheit noch nie. Erstmals leben mehr Fettleibige als Untergewichtige auf der Welt*“ schreiben sollten. Beispielantworten der Studierenden waren „Trotz stetig steigendem Gewicht der Menschheit, gibt es immer noch Hunger in vielen Ländern“, „Fettleibigkeit der Weltbevölkerung nimmt seit 1975 zu“, „Der BMI zeigt verwirrende Ergebnisse über Fettleibigkeit“ oder „Einer von vier Menschen mit starker Fettleibigkeit lebt in den USA“. Einerseits zeigen sie die Unterschiede auf, die durch subjektive Entscheidungen entstanden sind. Andererseits ähnelten sich viele der Schlagzeilen, da sich die Studierenden oftmals an der Originalüberschrift orientierten.

Um sich unabhängig eine eigene Meinung über (kontroverse) Themen bilden zu können, ist es wichtig andere Meinungen und Deutungen zu diesen Themen als solche auch zu erkennen. Innerhalb der Diskussion im Plenum stellten die Studierenden fest, dass manche Wörter bereits eine gewisse Wertung innehaben. Vor allem aber tragen wertende Wörter wie Adjektive, „die einem eine bestimmte Wertung suggerieren“, zur jeweiligen Auslegung bei. Auf die Frage der Lehrkraft, was denn für die jeweilige negative, neutrale oder positive Auslegung der Informationen verantwortlich sei, gaben die Studierenden das eigene Vorwissen sowie die eigene Erfahrung mit dem Thema an.

3 Reflexion

Das CQS-Modul bietet den Studierenden durch „begleitetes“ Erfassen datenbasierter Aussagen und

durch die Diskussionen im Plenum die Möglichkeit zu erfahren, welche Bereiche für eine angemessene Bewertung einer datenbasierten Aussage wichtig sind und wo mögliche Stolpersteine liegen.

Nach viermaliger Durchführung des beschriebenen CQS-Moduls kann festgehalten werden, dass den meisten Studierenden das Einschätzen und Bewerten von datenbasierten Aussagen aus den Medien zunächst schwerfällt, ebenso selbstständig relevante statistische Fragen zu stellen. Die meisten Studierenden hinterfragen zu Beginn bei einer Aussage hauptsächlich die Bereiche Datenerfassung („wie wurden die Daten erhoben“) und statistische Berechnungen („Wie viele Fettleibige und wie viele Untergewichtige gibt es?“). Die Operationalisierung von Variablen („Wie ist der BMI definiert“) wurde anfangs nur von wenigen berücksichtigt. Auch ein kritischer Bezug zum Kontextwissen („Aussagekraft des BMI-Wertes hinterfragen“) wurde zu Beginn nur vereinzelt hergestellt. Im Verlauf des Moduls erfuhren die Studierenden innerhalb der Diskussionsrunden, dass sowohl die Bewertung der Datenerfassung als auch der statistischen Berechnungen ohne einen Bezug zum Kontextwissen oder der Operationalisierung von Variablen wenig Aussagekraft hat. Erst durch die Berücksichtigung und Überprüfung der Datenerfassung, der statistischen Berechnungen, der Repräsentation und der Operationalisierung der Variablen und der Einbeziehung von Kontextinformationen kann eine evidenzbasierte Einordnung bezüglich der Aussage zu der jeweiligen Thematik stattfinden.

Zu Beginn des Moduls sorgte bei manchen Studierenden für Verwirrung, dass selbst gegensätzliche Auslegungen sich auf gleiche statistische Ergebnisse beziehen können. Durch Diskussionen im Plenum konnte den Studierenden aufgezeigt werden, wie viele unterschiedliche Interpretationen allein mit einer Schlagzeile entstehen können, wenn nur wenige Informationen vorhanden sind und diese sich mit Vorwissen vermischen. Ähnliches gilt auch für Ergebnisse statistischer Berechnungen. Eine sachliche Begründung der Bewertung, sowie die Nachvollziehbarkeit der Argumente sorgen für eine hohe Aussagekraft der Interpretation.

