

Diagnose von Statistical Literacy bei Abiturient*innen

Die Leitidee „Daten und Zufall“ der KMK-Bildungsstandards zieht sich von der Primarstufe bis in die gymnasiale Oberstufe und bietet damit die Grundlage für den Erwerb stochastischer Allgemeinbildung von Abiturient*innen. Im angelsächsischen Raum wird dabei besonders die Rolle von Daten und Statistik zur Förderung von *Statistical Literacy* betont. Auf Deutschland übertragen, lassen sich unter *Statistical Literacy von Abiturient*innen* frei nach Gal (2002) und dem Allgemeinbildungskonzept von R. Fischer (2003) diejenigen Kenntnisse, Fertigkeiten, Fähigkeiten sowie Einstellungen und insbesondere Grund- und Reflexionswissen fassen, die benötigt werden, um

- statistische Informationen, datenbasierte Argumente und stochastische Phänomene, denen man in verschiedenen Kontexten begegnet, zu interpretieren und kritisch zu beurteilen sowie
- eigene Ansichten über die Auswirkungen sowie Bedenken bezüglich der Angemessenheit von Schlussfolgerungen gegenüber Lai*innen und Expert*innen zu kommunizieren und zu diskutieren.

Dabei ist es unstrittig, dass entsprechendes Wissen und Können in der heutigen datengesteuerten Gesellschaft von zentraler Bedeutung ist und daher in der Schule ausgebildet werden sollte (Schüller et al., 2021). Trotz der Relevanz von *Statistical Literacy* erscheinen geeignete Diagnoseinstrumente für Abiturient*innen rar. So gibt es zwar Testinstrumente aus verschiedenen Studien, allerdings adressieren diese häufig andere Zielgruppen und spezielle Lernvoraussetzungen (z. B. Studierende oder Schüler*innen der Sekundarstufe I). Hinzu kommt, dass ein Teil der bestehenden Tests sich eher für summative Evaluationen eignet und nicht immer handlungsweisende Ergebnisse für Lehrende und/oder Lernende liefert (bspw. Sawatzky, 2020, S. 55f).

Entwicklung eines Diagnosetests

Um diesem Desiderat nachzugehen, wurden Testaufgaben mit dem Ziel der Individualdiagnose von Wissen und Können im Hinblick auf *Statistical Literacy* bei Abiturient*innen entwickelt. Dabei wurde sich auf die kognitiven Komponenten dieses Konstrukts fokussiert. Das Vorgehen orientierte sich an den Konstruktionsphasen von Brandt & Moosbrugger (2020, Abb. 1).



Abb. 1: Konstruktionsphasen des Diagnosetests

In: L. Schick, M. Platz & A. Lambert (Hrsg.),
Beiträge zum Mathematikunterricht 2025.

58. Jahrestagung der Gesellschaft für Didaktik der Mathematik. WTM.

<https://doi.org/10.37626/GA9783959873307.0>

Bei der *Entwicklung eines ersten Entwurfs* der Diagnoseaufgaben erwies sich die nachfolgende Inhalt-Prozess-Matrix (Klauer & Leutner, 2012, S. 29–30) zur Ausdifferenzierung des Wissens und Könnens als hilfreich (Abb. 2). Die Zellen beschreiben dabei Komponenten des für *Statistical Literacy* relevanten Wissens und Könnens. Die Prozessdimension greift die beiden zentralen Fähigkeiten „Interpretieren“ und „Beurteilen“ aus der Arbeitsdefinition auf und bezieht sich auf die beiden höheren Kategorien des dreistufigen Modells von *Statistical Literacy* von Watson (1997). Für die Inhaltsdimension wurden zentrale statistische Konzepte mit Anknüpfungspunkten in den Bildungsplänen ausgewählt.

