

Data Literacy im Wissenschaftsjournalismus – Facetten journalistischer Datenkompetenz und Fortbildungsbauusteine zu deren Schulung

KARIN BINDER, MÜNCHEN & MARKUS VOGEL, HEIDELBERG

Zusammenfassung: Die Data Literacy Charta des Stifterverbandes (2021) betont die Wichtigkeit der Förderung von Datenkompetenz. Journalist:innen spielen durch die Verbreitung von Daten und Fakten eine besondere Rolle und können maßgeblich zur Meinungsbildung in der Gesellschaft beitragen. Erstaunlicherweise gibt es bislang zwar anekdotische Berichte zu typischen Fehlern in medialen Darstellungen, aber kaum empirisch gesicherte Erkenntnisse zur Data Literacy von Journalist:innen. Im vorliegenden Beitrag werden inhaltsbezogene Kompetenzfacetten zur Konzeptualisierung der Data Literacy vorgeschlagen sowie ein entsprechendes Training zur Förderung der Data Literacy von Wissenschaftsjournalist:innen vorgestellt (evaluiert mit Prä- und Post-Test).

Abstract: The Stifterverband's Data Literacy Charter (2021) emphasizes the importance of promoting data literacy. Journalists play a special role by disseminating data and facts and can make an important contribution to shaping public opinion. Surprisingly, there are so far only anecdotal reports on typical errors in media representations but hardly any empirically verified findings on the data literacy of journalists. In this article, content-related competence facets for the conceptualization of data literacy are proposed and a corresponding training to enhance the data literacy of science journalists is presented (evaluated with pre- and post-test).

1. Einleitung

Die vom Stifterverband veröffentlichte Data Literacy Charta (2021) verdeutlicht die Bedeutung der Förderung der Datenkompetenz in unserer Gesellschaft. Schüller et al. (2019) führen aus, dass Data Literacy auch Fähigkeiten beinhaltet, „die ein mündiger Bürger in der Digitalisierung benötigt, um sich in einem Überangebot von Daten und Informationen zurechtzufinden und fundierte Entscheidungen zu treffen – im Alltag wie auf verschiedenen politischen Ebenen“. Die im Zitat beschriebenen Daten und Informationen, mit denen wir regelmäßig konfrontiert werden, werden in vielen Fällen von Journalist:innen recherchiert, aufbereitet, kondensiert dargestellt und verbreitet. Die Relevanz des Blicks auf die Data Literacy von Journalist:innen ergibt sich außerdem dadurch, dass die in Medien gezeigten und

möglicherweise manipulativen Darstellungen das Potenzial besitzen, die öffentliche Meinung zu beeinflussen. Obwohl eine umfassende Sammlung manipulativer statistischer Darstellungen aus den Medien in populärwissenschaftlicher Literatur existiert (siehe z. B. Krämer, 2015; Christensen & Christensen, 2015; Bosbach & Korff, 2011, oder die Unstatistik des Monats von Bauer, Gigerenzer, Krämer und Schüller, o. D.), sind erstaunlicherweise kaum empirische Befunde bezüglich der typischen Fehler und der Data Literacy von Journalist:innen vorhanden. Lück und Boczek (2020) kritisieren ebenfalls die bislang fehlende Beachtung, welche Rolle die Lehre empirischer Methoden und der Statistik in der Ausbildung von Journalist:innen in Deutschland spielt.

Dieser Beitrag will dieses Forschungsdesiderat in den Blick nehmen und hat zum Ziel 1) einen ersten Vorschlag für inhaltsbezogene Kompetenzfacetten zur Messung der berufsrelevanten Data Literacy von Journalist:innen zu erarbeiten, und 2) entwickelte Trainings-Bausteine für einen 6-stündigen Kurs zur Förderung der Data Literacy von Journalist:innen im Sinne der Mathematikdidaktik als Design Science (Wittmann, 1998) zu konstruieren. Der entwickelte Kurs wurde schließlich mit 19 Wissenschaftsjournalist:innen durchgeführt, von denen 12 sowohl einen Prä-Test als auch einen Post-Test zur Erfassung der Data Literacy ausgefüllt haben. Im Projekt *Journalistische Data Literacy (JouDaLi)* der beiden Autor:innen wurde ein Training sowie ein Erhebungsbogen entwickelt, welche die folgenden fünf Inhaltsbereiche fokussieren, die in einer informellen Expertenbefragung als zentral für die Arbeit von (Wissenschafts-)Journalist:innen angesehen wurden: 1) manipulative statistische Diagramme (z. B. Achsen- und Flächenmanipulationen), 2) Probleme im Umgang mit Anteilen (Bezugsgröße), 3) Probleme im Umgang von Anteilen von Anteilen (Prozent-Prozentpunkte-Problematik und Verwechslung bedingter Wahrscheinlichkeiten), 4) die Bedeutung der Korrelation (z. B. ihre Verwechslung mit Kausalität) sowie 5) ein Verständnis für Signifikanztests (z. B. die Bedeutung von p-Werten und die Frage, welche Forschungsfragen damit beantwortet werden und welche nicht). Die deskriptiven Ergebnisse geben erste Hinweise darauf, welche Facetten journalistischer Data Literacy durch die entwickelten Trainingsbausteine gefördert werden konnten. Zudem geben die

Rückmeldungen der teilnehmenden Wissenschaftsjournalist:innen Hinweise darauf, welche weiteren Aspekte von Data Literacy zusätzlich in den Blick genommen werden könnten. Langfristiges Ziel der Forschungsbemühung wird eine Steigerung der Data Literacy von Journalist:innen sein, die sich nachfolgend in einer Qualitätssteigerung medialer Produkte zeigen sollte.

2. Mathematikhaltige Kommunikation durch Journalist:innen

Bei medialer Berichterstattung handelt es sich um eine Expert:innen-Lai:innen-Kommunikation, die in vielen Fällen mathematikhaltig bzw. statistikhaltig ist (vgl. Vohns, 2018). Die Trennung der Rollen von Expert:innen und Lai:innen in der Informationsvermittlung ist eine wichtige Erkenntnis hinsichtlich der Frage, welche Art von „Data literacy“ Journalist:innen erwerben müssen, denn die Mathematik aus der Produzent:innensicht ist eine andere als die, die in der Konsument:innensicht erlernt werden muss (Fischer, 2012; Gal, 2002). Wissenschaftsjournalist:innen tauchen in ihrer täglichen Arbeit tatsächlich in beide Rollen ein, da sie beim Verstehen wissenschaftlicher Arbeiten und bei der Befragung von Expert:innen in der Konsument:innenrolle sind, dann aber in eine Produzent:innenrolle wechseln, wenn sie die Informationen für die Rezipient:innen aufbereiten.

Die Kommunikation von Daten und statistischer Befunde oder mathematikhaltige Kommunikation im Allgemeinen betrifft Journalist:innen, aber im besonderen Maße auch die spezialisierteren Berufsgruppen der Wissenschaftsjournalist:innen und Datenjournalist:innen, die im Folgenden in Abschnitt 2.1 unterschieden werden. In Abschnitt 2.2 wird beschrieben, welche Rolle die Data Literacy in diesem Berufsfeld einnimmt oder einnehmen sollte. Daran anschließend werden in Abschnitt 2.3 grundlegende prozessbezogene Facetten von Data Literacy ausdifferenziert und in den Bereich der medialen Berichterstattung hineingedacht.

2.1 Journalismus, Datenjournalismus und Wissenschaftsjournalismus

Beim Begriff *Journalist:in* handelt es sich nicht um eine geschützte Berufsbezeichnung. Damit ist verbunden, dass es vielfältige Möglichkeiten der Journalismusausbildung gibt: Neben einem Journalismusstudium bestehend aus einem Bachelor- und einem konsekutiven Masterstudiengang sind auch die Ausbildung in Journalist:innenschulen, nicht-konsekutiven Masterstudiengängen oder über

Voluntariate möglich und gängig (Auermann & Elitz, 2013; Lorenz, 2009). Je nach gewähltem Bildungsweg unterscheidet sich daher das statistische Vorwissen in Tiefe und Breite von Journalist:innen immens.

Einige Journalist:innen, die in den verschiedenen Medienhäusern tätig sind, ordnen sich dem Wissenschaftsjournalismus (engl. Science Journalism) und/oder dem Datenjournalismus (engl. Data Journalism oder Data Driven Journalism) zu. Dabei gilt grundsätzlich: „Wissenschaftsjournalismus ist Journalismus, der das Themen- und Ereignisfeld Wissenschaft behandelt.“ (Blöbaum, 2017). Dabei ist auch hier eine Abgrenzung zu Journalismus im Allgemeinen nicht ganz einfach, da beispielsweise viele freie Journalist:innen unter anderem – aber eben nicht ausschließlich – über wissenschaftliche Themen schreiben. Trennt man das Berufsfeld „Journalismus“ hinsichtlich des Grades der Automatisierung auf, werden weitere Bezeichnungen wie „Datenjournalismus“ und „computational journalism“ vorgeschlagen und von Wissenschaftsjournalismus abgegrenzt (Weinacht und Spiller, 2014; Gutouning, Radkohl, Goldgruber & Stoiber, 2022). Während 2013 etwa 35 Datenjournalist:innen in Deutschland bekannt waren, gab es im Jahr 2005 wohl bereits etwa 3000 Personen, die sich dem Wissenschaftsjournalismus zugeschrieben haben (Weischenberg et al., 2006; Weinacht und Spiller, 2014). Allerdings ist eine eindeutige Quantifizierung schwierig, weil es sich auch hier nicht um geschützte Berufsbezeichnungen handelt. Studiengänge für Wissenschaftsjournalist:innen bieten inzwischen teils Datenjournalismus als eine mögliche Schwerpunktsetzung im Studium an. Umgekehrt werden Datenjournalist:innen auch benötigt, wenn nicht über wissenschaftliche Themen berichtet wird, wie beispielsweise bei Wahlberichterstattungen.

Da das im vorliegenden Artikel vorgeschlagene Training zur Förderung der Data Literacy für Wissenschaftsjournalist:innen entwickelt und mit Personen sowohl durchgeführt als auch evaluiert wurde, die sich (auch) diesem Teilgebiet des Journalismus zuschreiben, soll der Fokus im Folgenden vor allem auf dem Wissenschaftsjournalismus liegen. Inwiefern eine Übertragung der Trainingsbausteine auf Journalist:innen im Allgemeinen denkbar ist, wird abschließend diskutiert.

Die massenmediale Wissenschaftsberichterstattung über verschiedene mediale Produkte beeinflussen „hauptsächlich wie Laien über Wissenschaft, ihre Prozesse, Ergebnisse und Maßnahmen denken,

welche Einstellungen sie ausbilden und zum Teil auch, wie sie sich verhalten“ (Guenter, 2017, S. 18; siehe auch Guenther, Froehlich, Milde, Heidecke & Ruhrmann, 2015). In diesem Sinne sind Wissenschaftsjournalist:innen weit mehr als der verlängerte Arm von Wissenschaftler:innen und fungieren nicht (nur) in einer dolmetschenden Funktion von Wissenschaftler:innen (Hettwer & Zotta, 2008).

Wissenschaftsjournalismus erfüllt in unserer Gesellschaft verschiedene Funktionen, die beispielsweise von Kohring (1997) beschrieben werden: 1. Informationstransfer von der Wissenschaft in die Öffentlichkeit, 2. Bildung und Aufklärung für die Gesellschaft, 3. Kritik und Kontrolle, auch im Sinne einer wissenschaftlichen Selbstkontrolle, 4. Schaffung von Akzeptanz für Wissenschaft und Technologie. Die von Kohring (1997) dargestellten Funktionen des Wissenschaftsjournalismus erfordern allesamt statistische Kompetenzen.

Der Transfer von Erkenntnissen (1.), die aus empirisch arbeitenden Wissenschaften abgeleitet werden, kann nur dann gewinnbringend erfolgen, wenn einerseits Journalist:innen die gefundenen Evidenzen ausreichend nachvollziehen können und andererseits Fähigkeiten zur Kommunikation entsprechender Befunde vorhanden sind. So umfasst der Aspekt des Informationstransfers eine rezeptive und eine produktive Komponente. Unter den Aspekt der Aufklärung (2.) fällt, dass Journalist:innen über etwaige Wissenschaftsmythen informieren können (z. B. der angeblich schädlich hohe Cholesteringehalt in Eiern oder der vermeintliche Einfluss von Lerntypen auf den Wissenserwerb). Hierzu sind ausreichende statistische Kenntnisse notwendig, um etwaige widersprüchliche Befunde richtig interpretieren zu können. Dabei stellt die Auswahl statistischer Informationen immer auch gleichzeitig ein Aufklärungshindernis dar, da diese in medialer Berichterstattung in kondensierter Form vorliegen und Empfänger dem Zustandekommen der Kennzahlen auch zu einem gewissen Grad vertrauen müssen (vgl. auch Vohns, 2018). Eine Kritik an wissenschaftlicher Evidenz (3.) ist nur dann möglich, wenn statistische Argumente in wissenschaftlichen Veröffentlichungen oder entsprechenden Pressemitteilungen ausreichend gut nachvollzogen werden können. Auch eine Schaffung von Akzeptanz von wissenschaftlichen Befunden (4.) ist ohne statistische Grundkompetenzen undenkbar. Beispielsweise gehört dazu auch die Kommunikation der Unsicherheit der gefundenen Befunde bei schlechter Datenlage (Feufel et al., 2022). So konnten Dries et al. (2024) am Beispiel der Wirksamkeit von Masken als Schutz vor

Corona-Infektionen zeigen, dass die bewusste Kommunikation der Unsicherheit empirischer Befunde (z. B. weil die Datenlage noch nicht ausreichend gut ist) vor einem etwaigen späteren Vertrauensverlust der Bevölkerung bewahren kann, wenn sich wissenschaftliche Befunde aufgrund neuer Erkenntnisse verändern.