Die Unterscheidung zwischen einer statistischen Information und einer Meinung, welche mit statistischen Informationen unterlegt ist, stellte sich für die Studierenden ebenfalls als ein Problem dar. Zu Beginn des Moduls wurde sehr häufig nur der Vergleich mit anderen Quellen genannt, wenn dort ähnliches steht, dann „ist es eine Information und keine Meinung“. Durch die Schlagzeilenvergleiche und die Diskussionen im Plenum wurde aufgezeigt, dass bei anderen Quellen nicht nur der Inhalt, sondern auch deren Quellenanga-

ben überprüft werden sollten. Zudem identifizierten die Studierenden zusätzliche Anhaltspunkte, die beim Unterscheiden zwischen statistischer Information und Meinung helfen: Die Verwendung von Adjektiven deutet stark auf eine Meinung hin, da durch diese zusätzlichen Worte Wertungen entstehen. Sofern der Datensatz inklusive Metadaten verfügbar ist, können die statistischen Angaben und anschließend die Aussagen auf deren Plausibilität überprüft werden.

4 Umsetzung in der Schule

Das Bewerten von Aussagen, die auf einer Datenanalyse basieren ist seit 2016 eine zu vermittelnde Kompetenz aller Niveaustufen der Leitidee Daten und Zufall für die Klassen 7, 8 und 9 des Bildungsplans für Mathematik der Sekundarstufe 1 in Baden-Württemberg⁶. Dies wird im ersten Teil des Moduls aufgegriffen. Das eigene Formulieren solcher Aussagen ist nur für Niveaustufe E vorgesehen, dies wird mit dem zweiten Teil umgesetzt.

Für zwei bis drei Unterrichtsstunden (Arbeit mit den Daten) sollten PCs mit Internetzugang zur Verfügung stehen. Für die Lehrkraft ist es von Vorteil, wenn sie sich bereits mit einer Visualisierungssoftware auskennt (z. B. CODAP, Fathom, TinkerPlots). Damit das Modul wie beschrieben durchgeführt werden kann, werden folgende Vorkenntnisse für die Schüler empfohlen:

- mit Begriffen wie Durchschnitt, Mittelwert oder Prozenten umgehen können,
- relevante Informationen aus Texten entnehmen können,
- mit dynamischen Visualisierungstools umgehen können (z. B.: CODAP, Fathom),
- grafische Darstellungen beurteilen können.

Inhaltliche Kompetenzen: Neben den Vorkenntnissen – die durch Wiederholung geübt werden – können folgende Kompetenzen erarbeitet werden:

- Aussagen, die auf einer Datenanalyse basieren, formulieren und bewerten,
- Einfluss der Datenerfassung und der Operationalisierung der Variablen auf die Daten kennen und in einer Beurteilung berücksichtigen.

Das Arbeiten mit dem Artikel gibt Gelegenheit das Entnehmen von Informationen aus Texten zu üben, diese zu bewerten und das Kontextwissen zu vertiefen. Bei selbstgewählten Artikeln könnten zusätzlich Tabellen oder andere statistische Darstellungen vorkommen, um weitere statistische Fähigkeiten einzubinden. Neben den inhaltlichen Kompetenzen werden

zudem prozessbezogene Kompetenzen (Problemlösen, Argumentieren und Kommunizieren), sowie verschiedene Leitperspektiven des Bildungsplans für Mathematik der Sekundarstufe 1 in Baden-Württemberg⁷ in das Modul einbezogen.

Die dem Modul zugrundeliegende Struktur (siehe Abb. 1) kann auch bestehen bleiben, wenn die eigentlichen Arbeitsmaterialien (downloadbar unter <https://www.ph-ludwigsburg.de/15972.html>) an die jeweilige Lerngruppe je nach Interesse, Aktualität, Fähigkeiten und Verfügbarkeit von Daten flexibel angepasst werden. Wichtig dabei ist, dass die Schlagzeile eine datenbasierte Aussage beinhaltet, die bewertet werden soll, sowie dass die Daten, auf die im Artikel verwiesen wird, zur Verfügung stehen und mit ihnen gearbeitet werden kann. Zudem können den Schüler*innen (Teil-)Fragen in den jeweiligen Arbeitsschritten vorgegeben werden, um beispielsweise die Datenanalyse zu unterstützen. Für das Modul sollte ein Zeitaufwand von ca. sieben Unterrichtsstunden à 45 Minuten eingeplant werden. Wobei die Arbeit mit der Schlagzeile und dem Artikel ca. zwei Unterrichtsstunden benötigt, für die Datenanalyse und die daran angeschlossene Erarbeitung der Kriterien sollten ca. drei Unterrichtsstunden eingeplant werden und für die Diskussion über Schlagzeilen und die selbstgeschriebenen Artikel ca. zwei Unterrichtsstunden. Alternativ kann das Modul auch in einem Projekt eingesetzt werden, um in einer interdisziplinären Thematik die statistische Bearbeitung zu motivieren.