	<i>Prozesse</i>		
<i>Inhalte</i>		Interpretieren	Beurteilen
Mittelwerte & Streuung		<input type="checkbox"/>	<input type="checkbox"/>
Wahrscheinlichkeiten & Häufigkeiten		<input type="checkbox"/>	<input type="checkbox"/>
Darstellungen (Diagramme und Tabellen)		<input type="checkbox"/>	<input type="checkbox"/>
Statistische Datenerhebungen		<input type="checkbox"/>	<input type="checkbox"/>
Induktive Statistik (Hypothesentests / Konfidenzintervalle)		<input type="checkbox"/>	<input type="checkbox"/>

Abb. 2: Inhalt-Prozess-Matrix zur Ausdifferenzierung von *Statistical Literacy*

Bei der Testkonstruktion wurden einerseits zu den verschiedenen Komponenten der Inhalt-Prozess-Matrix passende Aufgaben neu entwickelt (siehe bspw. Abb. 3). Andererseits wurden geeignete Aufgaben aus bereits bestehenden Instrumenten zu *Statistical Literacy* und eng verwandten Konstrukten, die zu den Komponenten der Matrix passten, adaptiert. Dabei wurde darauf geachtet, dass die Aufgaben in möglichst realitätsnahen Kontexten mit persönlicher oder gesellschaftlicher Relevanz eingebettet sind. Infolgedessen wurde beispielsweise auf Aufgaben verzichtet, die im Kontext von Münz- und Würfelwürfen situiert sind. Außerdem wurden primär offene Aufgabenformate gewählt, die ein höheres Diagnosepotential als geschlossene Aufgaben bieten (vgl. u.a. Leuders, 2023, S. 503).

Die so entwickelte Aufgabensammlung wurde anschließend sukzessive revidiert. Zunächst fand zur Sicherstellung der inhaltlichen Validität der Aufgaben eine *Evaluierung durch Expert*innen* mit Wissenschaftler*innen aus den Bereichen Stochastik, Didaktik und Diagnose in Form eines Online-Ratings statt. Infolge einer quantitativen und qualitativen Auswertung der Einschätzungen wurden so einige Aufgaben eliminiert oder gründlich überarbeitet. Anschließend wurden im Rahmen *qualitativer Verständlichkeitsanalysen* Think-Aloud-Interviews durchgeführt, um mögliche Verständnisschwierigkeiten und nicht-intendierte Bearbeitungsstrategien aufzudecken und durch Änderungen der Aufgabenstellungen zu beheben. Nach einer *Pilotierung* in einem Mathematik-Leistungskurs und weiteren kleineren

Überarbeitungen fand im Frühjahr 2024 eine *Evaluationsstudie* an mehreren hessischen Gymnasien mit etwa 275 Schüler*innen sowie eine zusätzliche Erhebung in einer universitären Einführungsveranstaltung mit etwa 45 Studierenden des gymnasialen Lehramts mit mindestens einem MINT-Fach statt.

Beispielaufgabe und Auswertung

Nachfolgend ist eine Aufgabensequenz des Diagnosetests dargestellt (Abb. 3), deren Kontext auf einer tatsächlichen Veröffentlichung und häufig kursierenden Fehlinformationen basiert (siehe Kerschner, 2023). Sie zielt auf eine kritische Auseinandersetzung mit der Durchführung sowie den berichteten Ergebnissen einer Studie und passt damit zu den Anforderungen aus den Bildungsstandards, dass Abiturient*innen „exemplarisch statistische Erhebungen [...] beurteilen“ sollen (KMK, 2012, S. 21).

Chlordioxid gegen Corona?

In einer von der „Genesis Foundation“ finanzierten Studie wurde 2020 untersucht, ob die orale Einnahme des Bleich- und Desinfektionsmittels Chlordioxid (ClO_2) bei der Behandlung von COVID-19 unterstützen könnte. Hierzu nahmen 20 von 40 Patienten mit einer zeitgleichen COVID-19-Infektion täglich eine Lösung mit Chlordioxid ein. Die anderen 20 Patienten bildeten eine Vergleichsgruppe. Alle 40 Teilnehmenden wurden unter anderem zu Beginn der Studie sowie nach einer Woche mithilfe eines Fragebogens zu ihren Symptomen und Schmerzen befragt und auf COVID-19 mithilfe eines PCR-Tests getestet.

a) **Erläutern** Sie mehrere mögliche Faktoren, die die Aussagekraft dieser Studie einschränken könnten.