2.2 Die Bedeutung von Data Literacy in der journalistischen Ausbildung

Lück und Boczek (2020) betonen die immense Relevanz, die Statistik in der akademischen Ausbildung von Journalist:innen (und zwar nicht nur von Wissenschaftsjournalist:innen) einnehmen sollte: „Grundlegende Kenntnisse in Statistik sind für Journalistinnen und Journalisten heute unerlässlich. In Anbetracht der stetig wachsenden Menge an verfügbaren Daten steigen auch die Anforderungen an einen sicheren Umgang mit den Möglichkeiten zur Beschaffung, Aufbereitung, Auswertung und Darstellung. Darüber hinaus werden Journalisten regelmäßig mit Studien und Statistiken Dritter (Politik, Wissenschaft, Unternehmen/PR, Interessensverbänden) konfrontiert und müssen die Qualität und Gültigkeit einschätzen, um unabhängige und sorgfältige Berichterstattung sicherstellen zu können.“. Gleichzeitig monieren Lück und Boczek (2020), dass eine grundlegende Statistikausbildung in Journalismusstudiengängen bislang nicht selbstverständlich ist. Zwar zeigen sie in ihren Analysen, dass zwei Drittel der Journalismusstudiengänge statistische Inhalte implementiert haben, diese seien aber häufig in empirischen Forschungsmethodenkurse zu Beginn des Bachelorstudiengangs vorgesehen und auch nicht immer verpflichtend.

Zudem konzentrieren sich viele Statistik-Lehrpläne an Schulen und Hochschulen nach wie vor auf die Vermittlung der Mechanik einer Mathematik, die sich für die Verwendung gut strukturierter numerischer Daten eignet um eine bestimmte Klasse von Problemen zu lösen, wohingegen die Beschäftigung mit neuen Klassen von realen Problemen, die neue Arten von Daten und die Notwendigkeit komplexerer Zugangsweisen mit sich bringen, oftmals unterbleibt (vgl. Batanero et al., 2011). Dies führt unter anderem dazu, dass die Adressat:innen nicht in die Lage versetzt werden, aktuelle Probleme ihrer Umwelt auf der Grundlage von Daten anzugehen. Das Fach Statistik wird daher zwar als mathematisch interessant und herausfordernd wahrgenommen, erscheint jedoch im Zusammenhang mit realen Daten zu Phänomenen und Zusammenhängen in unserer natürlichen, technischen und sozialen Welt als

wenig relevant (vgl. Engel & Ridgway, 2022). Vor diesem Hintergrund wird deutlich, dass eine grundlegende statistische Ausbildung, die sich lediglich in der Beherrschung statistischer algebraisch-technischer Fertigkeiten erschöpft, nur ein (allerdings nicht unwesentlicher) Teil einer Kompetenz des verständigen Umgangs mit Daten im Sinne von Data Literacy sein kann. Für die journalistische Auseinandersetzung mit gesellschaftspolitischen Themen ist es wichtig, dass die statistische Kompetenz über den Konsum oder die Produktion von Statistiken und statistischen bzw. datenbasierten Argumenten hinausgeht und auch die Frage einschließt, wie solche Argumente die Welt um uns herum formen, und wie Statistiken genutzt werden können, um die Realität, in der wir leben, besser zu verstehen (vgl. Weiland, 2017).

Seit dem Beginn des neuen Jahrtausends hat die rasche Zunahme der Digitalisierung wesentlich dazu beigetragen, die analoge Realität zunehmend in digitale Formate zu überführen und immer mehr Lebensbereiche datenbasiert abzubilden (Gutouning et al., 2022; Schüller et al., 2019). So ist die Bedeutung von Daten und deren digitale Verarbeitung in unserer Gesellschaft noch stärker in den Vordergrund gerückt. KI-Systeme, die auf der Analyse immer größerer Datenmengen basieren, werden zunehmend leistungsfähiger und üben einen immer stärkeren Einfluss auf unser Verhalten und unsere Entscheidungsprozesse aus (Ng et al., 2021). Vor dem Hintergrund dieses Wandels wird die journalistische Ausbildung in den Bereichen Datenaufbereitung, -repräsentation, -interpretation und -kommunikation in Feldern des wissenschaftlichen und öffentlichen Diskurses eine drängende und weitreichende Aufgabe, insbesondere auch deshalb, weil im Unterschied zu vor etwa 20 Jahren mittlerweile prinzipiell jede Person über die sozialen Medien Einfluss auf diesen Diskurs nehmen kann, ob in redlicher oder nicht-redlicher Weise.

Angesichts der zunehmenden inhaltlichen Komplexität des Berufsfeldes Journalismus und wachsenden Anforderungen ist nicht verwunderlich, dass in der Berichterstattung auch Fehler entstehen; Bauer, Gigerenzer, Krämer & Schüller (2018) sprechen hierbei auch von der „Unstatistik des Monats“. Für einen produktiven Umgang mit irreführenden Pressemitteilungen erscheint jedoch eine ko-konstruktive Zusammenarbeit von Journalismus und Wissenschaft zielführend, wie dies bereits in Einrichtungen wie dem ScienceMediaCenter (SMC), der Wissenschaftspressekonferenz (WPK) oder dem Nationalen Institut für Wissenschaftskommunikation (NaWik)

erfolgt. Ein paritätischer Austausch scheint auch deshalb geboten, da auch in wissenschaftlichen Darstellungen irreführend kommuniziert werden kann, etwa in der fehlerhaften Verwendung oder Deutung des Begriffs der statistischen Signifikanz (vgl. Dirnagl, 2024).

2.3 Prozessbezogene Facetten von Data Literacy und ihre Bedeutung in der medialen Berichterstattung

Mit den komplexer werdenden Anforderungen geht die Entwicklung der Frage einher, was unter Data Literacy zu verstehen ist. Während Gal (2002) noch von Statistical Literacy spricht und darunter die Kompetenz versteht, korrekte Schlussfolgerungen aus statistischen Daten zu ziehen, diese kritisch zu beurteilen und sachgerecht darüber zu kommunizieren, erweitert Gould (2017) diese Kompetenz zu einer Data Literacy, bei der auch die Datenerfassung und die Erstellung von Datendarstellungen (z. B. Diagrammen) berücksichtigt wird. In aktuellen Rahmenwerken wurde das Konzept der Datenkompetenz um die Konzepte des Datenschutzes und der Datenethik erweitert (z. B. Schüller et al., 2019; Wolff et al., 2017).

Die Definitionen der Begriffe Statistical Literacy und Data Literacy haben sich im Laufe der Zeit ausgeweitet und werden zumindest derzeit in der Literatur nicht einheitlich verwendet oder konzeptualisiert (Schreiter et al., 2023; Friedrich et al., 2024). Vor diesem Hintergrund ist zu verstehen, dass verschiedene Arbeiten darauf zielen, die mannigfaltigen und teils ineinander greifenden prozessbezogenen Facetten einer Datenkompetenz unter einem Gesamtkonstrukt von Data Literacy zu fassen (z. B. Schüller et al., 2019; Ridsdale et al., 2015; Parasie & Dagiral, 2013; Mandinach & Gummer, 2016; Koltay, T., 2017). Wenn sich diese Arbeiten auch in ihren Konzeptualisierungen unterscheiden, lässt sich doch eine gemeinsame Grundstruktur nachzeichnen, die sich in fünf zentrale prozessbezogene Facetten von Data Literacy unterteilen lassen (Ridsdale et al., 2015). Diese beziehen sich auf die verschiedenen Teilprozesse in der Datenarbeit: 1) Daten sammeln und erzeugen, 2) Daten verstehen und interpretieren, 3) Daten analysieren und bewerten, 4) Daten kommunizieren sowie 5) ethische und rechtliche Aspekte der Datennutzung berücksichtigen. Die erste Kompetenzfacette 1), das Sammeln und Erzeugen von Daten, beschreibt die Fähigkeit, relevante Datenquellen zu identifizieren und eigenständig Daten zu erheben. In einem digitalen Zeitalter, das von einer enormen Datenflut geprägt ist, erfordert dies

nicht nur technisches Wissen, sondern auch die Fähigkeit, die Qualität und Verlässlichkeit von Datenquellen kritisch zu bewerten (Calzada Prado & Marzal, 2013). Eng damit verbunden ist die Fähigkeit 2), Daten zu verstehen und zu interpretieren. Diese Facette bezieht sich auf das Wissen und die Fähigkeiten, die benötigt werden, um Daten korrekt zu lesen, in den jeweiligen Kontext einzuordnen und sinnvolle Schlüsse aus ihnen zu ziehen (Schüller et al., 2020). Dazu gehört nicht nur das Verstehen von Zahlen und statistischen Kennzahlen, sondern auch die Fähigkeit, Muster, Trends oder Ausreißer in Daten zu erkennen. Eine weitere wichtige Facette ist die Fähigkeit 3), Daten zu analysieren und zu bewerten. Dies umfasst die Anwendung geeigneter Methoden der Datenanalyse sowie das kritische Hinterfragen von Datenquellen und Ergebnissen (Ridsdale et al., 2015). Insbesondere in einer Welt, in der „Fake News“ und verzerrte Statistiken zunehmend ein Problem darstellen, wird die kritische Bewertung von Daten immer wichtiger, um fundierte und vertrauenswürdige Aussagen treffen zu können. Neben der Analyse ist auch die 4) Kommunikation von Daten eine Schlüsselkompetenz der Data Literacy. Daten allein sind oft schwer verständlich oder unzugänglich für Laien. Daher kommt der Fähigkeit, Daten in verständliche und anschauliche Formate zu übersetzen, eine entscheidende Bedeutung zu (Schüller et al., 2020). Dies kann durch die Nutzung von Visualisierungen (wie Diagrammen oder Infografiken) geschehen, die komplexe Daten auf einfache Weise vermitteln. Schließlich gehört auch die Berücksichtigung von 5) ethischen und rechtlichen Aspekten zur Data Literacy. Der Umgang mit Daten, insbesondere mit personenbezogenen oder sensiblen Daten, wirft viele Fragen zu Datenschutz und Datensicherheit auf (Calzada Prado & Marzal, 2013).

Die Bedeutung von Data Literacy erfährt im Journalismus eine Ausdifferenzierung dahingehend, dass die Protagonist:innen nicht nur für sich selbst Datenliteralität entwickeln und nutzbar im Verstehen der Welt machen, sondern getreu ihrer Funktionen nach Kohring (1997) Teil der Vermittlung von Daten und evidenzbasierten Schlussfolgerungen sind und im Idealfall selbst dazu beitragen, Datenliteralität in ihrer Leserschaft zu fördern. Es ist also zusätzlich „[...] zum eigenen Lernfortschritt eine Transfer- und Übersetzungsleistung zu vollbringen: Wissenschaftsjournalist:innen müssen Statistik selbst verstehen und lernen, Statistik in unterschiedlichen fachlichen Kontexten anwenden können und sie schlussendlich in eine für das Publikum zugängliche textuelle und/oder visuelle Form zu bringen.“ (Lück & Boczek,

2020, S. 110). Diese Ausdehnung von Data Literacy entspricht dem in Abschnitt 2.1 genannten Informationstransfer, der eine rezeptive und eine produktive Komponente beinhaltet.