Danksagung

Ich möchte Rolf Biehler, Andreas Prömmel und Philipp Ullmann für die hilfreichen konstruktiven Anmerkungen und Kommentare danken.

Anmerkungen

- 1 Die Schlagzeile und der dazugehörige Artikel stammen von Zeit Online. <https://www.zeit.de/wissen/gesundheit/2016-03/uebergewicht-adipositaernaehrung-bmi-entwicklung> (aufgerufen am 25.01.19)
- 2 Definition der BMI Stufen nach der WHO. <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>
- 3 Hintergrundinformationen zum BMI gibt es zum Beispiel auf folgender Webseite: <https://www.bmi-rechner.net/artikel/bmi-kritik.htm> (aufgerufen 25.01.19)
- 4 CODAP ist eine kostenlose, browserbasierte Datenanalysesoftware und steht für Common Online Data

Analysis Platform, <https://codap.concord.org/> (aufgerufen am 25.01.19)

- 5 http://www.ey.com/Publication/vwLUAssets/EY_Studentenstudie_2014_-_Branchen/%24FILE/EY-Studentenstudie-2014-Branchen.pdf (aufgerufen am 25.01.19)
- 6 <http://www.bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/SEK1/M/IK/7-8-9/05 S. 42> (aufgerufen am 25.01.19)
- 7 <http://www.bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/SEK1/M/IK/7-8-9/05 S. 4-5> (aufgerufen am 25.01.19)

Literatur

- Adorno, T. W. (2015): *Erziehung zur Mündigkeit* (25. Aufl.). Frankfurt: Suhrkamp.
- Di Cesare, M.; Bentham, J.; Stevens, G.A. et al. (2016): Trends in adult body-mass index in 200 countries from 1975 from 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387 (10026), S. 1377–1396.
- Engel, J. (2014): Open data, civil society and monitoring progress: challenges for statistics education. In: Makar, K.; de Sousa, B.; Gould, R. (eds.): *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014)*. Voorburg, The Netherlands: International Statistical Institute. http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_4F4_ENGEL.pdf (Stand: 25.01.19)
- Fischer, R. (2012): Fächerorientierte Allgemeinbildung: Entscheidungskompetenz und Kommunikationsfähigkeit mit ExpertInnen. In: Fischer, R.; Greiner, U.; Bastel, H. (Hrsg.). *Domänen fächerorientierter Allgemeinbildung*. Linz: Trauner, S. 9–17.
- Frech, S.; Richter, D. (Hrsg.) (2013): *Politische Kompetenzen fördern* (Didaktische Reihe). Schwalbach: Wochenschau Verlag.
- Gal, I. (2002): Adults' Statistical Literacy: Meanings, Components, Responsibilities, In: *International Statistical Review* 70(1), S. 1–25.
- Gramm, C. (2010): *Schlüsselqualifikationen für Staatsbürger*. München: Olzog.
- Kant, I. (1784): Was ist Aufklärung? In: *Berlinische Monatsschrift*. Dezember 1784, S. 481–494.
- Ridgway, J. (2015): Implications of the data revolution for statistics education. *International Statistical Review* 84(3), S. 528–549.

Anschrift des Verfassers

Achim Schiller
 Institut für Mathematik und Informatik
 Pädagogische Hochschule Ludwigsburg
 Reuteallee 46
 71634 Ludwigsburg
 Schiller01@ph-ludwigsburg.de