Mit dem enormen Anwachsen der zur Verfügung stehenden Datenmasse stehen Journalist:innen vor der zunehmend komplexeren Aufgabe, große Mengen an Daten zu analysieren und in verständliche Geschichten zu übersetzen. Einigen prozessbezogenen Facetten der zuvor ausbuchstabierten Data Literacy kommen dabei eine spezifische Bedeutung zu: Die Datenbeschaffung und die Recherche sind in der journalistischen Arbeit von entscheidender Bedeutung. Dies umfasst das Auffinden vorhandener Datensätze und das Anfordern neuer Daten, beispielsweise über Anfragen nach dem Informationsfreiheitsgesetz (Parasie & Dagiral, 2013). Datenquellen sind kritisch zu hinterfragen, um ihre Verlässlichkeit und Relevanz für die Berichterstattung zu bewerten. Darüber hinaus spielen Dateninterpretation und -analyse eine zentrale Rolle: Journalist:innen müssen aus Datensätzen Schlüsse ziehen, die gesellschaftlich relevante Trends oder Missstände aufzeigen (Hermida & Young, 2019). Dazu gehört auch das kritische Hinterfragen möglicher Verzerrungen. Ein weiteres wichtiges Element ist die Datenvisualisierung. Da visuelle Kommunikation an Bedeutung gewinnt, ist in der journalistischen Arbeit unerlässlich, Daten so darzustellen, dass sie für das Publikum zugänglich und verständlich sind (Fink & Anderson, 2015). Im Zeitalter der Digitalisierung gehören Infografiken und interaktive Darstellungen inzwischen wie selbstverständlich zu dem journalistischen Instrumentarium, mit dem komplexe Daten anschaulich aufbereitet und präsentiert werden. Im investigativen Journalismus stehen zudem ethische Aspekte im Vordergrund, gerade bei der Arbeit mit sensiblen Daten, wenn der Schutz der Privatsphäre und die Einhaltung rechtlicher Bestimmungen im Fokus stehen (Hermida & Young, 2019). Journalistische Datenkompetenz ist im Zeitalter von Fake News die Voraussetzung für das Gelingen einer seriösen Medienberichterstattung, die sich einem idealtypischen Anspruch zufolge als Vermittlungsinstanz zum Verständnis von ökonomischen, politischen und sozialen Zusammenhängen für eine offene, freie und möglichst vollständige Diskussion verschiedener Standpunkte versteht (vgl. Schicha, 2003) und die Falschmeldungen oder irreführende Darstellungen von Sachverhalten entgegentritt, die in sozialen Medien bewusst oder unbewusst kommuniziert werden. Im Sinne einer sozial verantwortlichen Nachrichtenarbeit können Journalist:innen in ihrer Rolle

als „guardian of the public's trust“ (Boyles & Meyer, 2016) den Bürger:innen helfen, die Informationsflut des digitalen Zeitalters zu bewältigen.

Mit der Informationsflut und der komplexen Natur von Big Data wachsen die Ansprüche an die Datenliteratilität der journalistischen Arbeit in der Aufbereitung und medialen Darstellung von Daten in Feldern des wissenschaftlichen und öffentlichen Diskurses (vgl. Boyles & Meyer, 2016). Dafür ist neben dem Blick auf die eben vorgestellten prozessbezogenen Facetten der Data Literacy jedoch auch eine Betrachtung inhaltsbezogener Facetten einer journalistischen Data Literacy erforderlich, mit denen (auch ungewollt) fehlerhaften und manipulativen medialen Darstellungen vorgebeugt werden soll.

3. Inhaltsbezogene Facetten journalistischer Data Literacy zur Vermeidung fehlerhafter oder manipulativer medialer Darstellungen

Vor dem Hintergrund der vorgestellten theoretischen Fundierung prozessbezogener Kompetenzen der Data Literacy im Wissenschaftsjournalismus sollen im Folgenden auch inhaltliche Kompetenzfacetten entfaltet werden, die Produkt einer Literaturrecherche sowie einer informellen Expertenbefragung sind. Denn die zuvor dargelegten prozessbezogenen Kompetenzen manifestieren sich in konkreten medialen Inhalten, die Kategorien typischer Fehlvorstellungen zugeordnet werden können. Bezogen auf die fünf zentralen prozessbezogenen Facetten von Data Literacy (Ridsdale et al., 2015; s. Abschnitt 2.3) betrachten wir hier eine inhaltsbezogene Ausdifferenzierung der Kompetenzfazette 4) Daten kommunizieren.

Die Daten, die hierbei medial kommuniziert werden, können fehlerhaft oder manipulativ sein. Eine Vielzahl populärwissenschaftlicher Publikationen beschreibt typische Klassen an Fehlern, die regelmäßig in medialen Darstellungen auftreten. Bücher wie „So lügt man mit Statistik“ (Krämer, 2015), „Lügen mit Zahlen“ (Bosbach & Korff, 2011), „Warum dick nicht doof macht und Genmais nicht tötet“ (Bauer, Gigerenzer & Krämer, 2014; erstes Buch zur Unstatistik des Monats), „Grüne fahren SUV und Joggen macht unsterblich“ (Bauer, Gigerenzer, Krämer & Schüller, 2022; zweites Buch zur Unstatistik des Monats), „Achtung Statistik“ (Christensen & Christensen, 2015; Buch zur Webseite und Kolumne „Achtung Statistik“) sind nur einige der empfehlenswerten Sammlungen, die interessierten Leser:innen einen kritischen Umgang mit Daten und einen Aufbau elementarer Data Literacy ermöglichen. Darüber hinaus

bieten die Plattformen „Unstatistik des Monats“ sowie „Achtung Statistik“ (<https://www.achtung-statistik.de/>) Fundgruben für aktuelle statistische Fehldarstellungen.

Im Folgenden werden typische Fehlerklassen erläutert, die in diesen Büchern wiederholt beschrieben werden und die spezifische inhaltliche Facetten *einer journalistischen Data Literacy* darstellen, nämlich:

- Manipulative Diagramme
- Probleme mit Anteilen (Fragliche Bezüge bei Prozenten)
- Probleme mit Anteilen von Anteilen (bedingte Wahrscheinlichkeiten und die Prozent-Prozentpunkte-Problematik)
- Verwechslung von Korrelation und Kausalität
- Probleme bei der Interpretation signifikanter Ergebnisse

Nachfolgend sollen zunächst einige begriffliche Unterscheidungen hinsichtlich manipulativer Darstellung statistischer Informationen in den Medien getroffen werden (Abschnitt 3.1.) und anschließend typische Klassen an Fehlern vorgestellt werden (Abschnitte 3.2). Im Sinne einer Mathematikdidaktik als Design Science (Wittmann, 1998) wurden mithilfe dieser Fehlerklassen anschließend Trainingsbausteine entwickelt, die von den Teilnehmer:innen als möglichst berufsrelevant eingeschätzt werden und die Elemente des 4C/ID-Modell aufweisen (a. learning tasks, b. supportive information, c. just-in-time information, und d. part-task practice, Merriënboer, Clark, & Croock, 2002, S. 44).

3.1 Statistische Fehler und manipulative Darstellungen

Eine zugangsnahe Möglichkeit sich dem Problem von irreführenden oder fehlerhaften Statistiken in der journalistischen Berichterstattung in einem Training zur Stärkung der Data Literacy zu nähern, ist das Aufgreifen typischer Fehler oder manipulativer Darstellungen statistischer Natur in den Medien. Im Folgenden sollen dabei grundsätzlich zwei verschiedene Arten manipulativer Darstellungen kontrastiert werden, indem eine rezeptive und eine produktive Komponente des Informationstransfers unterschieden wird. Betrachtet man die rezeptive Komponente des Informationstransfers, so soll unter einer *manipulativen Darstellung* einer statistischen Information eine solche Darstellung verstanden werden, die grundsätzlich das Potenzial hat Leser:innen,

Zuhörer:innen in eine bestimmte Richtung zu lenken oder zu beeinflussen. Dabei gilt jedoch zu beachten, dass eine Vielzahl an medialen Darstellungen dieses Potenzial besitzt, ohne dass die Person, die die Darstellung erstellt hat, überhaupt die Absicht einer solchen Einflussnahme hatte. Daher kann neben der rezeptiven Komponente auch eine produktive Komponente des Informationstransfers betrachtet und damit fokussiert werden, ob eine (manipulative) einflussnehmende Absicht hinter der Erstellung der manipulativen Darstellung steckt oder nicht. In vielen Fällen lässt sich jedoch beim Betrachten manipulativer Darstellungen nicht mehr feststellen, ob tatsächlich die Absicht einer bewussten Manipulation unterstellt werden kann oder nicht.

Als weitere Unterscheidung muss beachtet werden, dass nicht jede manipulative Darstellung auch als statistische Fehldarstellung gewertet werden kann. In der medialen Berichterstattung finden sich einerseits klare statistische Fehldarstellungen, wenn etwa bedingte Wahrscheinlichkeiten verwechselt werden (z. B. die Verwechslung des Anteils der Infizierten unter den positiv Getesteten mit dem Anteil der positiv Getesteten unter den Infizierten). Andererseits finden sich auch Darstellungen, die manipulativ wirken, aber nicht durch normative Kriterien als eindeutig falsch klassifiziert werden können. Ein Beispiel hierfür sind die zwei verschiedenen Diagramme zu den Einschulungszahlen in Deutschland in Abb. 1.

In beiden Fällen werden die gleichen Daten dargestellt. Der Unterschied in der Art der Darstellung ist a) der Platz, der im Diagramm jeweils über den Säulen abgebildet wird und b) die unterschiedlich lang gewählte horizontale und vertikale Achse, die auch zu Unterschieden in die Säulenhöhen führt. Die beiden Diagramme transportieren dabei für Rezipient:innen unterschiedliche Botschaften: Im ersten Diagramm nimmt man etwa gleichbleibend niedrige Einschulungszahlen wahr, im zweiten Diagramm hingegen etwa gleichbleibend hohe Einschulungszahlen. Bei den Entscheidungen hinsichtlich der Länge der vertikalen oder horizontalen Achse oder dem Platz über den Säulen gibt es keine normativen Kriterien, ab denen von einem statistischen Fehler gesprochen werden kann. Die Übergänge zu einer manipulativen Darstellung sind hier also fließend. Auch die Wahl des Anfangspunktes (d. h. welches Jahr als erstes in der Achse abgebildet wird, z. B. 2000 oder 2005) kann entscheidend sein, um beispielsweise gerade noch einen deutlichen Rückgang der dargestellten Zahlen in den Daten abbilden zu können.

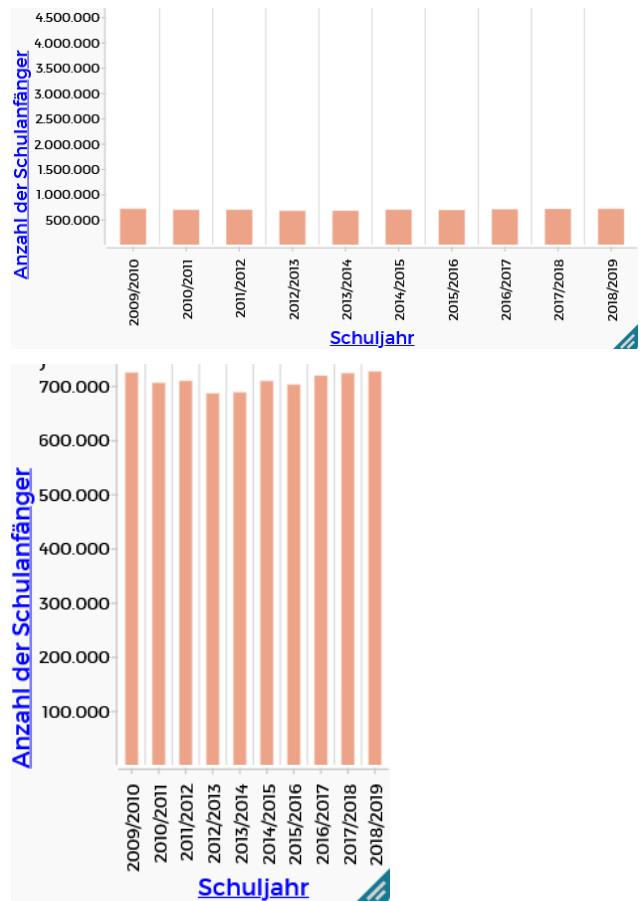


Abb. 1: Einschulzungszahlen, die einmal die Botschaft gleichbleibend niedrig (oben) oder gleichbleibend hoch (unten) vermitteln (Daten: Statistisches Bundesamt; Software für die Diagrammerstellung: CODAP)

3.2 Typische mediale Fehler im Umgang mit Daten und Datendarstellungen als Grundlage der Ausdifferenzierung inhaltsbezogener Facetten journalistischer Data Literacy

Im Folgenden werden typische und wiederkehrende mediale Fehler im Umgang mit Daten und Datendarstellungen vorgestellt, die empirische Anknüpfungspunkte für eine Ausdifferenzierung von inhaltsbezogenen Facetten einer Data Literacy darstellen, welche spezifisch die journalistische Arbeit adressieren.

1) Die erste Fehlerklasse, die wiederholt in populärwissenschaftlicher Literatur zu medialen statistischen Fehlern vorgestellt wird, beschreibt *manipulative statistische Diagramme* (siehe z. B. Krämer, 2015; Bosbach & Korff, 2011, Binder, Krauss und Krämer, 2019). Besonders häufig werden dabei zwei Fehlerkategorien beschrieben: 1) Abgeschnittene oder verzerrte Achsen in statistischen Diagrammen und 2) manipulative Flächendiagramme. Abgeschnittene Achsen führen in Säulen- und Liniendiagrammen zu einer verzerrten Darstellung des Sachverhalts. Auf den ersten Blick wirken Anstiege (oder Abfälle) oder auch Unterschiede zwischen

verschiedenen Klassen auf diese Weise stärker, als sie durch eine neutrale Darstellung wirken würden, die einen natürlichen Nullpunkt berücksichtigt (falls dies durch das vorgegebene Skalenniveau überhaupt möglich ist). Trotz vielfacher Erwähnung der manipulativen Wirkung solcher Darstellungen, finden sich diese nach wie vor in Digital- und Printmedien. In Videos, in denen die jeweiligen Diagramme oft nur wenige Sekunden eingeblendet werden, dürfte der verzerrten Darstellung eine besondere Bedeutung zukommen, da die Möglichkeit der Entlarvung des manipulativen Potenzials des Diagramms zeitlich häufig stark begrenzt wird.

Überdies wird in der Literatur häufig von statistischen Diagrammen berichtet, die eine Flächenmanipulation beinhalten. Ein Beispiel, das an eine Abbildung aus dem BMEL-Ernährungsreport (2018) angelehnt ist, findet sich in Abbildung 2. Häufig werden bei derartigen Diagrammen Verdopplungen (bzw. Verdreiblungen oder wie in Abb. 2 Ver-82/61-fachungen) nicht dargestellt, indem der Flächeninhalt verdoppelt (verdreiblacht bzw. ver-82/61-facht) wird, sondern es wird absichtlich oder versehentlich der Radius verdoppelt (verdreiblacht bzw. ver-82/61-facht), was bei einem zweidimensionalen Objekt jedoch zu einem vierfachen (bzw. 9-fachen, bzw. $(82/61)^2$ -fachen) Flächeninhalt führt. Dies ist in der Darstellung problematisch, da die meisten Menschen bei der Betrachtung entsprechender Diagramme die Flächeninhalte der Diagramme wahrnehmen und nicht die eindimensionale Ausdehnung der Objekte (Rolfes, 2021).

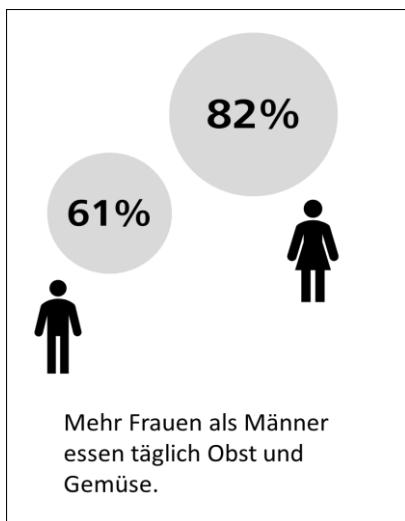


Abb. 2: Irreführendes Flächendiagramm zum Konsum von Obst und Gemüse von Männern und Frauen (angelehnt an eine Grafik des BMEL-Ernährungsreports, 2018, Seite 7).

2) Die zweite Fehlerkategorie, die häufig in der Literatur beschrieben wird, bezieht sich auf *Fehlbezüge bei Anteilen*. Ein prominentes Beispiel, das jedes Jahr mehrfach in medialen Darstellungen vorzufinden ist, sind Aussagen der Form „Jeder zweite Deutsche lebt allein.“ Hierbei wird häufig fälschlicherweise geschlossen, dass die Aussagen „50 % der Haushalte sind 1-Personen-Haushalte“ und „50 % der Menschen leben alleine“ den gleichen Anteil beschreiben, was jedoch aufgrund unterschiedlicher Bezugsgröße (Haushalte vs. Personen) nicht richtig ist. Dass diese Gleichsetzung unzulässig ist, verdeutlicht auch Abb. 3 (Binder, Krauss & Krämer, 2019).

Ein weiteres Beispiel, das wir dieser Kategorie zuordnen, ist die immer wiederkehrende Verwechslung „x % mehr \leq x % weniger“, die beispielsweise regelmäßig beim Gender-Pay-Day zu beobachten ist (vgl. auch Loos, 2014; Krämer, 2015). So war 2022 auf tagesschau.de im Titel einer Meldung zu lesen „Frauen verdienen 18 Prozent weniger“, während in der Unterüberschrift die Information zu finden war „[...] lag der durchschnittliche Stundenlohn von Männern 18 Prozent höher“. Nur eine der beiden Aussagen kann jedoch richtig sein, weil sich die 18 Prozent im ersten Fall auf das Gehalt der Männer beziehen und im zweiten Fall auf das niedrigere Gehalt der Frauen und somit zwei unterschiedliche Bezugsgrößen zu grunde liegen. Wenn in medialen Darstellungen lediglich über eine *Lohnlücke* von beispielsweise 18 Prozent berichtet wird, ist ebenfalls unklar, welche Bezugsgröße für diese Lohnlücke gewählt wird (das Gehalt der Männer oder das Gehalt der Frauen). Diese Art der Darstellung ist also doppeldeutig.

3) Aufbauend auf der Kategorie von Problemen mit Anteilen kann eine Fehlerkategorie beschrieben werden, die sich mit *Anteilen von Anteilen* beschäftigt. Darunter fallen die berühmte Prozent-Prozentpunkte-Problematik sowie Probleme im Umgang mit bedingten Wahrscheinlichkeiten. Die Prozent-Prozentpunkte-Problematik beschreibt Verwechslungen von absoluten und relativen Veränderungen von Anteilen.

Während absolute Veränderungen von Anteilen in Prozentpunkten ausgedrückt werden, werden relative Veränderungen in Prozent ausgedrückt. Wenn also mit dem alten Medikament 4 von 100 Personen geheilt werden und mit dem neuen Medikament 5 von 100 Personen, dann ist das neue Medikament um einen Prozentpunkt besser als das alte Medikament (absolute Veränderung). Drückt man die Veränderung jedoch relativ ist, so ist das neue Medikament um 25 Prozent besser als das alte (bzw. das

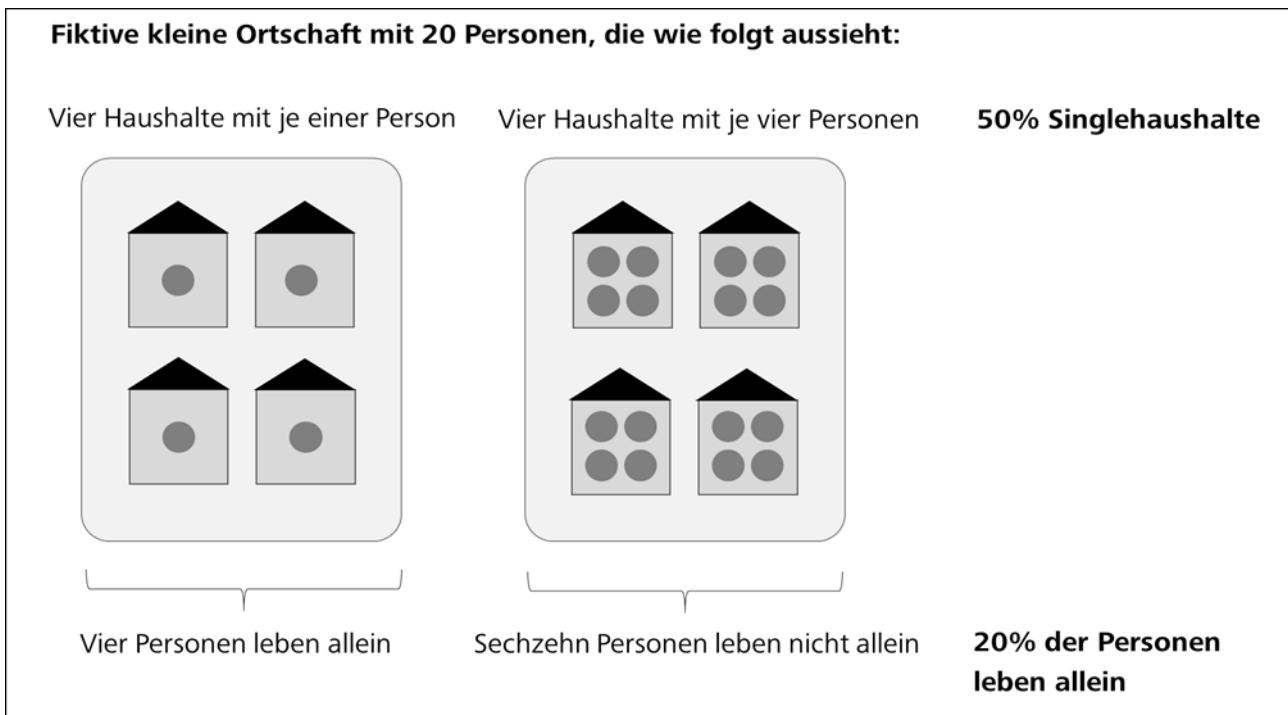


Abb. 3: Gegenüberstellung von Haushalten und Personen bei der Aussage „Jeder zweite Deutsche lebt alleine“ (aus Binder, Krauss & Krämer, 2019)

alte Medikament um 20 Prozent schlechter als das neue). Weil in Medien häufig nur von „Prozent“ gesprochen wird – unabhängig davon, ob eigentlich Prozentpunkte gemeint sind, oder nicht – lässt sich in Berichterstattungen oft nur schwer oder auch gar nicht nachvollziehen, ob eine absolute oder relative Veränderung ausgedrückt werden soll.

Da relative Veränderung oft effektvoller wirken als absolute Veränderungen (vgl. auch obiges Beispiel), werden diese in medialer Berichterstattung häufig präferiert, obwohl die absolute Veränderung die neutralere Berichterstattung bedeuten würde (vgl. auch Binder, Krauss & Gigerenzer, 2020; Wegwarth & Gigerenzer, 2011). Das Harding-Zentrum für Risikokompetenz empfiehlt für die Kommunikation entsprechender statistischer Informationen (z. B. in Bezug auf Gesundheitsrisiken) die Darstellung in Form von Fakten- oder Iconboxen. Ähnlich problematisch ist die Kommunikation bedingter Wahrscheinlichkeiten, die medial häufig mit der zugehörigen invertierten bedingten Wahrscheinlichkeit verwechselt werden: Nur weil viele befragte SUV-Fahrer Die Grünen wählen, darf nicht gefolgert werden, dass viele Grünen-Wähler auch SUV fahren (vgl. Bauer et al., 2022).

4) Eine weitere Fehlerkategorie, die oft in medialen Darstellungen vorkommt und in der Literatur

wiederholt beschrieben wurde, sind die *Überinterpretationen von Korrelationen* im Sinne von Kausalitäten (vgl. Weber-Stein & Engel, 2021; Bauer et al., 2022, Krämer 2015). Aus einem Zusammenhang zwischen zwei Merkmalen darf noch nicht geschlossen werden, dass eines der beiden Merkmale auch ursächlich für das andere Merkmal ist. Dieser Zusammenhang kann auch zufällig entstanden sein oder indirekt durch ein anderes Merkmal beeinflusst worden sein. So wird einem Zusammenhang des täglichen Konsums von Äpfeln und einer erhöhten Lebenserwartung vermutlich ganz allgemein ein gesünderer Lebensstil zugrundliegen (mit regelmäßigen Sport, Verzicht auf Alkohol- und Zigarettenkonsum), der sich sowohl auf den Konsum der Äpfel als auch auf die Lebenserwartung auswirken dürfte.

5) Zuletzt betrachten wir die Fehlerkategorie im Zusammenhang mit *statistischer Signifikanz*. Signifikanztests und p-Werte sind bekanntlich mit einer Vielzahl typischer Fehlkonzepte verbunden (vgl. auch Haller & Krauss, 2002, Oakes, 1986, Passon & von der Twer, 2020). Da sich Wissenschaftsjournalist:innen sowohl mit Pressemitteilungen als auch mit den zugrundeliegenden wissenschaftlichen Artikeln selbst in ihrer Recherche auseinandersetzen, ist ein Verständnis der Bedeutung signifikanter und nicht-signifikanter Ergebnisse eine wichtige Facette der Data Literacy im Wissenschaftsjournalismus. Je

nach Anzahl herausgearbeiteter Fehlerkategorien wird bei den Fehlvorstellungen von p-Werten auch von den fantastischen Vier (Passon & von der Twer, 2020), den big five (Kline, 2013) oder dem dirty dozen (Goodman, 2008) gesprochen. Ein einfacher Fehler besteht bereits darin, die Ergebnisse von Signifikanztests im Sinne eines eindeutigen Beweises überzuinterpretieren. Sehr bekannt ist auch die inverse fallacy, indem die eigentliche Bedeutung des p-Wertes mit seiner Invertierung verwechselt wird (vgl. Passon & van der Twer, 2020). Die tatsächliche Bedeutung eines p-Wertes lässt sich anhand des Vergleichs zweier Mittelwerte mithilfe eines t-Tests betrachten. Der p-Wert gibt in diesem Fall die Wahrscheinlichkeit an, einen solchen oder einen noch extremeren Unterschied in den Daten zu erhalten, obwohl in Wirklichkeit die Nullhypothese wahr ist (es gibt keinen Unterschied). Personen, die der „inverse fallacy“ unterliegen, verwechseln diese Wahrscheinlichkeit mit der Wahrscheinlichkeit, dass die Nullhypothese wahr ist, unter der Annahme diese (oder noch extremere) Daten zu erhalten. Die Meldung „Kinder sind genauso infektiös wie Erwachsene“ mit dem Zusatz „kein signifikanter Unterschied.“ (FAZ, 30.4.2020) beinhaltet hingegen einen anderen häufigen Fehlschluss im Zusammenhang mit statistischer Signifikanz. Dieser besteht darin, aus der Nicht-Signifikanz einer Unterschiedshypothese zu folgern, dass statistisch valide gezeigt wurde, die beiden Gruppen unterscheiden sich nicht. Je nach Effektstärke, die dem Signifikanztest zugrundlag, kann der deskriptive Unterschied der Mittelwerte in den Gruppen durchaus beträchtlich sein. Signifikanztests sind in eine bestimmte Richtung angelegt. Werden Unterschiedshypothesen getestet, kann nur erfasst werden, ob ein signifikanter Unterschied besteht oder kein signifikanter Unterschied besteht. Soll hingegen getestet werden, ob zwei Gruppen signifikant gleich sind, könnte ein Äquivalenztest herangezogen werden (Lakens et al., 2018). Erst dann beruhen Meldungen wie „Kinder sind genauso infektiös wie Erwachsene“ auch auf empirischer Evidenz.

Die fünf beschriebenen Klassen typischer medialer Fehler bilden die Grundlage für eine Ausdifferenzierung von spezifischen inhaltlichen Facetten einer journalistischen Data Literacy, auf die hin die im Folgenden vorgestellte Studie konzipiert wurde. Wir erheben mit diesen fünf Facetten keinen Anspruch auf Vollständigkeit, sondern möchten einen ersten Schritt zur empirischen Untersuchung der Data Literacy und der Entwicklung entsprechender Förderkonzepte im (Wissenschafts-)Journalismus vorstellen. Aufgrund der bisherigen fehlenden empirischen

Evidenz in Bezug auf die Data Literacy von Wissenschaftsjournalist:innen sollen diese Kategorien eine erste Diskussionsgrundlage bilden.

4. Konzeption einer Fortbildung zur Förderung journalistischer Data Literacy

Die Themenauswahl für die Fortbildung erfolgte in Rücksprache mit zwei Experten aus dem Bereich des Wissenschaftsjournalismus. Zunächst wurde hierzu literaturgestützt eine Sammlung möglicher Themen erarbeitet und mit den Experten diskutiert. Grundlage hierfür boten populärwissenschaftliche Abhandlungen zu typischen medialen Fehlern, aus denen typische Fehlerklassen abgeleitet wurden. Mit den Experten wurde anschließend reflektiert, ob 1) die Fehlerklassen auch für Wissenschaftsjournalist:innen als berufsrelevant erachtet werden, 2) welcher Schwierigkeitsgrad je Kompetenzfacette für die Zielgruppe angestrebt werden soll und auf welches Vorwissen aufgebaut werden darf, 3) ob wesentliche Kompetenzfacetten fehlen, die im Rahmen einer solchen Fortbildung von den Teilnehmer:innen erwartet würden. Anschließend erfolgten auf dieser Basis eine Fokussierung, Schwerpunktsetzung und Planung eines ersten möglichen Fortbildungskonzepts. In Abgleich mit den Erfordernissen der mit der Wissenschaftspressekonferenz (www.wpk.org) getroffenen Vereinbarung wurde die Fortbildung als Webinar konzipiert und erstreckte sich über einen Tag.

Die inhaltlichen Schwerpunkte, die sich aus der Expertendiskussion und der Rückkopplung mit informell geäußerten Erwartungen und Wünschen von Interessent:innen ergaben, waren die Themenfelder fehlleitende Diagramme, verwirrende Angaben von Anteilen, Fehlerquellen bei den Deutungen von Korrelation und Signifikanz sowie typische, sich wiederholende saisonale Fehler in der Berichterstattung. Mit Blick auf die in Abschnitt 3.2 literaturgestützt eruierten Fehlerquellen lässt sich eine weitgehende Themendeckung feststellen. Die zusätzlich betrachteten typisch wiederkehrenden saisonalen Fehler bilden eine kontextuell-inhaltlich übergreifende Kategorie, die unterschiedliche statistische Fehlerquellen subsummieren. Entsprechend steckten diese Themenfelder die inhaltlichen Bausteine der Fortbildung unter den Titeln „Lügen mit Diagrammen“, „Anteile – teils schön tricky“, „Warum Joggen nicht unsterblich macht und Kirchgang nicht den Blutdruck senkt“, „Signifikant oder nicht signifikant – Das ist hier die Frage“ und „Saisonale typische Fehler“.

Formulierte Ziele der Fortbildung waren einerseits der Schutz vor typischen und häufig

wiederkehrenden Fehlern in der medialen Berichterstattung. Andererseits sollte die Fortbildung aber auch Anregungen geben, statistische Manipulationen in Wirtschaft, Wissenschaft, Medien und Politik aufdecken zu können, um diese entdeckten Inkonsistenzen und Fehlmeldungen selbst in mediale Produkte umwandeln zu können, die die Sachverhalte wieder richtigstellen (vgl. auch Ickstadt, Müller & Weinert, 2022). Um dieses Ziel in der gebotenen Kürze in effizienter Weise anzustreben, wurde die Zeitplanung entsprechend um die fünf genannten inhaltlichen Schwerpunkte gestaltet, die auch in Tab. 1 dargestellt sind: In einem Vormittagsblock wurden die Bausteine „Lügen mit Diagrammen“ und „Anteile

„teils schön tricky“ in je einstündigen Zeitslots ausgetragen, die durch eine 15-minütige Pause geteilt wurden. Nach einer 90-minütigen Mittagspause wurden die Bausteine „Warum Joggen nicht unsterblich macht und Kirchgang nicht den Blutdruck senkt“, „Signifikant oder nicht signifikant – Das ist hier die Frage“ und „Saisonale typische Fehler“ thematisiert. Eingerahmt wurde der inhaltliche Kern der Fortbildung durch eine halbstündige Einführung (inklusive Vorstellungen und Programmüberblick) und eine abschließende, ebenfalls halbstündige Zusammenfassung mit einem Austausch und Feedback der Teilnehmer:innen.

Dauer	Titel des Fortbildungsabschnitts	Inhalte
15 min	Prä-Test	
1 Std	„Lügen mit Diagrammen“ Manipulative Diagramme	<ul style="list-style-type: none"> • Achsenmanipulation • Flächenmanipulation • Weitere Fehler und manipulative Darstellungen • Eigenständig manipulative Darstellungen in Videos erkennen
1 Std	„Teils schön Tricky I“ Probleme mit Anteilen	<ul style="list-style-type: none"> • Anteile in verschiedene Schreibweisen umrechnen (z.B. Prozent in „Jeder wievielte“) • Wie viel x % ist, hängt von der Basis ab; vgl. Beispiel „X verdient 18 % weniger als Y“ ≠ „Y verdient 18 % mehr als X“
1 Std	„Teils schön Tricky II“ Probleme mit Anteilen von Anteilen	<ul style="list-style-type: none"> • Bedingte Wahrscheinlichkeiten • Verwechslung von $P(B A)$ mit $P(A B)$ • Eigenständige Konstruktion eines Doppelbaums • Prozent-Prozentpunkte-Problematik • Kritische Reflexion verschiedener Berichterstattung zur Wirksamkeit neuer Behandlungsmethoden
45 min	„Warum Joggen nicht unsterblich macht und Kirchgang nicht den Blutdruck senkt“ Verwechslung von Korrelation und Kausalität	<ul style="list-style-type: none"> • Unterschied Korrelation und Kausalität aus statistischer Perspektive • Unterschied in sprachlicher Hinsicht • Einfluss von Hintergrundvariablen und Scheinkorrelation • Simpson-Paradoxon
45 min	„Signifikant oder nicht signifikant – Das ist hier die Frage“ Probleme bei der Interpretation signifikanter Ergebnisse	<ul style="list-style-type: none"> • Vorstellungsorientierte Erklärung des Begriffs „signifikant“ • Typische Fehlinterpretationen
15 min	„Saisonale typische Fehler: Vom Gender-Pay-Day über die polizeiliche Kriminalitätsstatistik bis hin zum Brustkrebsmonat „Oktober““	Weitere typische Fehlerquellen und Vorschläge für weitere Literatur zur Fortbildung in diesem Themenbereich
15 min	Post-Test	

Tab. 1: Zeitliche und grobe inhaltliche Planung der Fortbildung

Die methodische Ausgestaltung der einzelnen Bausteine folgte der Logik des Instruktionsdesigns des 4C/ID Modells nach van Merriënboer, Clark, & Croock (2002, S. 44), da sich dieses Modell in besonderer Weise für die Entwicklung komplexer Fähigkeiten und spezifischer beruflicher Fähigkeiten eignet (van Merriënboer, 2020). Im Sinne des 4C/ID Modells wurden die konstitutiven Elemente von a. learning tasks, b. supportive information, c. just-in-time information, und d. part-task practice in mikroadaptiver Form aufgenommen und im Kontext von spezifischen statistischen Fragestellungen ausgearbeitet.

Die Lernaufgaben (a.) stehen im Zentrum des 4C/ID Modells und spiegeln reale Probleme wider, die schrittweise in ihrer Komplexität zunehmen sollen. Im Rahmen der Kurskonzeption zur Förderung einer journalistischen Data Literacy wurde für die Aufgabengestaltung auf authentische Beispiele aus den Tagesnachrichten zurückgegriffen, die zum Teil auch aus den Reihen der Teilnehmer:innen mitgebracht wurden. So dienten diese zu Aufgaben umgearbeiteten Beispiele der Förderung eines authentischen Problemlösens (vgl. van Merriënboer & Kirschner, 2018) und unterstützten die Kursteilnehmer:innen dabei, zielgerichtet ihre Fähigkeiten im Umgang mit Daten weiter zu entwickeln.

Unterstützende Informationen (b.) lieferten den Teilnehmer:innen das notwendige Hintergrundwissen und theoretische Verständnis, um die Lernaufgaben erfolgreich zu bewältigen. Diese wurden vor der Bearbeitung der Aufgaben in kurzen Inputphasen vermittelt und boten somit den Kursteilnehmer:innen konzeptionelle und kognitive Unterstützung (vgl. Kirschner et al., 2006).

Dagegen wurden prozedurale Informationen (c.) während der Aufgabenbearbeitung auf Anfrage mikroadaptiv dann („just in time“) bereitgestellt, wenn einzelne Kursteilnehmer:innen spezifische Handlungsanweisungen benötigten (z. B. in Break-Out-Räumen). Dem 4C/ID Modell zufolge helfen diese Anweisungen insbesondere bei der schrittweisen Ausführung regelbasierter Prozesse.

Zum Abschluss eines jeden Arbeitsbausteines boten spezifische Teilaufgabenübungen (d.) den Kursteilnehmer:innen die Gelegenheit, die gelernten Facetten von Data Literacy gezielt zu automatisieren und zu festigen. Solche Teilaufgabenübungen sind darauf ausgerichtet, spezifische Fertigkeiten in isolierten Übungen zu trainieren (van Merriënboer & Kirschner, 2018).

5. Erste Messung der Data Literacy-Facetten von Wissenschaftsjournalist:innen (Prä-Post)

In Anlehnung an die in Abschnitt 3.2 vorgeschlagene Konzeptualisierung der inhaltsbezogenen Facetten der Data Literacy im Journalismus wurde eine Operationalisierung vorgenommen und parallelisierte Prä- und Post-Tests entwickelt, die im Rahmen einer ersten Fortbildung mit Wissenschaftsjournalist:innen zum Einsatz kamen. Aufgrund der in Abschnitt 3.2 beschriebenen Fehlertypen, die regelmäßig in populärwissenschaftlicher Literatur beschrieben wurden und auch mediale Berichterstattung der Kategorie Wissenschaftsjournalismus betrifft, wurde vermutet, dass die Teilnehmer:innen in einigen Fällen den Fehlvorstellungen unterliegen. Am meisten Vorwissen und am wenigsten Schwierigkeiten wurden in Bezug auf das Thema „Irreführende Diagramme“ erwartet, weshalb dieses Thema als Einstiegsthema der Fortbildung gewählt wurde.

5.1 Beschreibung der Kursteilnehmer:innen

An der eintägigen Fortbildung haben 19 Journalist:innen teilgenommen, die sich alle selbst als Wissenschaftsjournalist:innen sehen. Aufgrund der unterschiedlichen möglichen Ausbildungswege unterschieden sich die Teilnehmer:innen hinsichtlich der statistischen Vorbildung, sowie auch in der Intensität, in der sie sich in der täglichen Arbeit mit statistischen Inhalten auseinandersetzen müssen. Da sich unter den Teilnehmenden auch Professor:innen und Datenjournalist:innen befanden und sich die Teilnehmer:innen bewusst für den Kurs entschieden haben, ist eher von einer Personengruppe auszugehen, die über überdurchschnittliches Vorwissen und Interesse in Bezug auf statistische Themen verfügt. Die Bewerbung der Fortbildung hatte explizit Wissenschaftsjournalist:innen angesprochen.

Da sich Teilnehmende verspäteten oder aufgrund terminlicher Überschneidungen schon früher aus der Fortbildung ausscheiden mussten, liegen uns die vollständig ausgefüllten Prä- und Posttests von 12 der 19 Teilnehmenden vor, von denen wir nachfolgend die Ergebnisse berichten.

Diese Teilnehmer:innen (8 Frauen, 2 Männer, 2 ohne Angabe) waren zwischen 26 und 75 Jahre alt (im Mittel 46,5 Jahre; $SD=13,6$) und gaben Abiturnoten zwischen 1,1 bis 3,0 an (im Mittel 2,0; $SD=0,5$). Es handelte sich also um eine sehr heterogene Gruppe, auch weil die Teilnehmenden nicht ausschließlich aus dem Bereich der Print-Medien stammten,

sondern beispielsweise auch Personen an der Fortbildung teilgenommen haben, die überwiegend im Hörfunk tätig sind.

5.2 Instrumente

Die inhaltsbezogenen Facetten von Data Literacy der Wissenschaftsjournalist:innen wurden mithilfe eines Prä-Tests und eines parallelisierten Post-Tests direkt vor und direkt nach der Fortbildung erhoben. Prä- und Post-Test bestanden jeweils aus 11 Items (7 offene und 4 geschlossene Items). Die offenen Items wurden von zwei geschulten Raterinnen unabhängig voneinander kodiert. Die Interraterreliabilität lag bei den Items zwischen 87,5 % und 100 %.

Tab. 2 zeigt die Zuordnung der Items zu den einzelnen Facetten: 2 Items zu irreführenden Diagrammen, 2 Items zur Bedeutung von Anteilen mit einem Fokus auf fehlerhafte Bezugsgrößen, 3 Items zu Anteilen von Anteilen mit einem Fokus auf bedingte Wahrscheinlichkeiten und die Prozent-

Prozentpunkte-Problematik, 2 Items zur Verwechslung von Korrelation und Kausalität sowie 2 Items zur Bedeutung von p-Werten (bei Über- bzw. Unterschreitung des Signifikanzniveaus).

Ziel des vorliegenden Beitrags ist noch nicht die Entwicklung und Validierung eines abschließenden Testinstruments zur Messung der Data Literacy von Wissenschaftsjournalist:innen, sondern die Erarbeitung eines ersten Vorschlags möglicher Kompetenzfacetten der Data Literacy von Wissenschaftsjournalist:innen und der Identifikation möglicher weiterer Aspekte, die in diesem Berufsfeld in Bezug auf Data Literacy interessant sind.

Für den Prä-Test ergibt sich ein Cronbachs Alpha von 0,7, für den parallelisierten Nachtest nur ein Cronbachs Alpha von 0,6. Auch hier gilt jedoch: Aufgrund der geringen Strichprobengröße ist der Schätzer des Reliabilitätsmaßes mit einem entsprechend großen Schätzfehler behaftet.

Kategorie	Items	Theoretische mögliche Bewertungseinheiten
Irreführende Diagramme	1 Item: Abgeschnittene Achse (offen) 1 Item: Flächendiagramm (offen)	0-1 0-2
Anteile I: Fraglicher Bezug des Prozentwerts	1 Item: Bezug auf Klasse oder Bezug auf Elemente einer Klasse (offen) 1 Item: x Prozent weniger vs. x Prozent mehr (offen)	0-1 0-2
Anteile II: Anteile von Anteilen (d. h. bedingte Wahrscheinlichkeiten und Prozent-Prozentpunkte-Problematik)	1 Item: Bedingte Wahrscheinlichkeiten (geschlossen) 1 Item: Verwechslung bedingter Wahrscheinlichkeiten (offen) 1 Item: Prozent-Prozentpunkte-Problematik (offen)	0-2 0-1 0-1
Korrelation vs. Kausalität	1 Item: Unzulässige kausale Formulierung entdecken (offen) 1 Item: Passende von überinterpretierten Formulierungen abgrenzen (geschlossen)	0-2 0-2
Signifikanztests	1 Item: Interpretation des p-Wertes größer dem Signifikanzniveau (geschlossen) 2 Item: Interpretation des p-Wertes kleiner dem Signifikanzniveau (geschlossen)	0-2,5 0-2

Tab. 2: Verschiedene Kompetenzaspekte von Data Literacy

Jedenfalls lässt sich aus inhaltlicher Perspektive nicht vertreten, bestimmte Items mit geringer Trennschärfe aus dem Post-Test zu entfernen. Ein weiterer Einsatz der entwickelten Prä- und Posttests mit einer größeren Stichprobe muss zeigen, ob die Instrumente die inhaltbezogene Data Literacy von Wissenschaftsjournalist:innen reliabel erfassen kann, die einzelnen Kompetenzfacetten trennscharf sind und ob sich dieses Messinstrument auch eignet, um die Data Literacy solcher Journalist:innen zu messen, die nicht im Wissenschaftsjournalismus tätig sind.

5.3 Ergebnisse

Aufgrund der kleinen Anzahl an Teilnehmer:innen (uns liegen nur die Prä- und Post-Test-Daten von 12 Versuchspersonen vor), sollten alle nachfolgenden Ergebnisse vorsichtig interpretiert werden und weitere empirische Evidenz gesammelt werden. Da die bisherige Befundlage in Bezug auf die Data Literacy von Journalist:innen im Allgemeinen und der Data Literacy von Wissenschaftsjournalist:innen im Speziellen bislang defizitär ist, sollen nachfolgende Ergebnisse einen ersten Eindruck darüber geben, welches Vorwissen Wissenschaftsjournalist:innen mitbringen (und welches nicht) und welche Facetten durch eine kurze Fortbildung gefördert werden können.

Tab. 3 fasst die Ergebnisse des Prä- und Post-Tests für die fünf gebildeten inhaltsbezogenen Kompetenzfacetten zusammen, wobei alle Facetten auf eine Gesamtpunktzahl von 5 Punkten normiert wurden, um eine bessere Vergleichbarkeit innerhalb der Kategorien zu gewährleisten, denen auch im Rahmen der Fortbildung eine vergleichbare zeitliche Gewichtung zukam. Die Wissenschaftsjournalist:innen erzielten im Prä-Test im Schnitt weniger als die Hälfte der möglichen Punkte bei typischen Fehlern in der Kategorie irreführender Diagramme und steigerten sich von Prä-Test auf Post-Test von 1,67 Punkten auf 2,92 (von 5 möglichen) Punkten. Die Beschäftigung mit typischen Diagrammanipulationen kann also zumindest als kurzfristig erfolgreich gewertet werden. Ein ähnliches Bild zeigt sich im Kompetenzaspekt Anteile I, indem die Teilnehmenden sich durchschnittlich von 1,60 auf 3,33 (von 5 möglichen) Punkten steigern konnten. Schwierig scheint für die Wissenschaftsjournalist:innen der Umgang mit Anteilen von Anteilen (siehe Abschnitt 3.2) im Prä-Test. Während die Teilnehmenden im Prä-Test im Schnitt 1,82 Punkte erzielten, konnten sie sich im Post-Test nur auf durchschnittlich 2,45 (von 5 möglichen!) Punkten steigern. Trotz intensiver Auseinandersetzung mit den Themen in der Fortbildung scheint dieser Themenbereich zu komplex, um den Schwierigkeiten mit einer eintägigen Fortbildung begegnen zu können.

Kategorie	MZP	M (SD)	Min-Max
Irreführende Diagramme	Prä	1,67 (1,18)	0-5
	Post	2,92 (1,04)	0-5
Anteile I: Fraglicher Bezug des Prozentwerts	Prä	1,60 (1,35)	0-5
	Post	3,33 (1,23)	0-5
Anteile II: Bedingte Wahrscheinlichkeiten und Prozent-Prozent-Punkte-Problematik	Prä	1,82 (1,18)	0-5
	Post	2,45 (1,40)	0-5
Korrelation vs. Kausalität	Prä	3,28 (1,20)	0-5
	Post	4,17 (0,97)	0-5
Signifikanztests	Prä	4,48 (0,99)	0-5
	Post	4,79 (0,97)	0-5
Gesamtpunktzahl	Prä	12,85 (4,10)	0-25
	Post	17,66 (3,70)	0-25

Tab. 3: Ergebnisse im Prä- und Post-Test; *M*: arithmetisches Mittel, MZP: Messzeitpunkt, *SD*: Standardabweichung, Min: Theoretisches Minimum an erzielbaren Punkten, Max: Theoretisches Maximum an erzielbaren Punkten

Beim Thema Korrelation vs. Kausalität steigerten sich die Teilnehmenden von durchschnittlich 3,28 auf 4,17 Punkte. Beim Thema Signifikanztest starteten die Teilnehmenden bereits mit einem hohen durchschnittlichen Wert von 4,48 (von 5 möglichen Punkten) und steigerten sich dann nur wenig auf 4,79 Punkte (siehe auch Diskussion). Es fällt auf, dass die Teilnehmenden im Prä-Test bereits etwa die Hälfte der Punktzahl erreichen konnten (nämlich 12,85 von 25 möglichen Punkten). Im Post-Test steigerten sich die Teilnehmer dann auf durchschnittlich 17,66 Punkte (wieder von 25 möglichen Punkten).

6. Diskussion und Resümee

6.1 Zusammenfassung

Insgesamt zeigt sich, dass die Wissenschaftsjournalist:innen bereits im Prätest durchschnittlich etwa die Hälfte der möglichen Punkte unseres entwickelten Tests erzielen konnten, was dafür spricht, dass eine gewisse Basiskompetenz in der von uns gemessenen inhaltsbezogenen Kompetenzfacetten durchaus bei der kleinen Gruppe von uns geschulten und getesteten Wissenschaftsjournalist:innen vorhanden war. Ein möglicher Grund dafür könnte sein, dass sich für den Kurs genau die Personen angemeldet haben, die ein großes Interesse an der Thematik haben und daher bereits über gutes Vorwissen zu typischen statistischen Fehlern besitzen (im Sinne eines positiven Sampling-Bias). Im Gesamttest, aber auch in allen einzelnen Kompetenzaspekten erzielten die Teilnehmenden durchschnittlich eine Leistungssteigerung, was im Hinblick auf die Kürze der einzelnen Kursbausteine erfreulich ist, aber noch keine Aussagekraft hinsichtlich langfristiger Effekte hat.

Betrachtet man einzelne Kompetenzfacetten zeigen sich jedoch einige Auffälligkeiten, die abschließend diskutiert werden sollen. Auffällig ist beispielsweise das gute Abschneiden der Wissenschaftsjournalist:innen bei den Aufgaben zur Interpretation signifikanter Ergebnisse anhand eines gegebenen p-Werts im Prä-Test. Aufgrund der Fülle der empirischen Forschungsbefunde zu typischen Missinterpretationen des p-Wertes und signifikante Ergebnisse hätte man hier von einer deutlich schlechteren Vortestleistung ausgehen können. Für eine Weiterentwicklung des Testinstruments könnten hier schwierigere Items entwickelt und eingesetzt werden und auch solche, die eine höhere Berufsrelevanz für die Wissenschaftsjournalist:innen aufweisen. Wichtig wäre beispielsweise die Verwechslung von

signifikanten Ergebnissen mit großen Effekten, die aus inhaltlicher Perspektive eine wertvolle Weiterentwicklung des Testinstruments darstellen dürfte.

Es ist aber davon auszugehen, dass sich die Ergebnisse stark unterscheiden, wenn man den Test nicht Wissenschaftsjournalist:innen, sondern Journalist:innen vorlegt, die eine andere journalistische Schwerpunktsetzung verfolgen und daher in ihrer täglichen Arbeit weniger mit signifikanten Testergebnissen bei wissenschaftlichen Befunden konfrontiert werden und sie infolgedessen im Falle einer Konfrontation mit p-Werten und signifikanten Ergebnissen möglicherweise größere Schwierigkeiten hinsichtlich einer entsprechenden Interpretation haben.

Ebenso mag die relativ gute Leistung der Teilnehmenden im Vortest im Themenbereich der Korrelation überraschen. Im Vergleich dazu hätten wir ein besseres Abschneiden der Teilnehmenden bei typischen Fehlinterpretationen von Diagrammen erwartet, bei denen die Teilnehmenden im Prä-Test im Schnitt nur die Hälfte der Punktzahl erreichten. Grund für diese Annahme war, dass in der Vorstellungsrunde zu Beginn der Fortbildung einige Teilnehmer:innen erwähnt hatten, dass sie gerade hinsichtlich manipulativer Diagramme bereits Vorwissen mitbringen.

Erwartungskonform sind hingegen die bereits vielfach belegten Schwierigkeiten mit Anteilen und mit Anteilen von Anteilen, die sich leider hinsichtlich bedingter Wahrscheinlichkeiten und der Prozent-Prozentpunkte-Problematik durch die Fortbildung nur wenig verbessern ließen.

6.2 Weiterentwicklungspotenziale des Fortbildungskonzepts und Identifikation weiterer Kompetenzfacetten

Ziel des vorliegenden Beitrags war noch nicht die abschließende Entwicklung und Validierung eines Testinstruments zur Messung der Data Literacy von Wissenschaftsjournalist:innen, sondern die Erarbeitung eines ersten Vorschlags für mögliche inhaltsbezogene Kompetenzfacetten sowie die Identifikation weiterer potentieller Facetten, die für die täglich Arbeit im Wissenschaftsjournalismus im Umgang und der Bewertung von Daten relevant sind.

Bei einer erneuten Durchführung der Fortbildung wäre aufgrund der Fülle der vermittelten Inhalte und der damit verbundenen kognitiven Belastung der Teilnehmer:innen eine Auslagerung des Post-Tests auf einen anderen Tag denkbar. Überdies wäre eine

Ausdehnung der Fortbildungskonzepts auf zwei oder mehrere Tage wünschenswert, um die Inhalte intensiver aufarbeiten zu können. Um eine bessere Berufsfeldpassung zu erreichen wäre eine gezieltere Auswahl hinsichtlich der medialen Darstellungen (z. B. Podcast vs. digitale Zeitung) wichtig, die für die jeweiligen Teilnehmer:innen von Belang sind. Beispielsweise waren Teilnehmer:innen teils auf den Hörfunk spezialisiert, was wiederum eigene Problemfelder entstehen lässt. Auch unterscheiden sich Printmedien von Darstellungen in Form von Videos, weil in Videos beispielsweise Diagramme nur sehr kurz eingeblendet werden, was wiederum schneller zu Fehlinterpretationen führen könnte.

Eine abschließende informelle Befragung der Wissenschaftsjournalist:innen nach dem Post-Test ergab tatsächlich mögliche Themenschwerpunkte, die für die Teilnehmenden ebenfalls interessant sein könnten, weil ihnen diese Aspekte in der täglichen Arbeit mit Daten begegnen. Bei einer weiteren Überarbeitung der Testinstrumente könnte somit auf eine noch bessere Passung zum Berufsbild geachtet werden. Einerseits könnte man hierbei auf sprachliche Aspekte achten, zum Beispiel ein stärkerer Fokus auf richtige und falsche mediale Formulierungen

signifikanter Ergebnisse, die sich üblicherweise in Print-Medien finden. Dies geschah zwar bereits punktuell in den Kursbausteinen, könnte aber in eigenen Übungsphasen deutlich ausgeweitet werden. In diesem Zusammenhang könnte außerdem – wie in Abb. 4 illustriert – dem Aspekt Rechnung getragen werden, dass sich Wissenschaftsjournalist:innen in ihrer täglichen Arbeit in zwei verschiedenen Rollen mit wissenschaftlichen Daten, Fakten, Diagrammen auseinandersetzen (vgl. auch Abschnitt 3 und Vohns, 2018): Einerseits befinden sie sich in einer *rezeptiven Rolle als Empfänger* statistischer Informationen, die richtig interpretiert werden müssen. In dieser Rolle sind Wissenschaftsjournalist:innen beispielsweise in der Phase der Recherche. Hierbei stellen sich etwa Fragen wie: Muss ich weitere Informationen recherchieren (z. B. die Basisrate bei medizinischen Tests; oder die Information ob Prozente oder Prozentpunkte kommuniziert wurden)? Liefert die vorliegenden Informationen wirklich schon ausreichende Antworten auf meine Fragen (z. B. reicht mir die Information, dass die Studie ein signifikantes Ergebnis zeigt, oder möchte ich lieber die Effekte zusätzlich deskriptiv darstellen?).

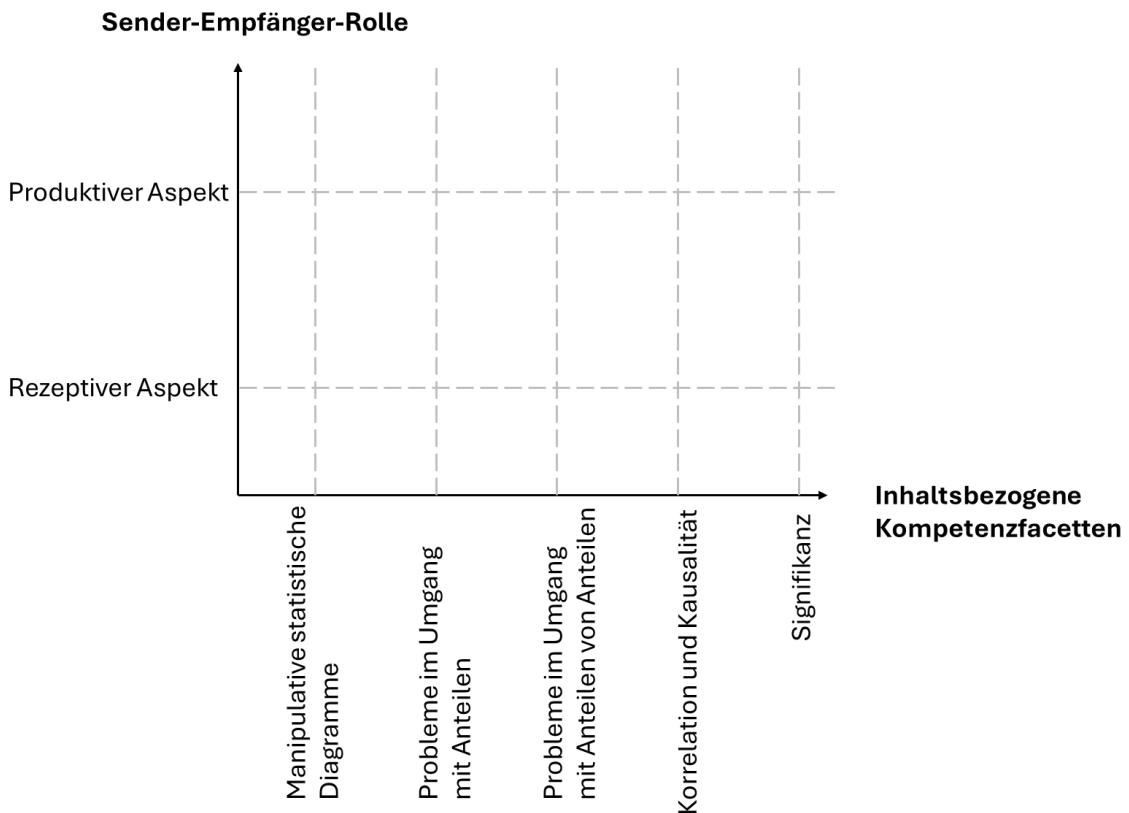


Abb. 4. Spezifische inhaltsbezogene Kompetenzfacetten der Data Literacy von Wissenschaftsjournalist:innen sowie die Unterscheidung der Sender-Empfänger-Rolle hinsichtlich eines rezeptiven und eines produktiven Aspekts

Andererseits befinden sich Wissenschaftsjournalist:innen aber auch in einer *produktiven Rolle als Sender* statistischer Informationen. Und in dieser Rolle kommen neue Aspekte hinzu, die die Darstellung von Daten betreffen. Eine stärkere Ausdifferenzierung dieser beiden Rollen sowohl im Kurskonzept als auch in den Messinstrumenten könnte gewinnbringend sein. Die beiden Rollen sind auch im Einklang mit den von Ickstadt, Müller & Weinert (2022) beschriebenen „ausgraben, verstehen, verkaufen“ als Kompetenzen, die Journalist:innen heute mitbringen müssen, um eine aktive Rolle in der Informationsvermittlung einzunehmen. Während hier „verstehen“ der Empfängerrolle zugeordnet werden kann, wäre „verkaufen“ dann entsprechend der Senderrolle zuzuordnen. Überdies wurde diese Unterscheidung der Rollen im Hinblick auf mathematische Bildungsziele auch wiederholt in der Mathematikdidaktik vorgeschlagen (siehe z. B. Vohns, 2018 oder Gal, 2002)

In Bezug auf eine Stärkung der Empfänger-Rolle wäre auch die Implementation weiterer Fortbildungsbauusteine denkbar – orthogonal zu jedem der fünf beschriebenen inhaltsbezogenen Kompetenzfazetten, wie beispielsweise typische Fehler und Schwierigkeiten bei der Interpretation von Pressemitteilungen (die oft von Universitäten und Hochschulen in Zusammenarbeit mit Wissenschaftler:innen formuliert werden) oder wissenschaftlichen Artikeln, beispielsweise nach dem aus der evidenzbasierten Medizin bekannten PICO-Prinzip (O’Sullivan, Wilk & Michalowski, 2013). Die Wissenschaftsjournalist:innen äußerten jedenfalls in der abschließenden Befragung, dass ihnen der konkrete Umgang mit Pressemitteilungen und wissenschaftlichen Artikeln ein wichtiger Aspekt ihrer täglichen Arbeit darstellt, der als besonders komplex wahrgenommen wird. Ein besonderes Augenmerk könnte dann auch auf dem Umgang mit widersprüchlichen Forschungsbefunden liegen und der Beurteilung der Qualität der jeweiligen wissenschaftlichen Erkenntnisse. Als weitere inhaltliche Bereiche können bei einer Weiterentwicklung des Kurskonzepts sowie im Prä- und Post-Test auch Themenbereiche wie Konfidenzintervalle und Effektstärken aufgenommen werden.

Während wir in der vorliegenden Studie mit Blick auf Wissenschaftsjournalist:innen einen Fokus auf die prozessbezogene Kompetenz „Daten kommunizieren“ (Ridsdale et al., 2015) gerichtet hatten, wäre zumindest für Datenjournalist:innen auch eine inhaltsbezogene Ausdifferenzierung der anderen vier Fazetten wünschenswert.

Die Ergebnisse der vorliegenden Studie bieten aufgrund der kleinen Gruppe teilnehmender Personen nur einen ersten Einblick darüber, ob die in der populärwissenschaftlichen Literatur beschriebenen Fehler tatsächlich auch Schwierigkeiten von Wissenschaftsjournalist:innen darstellen. Darüber hinaus soll ein erster Eindruck bezüglich der Schulbarkeit der verschiedenen inhaltsbezogenen Kompetenzfazetten in kurzen Fortbildungsbausteinen gewonnen werden. Der vorliegende Artikel versteht sich als erster Schritt einer sukzessiven forschungsbasierten Entwicklung von Fortbildungen, Materialien und Testinstrumenten, die die Ausbildung und Festigung einer journalistisch gefärbten Data Literacy zum Ziel hat. Weitere ähnliche Fortbildungskonzepte für Wissenschaftsjournalist:innen sind geplant und sollten auch auf Journalist:innen ausgedehnt werden. Dieses Anliegen begreift sich über das Forschungsinteresse hinaus als konkretes Maßnahmenbündel, mit denen Wissenschaftsjournalist:innen in ihrer wichtigen Tätigkeit für die Gesellschaft unterstützt werden sollen.

Danksagung

Wir danken den an der Fortbildung und Prä- und Post-Test teilnehmenden Wissenschaftsjournalist:innen für den regen und intensiven Austausch, von dem vermutlich beide Seiten gleichermaßen profitiert haben, der Wissenschaftspressekonferenz für die produktive Zusammenarbeit und den beiden anonymen Reviewern für die wertvollen Überarbeitungshinweise für den vorliegenden Artikel.

Literatur

- Aufermann, J. & Elitz, E. (Hrsg.). (2013). *Ausbildungsweg zum Journalismus: Bestandsaufnahmen, Kritik und Alternativen der Journalistenausbildung* (Vol. 18). Springer-Verlag.
- Batanero, C., Burrill, G., & Reading, C. (2011). *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study: the 18th ICMI study* (Vol. 14). Dordrecht: Springer.
- Bauer, T., Gigerenzer, G. & Krämer, W (2014). *Warum dick nicht doof macht und Genmais nicht tötet – über Risiken und Nebenwirkungen der Unstatistik*. Campus Verlag.
- Bauer, T., Gigerenzer, G., Krämer, W. & Schüller, K. (2022). *Grüne fahren SUV und Jogen macht unsterblich: Über Risiken und Nebenwirkungen der Unstatistik*. Campus Verlag.
- Bauer, T., Gigerenzer, G., Krämer, W. & Schüller, K. (o. D.) *Unstatistik des Monats*, <https://www.rwi-essen.de/presse/wissenschaftskommunikation/unstatistik> (abgerufen am 19.09.2024)
- Blöbaum, B. (2017). Wissenschaftsjournalismus. In H. Bonfadelli, B. Fähnrich, C. Lüthje, J. Milde, M. Rhomberg & M. Schäfer (Hrsg.), *Forschungsfeld Wissenschaftskommunikation* (S.

- 221-238). Springer. https://doi.org/10.1007/978-3-658-12898-2_12
- Bosbach, G., & Korff, J. J. (2011). *Lügen mit Zahlen: Wie wir mit Statistiken manipuliert werden*. Heyne Verlag.
- Boyles, J. L. & Meyer, E. (2016). Letting the Data Speak. *Digital Journalism*, 4(7), 944–954. <https://doi.org/10.1080/21670811.2016.1166063>
- Binder, K., Krauss, S. & Gigerenzer, G. (2020). Risikoveränderungen: Wie absolute und relative Veränderungen von Risiken mit Bildgittern unterrichtet werden können. *Mathematik lehren*, 220, 12-15.
- Binder, K., Krauss, S. & Krämer, W. (2019). Sonderbare Avocado-Vermehrung und kriminelles Frankfurt–Aktuelle statistische Fehler in den Medien unterrichtlich nutzen. *Stochastik in der Schule*, 39(2), 11-21.
- Bundesministerium für Ernährung und Landwirtschaft BMEL (2017). Deutschland wie es isst – Der BMEL-Ernährungsreport 2018. https://www.bmel.de/DE/Ernaehrung/_Texte/Ernaehrungsreport2018.html (abgerufen am 14. Mai 2018).
- Calzada Prado, J. & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123-134.
- Christensen, B. & Christensen, S. (2015). *Achtung: Statistik*. Springer.
- Cobb, G. W., & Moore, D. S. (1997). Mathematics, Statistics, and Teaching. *The American Mathematical Monthly*, 104(9), 801. <https://doi.org/10.2307/2975286>
- Data Literacy Charta (2021). *Data-Literacy-Charta*. www.data-literacy-charta.de
- Dirnagl, U. (2024). Ein signifikantes Problem. *Forschung & Lehre*, 4(24), 241.
- Dries, C., McDowell, M., Rebitschek, F. G. & Leuker, C. (2024). When evidence changes: Communicating uncertainty protects against a loss of trust. *Public Understanding of Science*, 09636625241228449.
- Engel, J. (2007). Daten im Mathematikunterricht: Wozu? Welche? Woher? [Data in mathematics education: What for? Which? Where from?]. *Der Mathe-matikunterricht*, 53(3), 12–22.
- Engel, J., & Ridgway, J. (2022). Back to the Future: Rethinking the Purpose and Nature of Statistics Education. In J. Ridgway (Hrsg.), *Statistics for empowerment and social engagement: Teaching civic statistics to develop informed citizens* (S. 17–36). Springer. https://doi.org/10.1007/978-3-031-20748-8_2
- Feufel, M. A., Antes, G. & Gigerenzer, G. (2010). Vom sicheren Umgang mit Unsicherheit: Was wir von der pandemischen Influenza (H1N1) 2009 lernen können. *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 53(12), 1283-1289.
- Fischer, R. (2012). Fächerorientierte Allgemeinbildung: Entscheidungskompetenz und Kommunikationsfähigkeit mit ExpertInnen. In R. Fischer, U. Greiner & H. Bastel (Hrsg.), *Domänen fächerorientierter Allgemeinbildung* (S. 9–17). Linz: Trauner
- Friedrich, A., Schreiter, S., Vogel, M., Becker-Genschow, S., Brünken, R., Kuhn, J., Lehmann, J. & Malone, S. (2024). What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies. *International Journal of STEM Education*, 11(1). <https://doi.org/10.1186/s40594-024-00517-z>
- Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review/Revue Internationale de Statistique*, 70(1), 1–25.
- Gigerenzer, G. (2014 [2013]). *Risiko. Wie man die richtigen Entscheidungen trifft* (6.Aufl.). Bertelsmann.
- Goodman, S. N. (2008). A dirty dozen: twelve P-value misconceptions. *Seminars in Hematology*, 45(3), 135–140.
- Guenther, L. (2017). Einleitung: Wissenschaftsjournalismus und wissenschaftliche Evidenz. *Evidenz und Medien: Journalistische Wahrnehmung und Darstellung wissenschaftlicher Un-sicherheit*, 11-39.
- Guenther, L., Froehlich, K., Milde, J., Heidecke, G. & Ruhrmann, G. (2015). Effects of valenced media frames of cancer diagnoses and therapies: Quantifying the transformation and establishing of evaluative schemas. *Health communication*, 30(11), 1055-1064.
- Gutounig, R., Radkohl, S., Goldgruber, E. & Stoiber, C. (2022). Datenjournalismus: Die Transformation journalistischer Arbeitsabläufe und Produkte durch Visualisierung und Analyse von Daten. In *Die digitale Transformation der Medien: Leit-medien im Wandel* (S. 325-345). Springer Fachmedien Wiesbaden.
- Haller, H. & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers. *Methods of psychological research*, 7(1), 1-20.
- Hermida, A., & Young, M. L. (2019). From peripheral to integral? A digital-born journalism not for profit in a time of crises. *Media and communication*, 7(4), 92-102.
- Hettwer, H. & Zotta, F. (2008). Eine Frage der Haltung: Oder: Warum Wissenschaftsjournalisten mehr sind als bloße Dolmetscher. *Gegenworte – Hefte für den Disput über Wissen*. https://edoc.bbaw.de/opus4-bbaw/frontdoor/deliver/index/docId/1321/file/20_Hettwer_Zotta.pdf
- Ickstadt, K., Müller, H. & Weinert, H. (2022). Data Literacy. In *Statistische Datenanalyse im Journalismus: Fallstudien und wissenschaftliche Anforderungen zum Einsatz fortgeschritten-er statistischer Methoden* (S. 29-42). Springer.
- Kline, R. B. (2013). *Beyond significance testing – statistics reform in the behavioral sciences*. Baltimore: United Book Press.
- Kohring, M. (1997). *Die Funktion des Wissenschaftsjournalismus. Ein systemtheoretischer Entwurf*. Opladen, 265.
- Krauss, S. & Wassner, C. (2001). Wie man das Testen von Hypothesen einführen sollte. *Stochastik in der Schule*, 21(1), 29–34.
- Krämer, W. (2015). *So lügt man mit Statistik*. Campus Verlag.
- Lakens, D., Scheel, A. M. & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in methods and practices in psychological science*, 1(2), 259-269.
- Loos, Andreas (2014). The never ending equal pay day (story). DMV-Blog, Eintrag vom 18. März 2014. <https://www.mathematik.de/dmv-blog/215-the-never-ending-equal-pay-day-story> (abgerufen am 14. Mai 2018).
- Lorenz, D. & Lorenz, D. (2009). Wege zum Journalismus. *Journalismus*, 174-179.
- Lück, J. & Boczek, K. (2020). Statistik in der akademischen Journalist*innen-Ausbildung. In J. Schützeneder, K. Meier, & N. Springer (Hrsg.), *Neujustierung der Journalistik/Journalismusforschung in der digitalen Gesellschaft: Proceedings zur Jahrestagung der Fachgruppe*

- Journalistik/Journalismusforschung der Deutschen Gesellschaft für Publizistik- und Kommunikationswissenschaft 2019, Eichstätt* (S. 105-116). Eichstätt: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e.V. <https://doi.org/10.21241/ssoar.70827>
- Mandinach, E. B. & Gummer, E. S. (2016). What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teaching and Teacher Education*, 60, 366-376.
- Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioral sciences*. Wiley.
- O'Sullivan, D., Wilk, S., Michalowski, W. & Farion, K. (2013). Using PICO to align medical evidence with MDs decision making models. In *MEDINFO 2013* (S. 1057-1057). IOS Press.
- Parasie, S. & Dagiral, E. (2013). Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. *New media & society*, 15(6), 853-871.
- Passon, O. & von der Twer, T. (2020). Evidenz, Signifikanz und das kleine p. *Zeitschrift für Bildungsforschung*, 10(3), 377-395.
- Ridgway, J. (2016). Implications of the data revolution for statistics education. *International Statistical Review*, 84(3), 528-549.
- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., Kelley, D., Matwin, S. & Wuetherick, B. (2015). *Strategies and best practices for data literacy education: Knowledge synthesis report*. Dalhousie University. <https://www.doi.org/10.1.3140/RG.2.1.1922.5044>
- Rolfes, T. (2021). Interpretation of Quantities Displayed in Pictorial Charts. *Frontiers in Psychology*, 12, 1-7. <https://doi.org/10.3389/fpsyg.2021.609027>
- Schicha, C. (2003). Medienethik und Medienqualität. *Zeitschrift für Kommunikationsökologie*, 2/2003, 44-53
- Schreiter, S., Friedrich, A., Fuhr, H., Malone, S., Brünken, R., Kuhn, J. & Vogel, M. (2023). Teaching for statistical and data literacy in K-12 STEM education: a systematic review on teacher variables, teacher education, and impacts on classroom practice. *ZDM – Mathematics Education*, 56, 31-45. <https://doi.org/10.1007/s11858-023-01531-1>
- Schüller, K., Busch, P. & Hindinger, C. (2019). *Future Skills: Ein Framework für Data Literacy – Kompetenzrahmen und Forschungsbericht*. Berlin: Hochschulforum Digitalisierung 47. DOI: 10.5281/zenodo.3349865
- Tagesschau (2022). Frauen verdienen 18 Prozent weniger. (7.3.2022). <https://www.tagesschau.de/wirtschaft/unternehmen/ungleiche-bezahlung-101.html> (abgerufen am 12. April 2024).
- van Merriënboer, J. J. G. (2020). Das Vier-Komponenten Instructional Design (4C/ID) Modell. *Handbuch Bildungstechnologie: Konzeption und Einsatz digitaler Lernumgebungen*, 153-170.
- van Merriënboer, J. J. G. & Kirschner, P. A. (2018). 4c/id in the Context of Instructional Design and the Learning Sciences. In F. Fischer (Hrsg.), *International handbook of the learning sciences* (S. 169-179). Routledge. [Https://doi.org/10.4324/9781315617572-17](https://doi.org/10.4324/9781315617572-17)
- van Merriënboer, J. J. G., Clark, R. E. & Croock, M. B. M. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational Technology Research and Development*, 50(2), 39-61. <https://doi.org/10.1007/BF02504993>
- Vohns, A. (2018). Mathematische Bildung am Ausgang ihrer Epoche? Eine nicht bloß rhetorisch gemeinte Frage. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, (105), 8-21.
- Weber-Stein, F. & Engel, J. (2021). Civic Statistical Literacy und politische Bildung im Informationszeitalter. Kooperative statistik- und politikdidaktische Erkundungen im Feld der Demokratiemessung. *Demokratie im Stresstest: Reaktionen von Politikdidaktik und politischer Bildung*, 165-192.
- Wegwarth, O. & Gigerenzer, G. (2011). Risiken und Unsicherheiten richtig verstehen lernen: Risikokommunikation. *Deutsches Ärzteblatt*, 108(9), 448-451
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96, 33-47. <https://doi.org/10.1007/s10649-017-9764-5>
- Weinacht, S. & Spiller, R. (2014). Datenjournalismus in Deutschland. *Publizistik*, 4(59), 411-433.
- Weischenberg, S., Malik, M. & Scholl, A. (2006). *Die Souffleure der Mediengesellschaft. Report über die Journalisten in Deutschland*. UVK.
- Wittmann, E. C. (1998). Mathematics education as a 'design science'. In *Mathematics Education as a Research Domain: A Search for Identity: An ICMI Study Book 1. An ICMI Study Book 2* (S. 87-103). Dordrecht: Springer Netherlands.
- Wolff, A., Gooch, D., Cavero, M., Jose J., Rashid, U. & Kortuem, G. (2017). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics*, 12(3), 9-26. <https://doi.org/10.15353/joci.v12i3.3275>

Anschrift der Verfasser:innen

Karin Binder
Ludwig-Maximilians-Universität München
Fakultät für Mathematik, Informatik und Statistik
Theresienstraße 39
80333 München

Seit 01.04.2025 neue Adresse:
Universität Paderborn
Institut für Mathematik
Warburger Straße 100
33098 Paderborn
Karin.Binder@uni-paderborn.de

Markus Vogel
Pädagogische Hochschule Heidelberg
Institut für Mathematik und Informatik
Im Neuenheimer Feld 561
69120 Heidelberg
vogel@ph-heidelberg.de