Im Ergebnisteil der Studie führen die Autoren an, dass eines der wichtigsten Ergebnisse sei, dass 100 Prozent der Patienten, die Chlordioxid einnahmen, nach einer Woche einen negativen PCR-Test aufwiesen.

b) **Formulieren** Sie mindestens eine kritische Nachfrage zu diesem Ergebnis der Studie.

Abb. 3: Chlordioxid-Aufgabe zum *Beurteilen* von *Statistischen Datenerhebungen*

Auf Grundlage der Vielzahl an Antworten aus den verschiedenen Erhebungen wurden für jede Aufgabe gestufte und teilweise mehrkategoriale Kodierschemata mittels einer qualitativen Inhaltsanalyse (Kuckartz & Rädiker, 2022) entwickelt. Diese geben einen Einblick in die Vielfalt an zu erwartenden Antworten und stellen einen Vorschlag zur Einordnung und Bewertung dieser dar. Bei der obigen Aufgabe erwies sich so beispielsweise eine gemeinsame Auswertung der Antworten zu Aufgabenteil a) und b) mit insgesamt sieben Kategorien als hilfreich. Hierbei zeigte sich u.a., dass nur ein geringer Bruchteil der Schüler*innen einen Einfluss der Auftraggebenden und Forschenden oder die kleine Stichprobengröße anmerkten, während deutlich mehr Schüler*innen mögliche verzerrende Gruppenunterschiede

problematisierten. Andere Aufgaben des Tests werden hingegen mithilfe einer einzelnen Skala ausgewertet (siehe hierzu Herzog & Krüger, 2025).

Ausblick

Neben den qualitativen sind quantitative Auswertungen in Vorbereitung, um Lernenden, Lehrenden und Forschenden eine grobe Orientierung zur *Statistical Literacy* von Abiturient*innen zu bieten. Ferner sind weitere aufgabenübergreifende sowie psychometrische Analysen geplant. Eine Herausforderung ist hierbei, die Antworten zu den unterschiedlichen Aufgaben sowohl inhaltlich als auch psychometrisch sinnvoll, für Lehrkräfte nachvollziehbar und handlungsleitend zusammenzufassen.

Literatur

- Brandt, H. & Moosbrugger, H. (2020). Planungsaspekte und Konstruktionsphasen von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Lehrbuch. Testtheorie und Fragebogenkonstruktion* (3. Auflage, S. 39–66). Springer.
- Fischer, R. (2003). Höhere Allgemeinbildung. In A. Fischer, A. Fischer-Buck, K. H. Schäfer & D. Zöllner (Hrsg.), *Situation—Ursprung der Bildung*, (S. 151–161). Universitätsverlag Leipzig.
- Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review*, 70(1), 1–25.
- Herzog, J. S. & Krüger, K. (2025, in Vorbereitung). Fehlvorstellungen zu bedingten Wahrscheinlichkeiten aus der Perspektive von Statistical Literacy. *Der Mathematikunterricht*, 71(4).
- Kerschner, B. (11. Dezember 2023). MMS & Chlordioxid: gefährliche Wundermittel. *Medizin transparent - Universität für Weiterbildung Krems*. <https://medizin-transparent.at/mms-das-gefaehrliche-wundermittel>
- Klauer, K. J. & Leutner, D. (2012). *Lehren und Lernen: Einführung in die Instruktionspsychologie*. (2. Auflage). Beltz.
- Kuckartz, U. & Rädiker, S. (2022). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung : Grundlagentexte Methoden* (5. Auflage). Grundlagentexte Methoden. Beltz Juventa.
- Kultusministerkonferenz. (2012). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.10.2012*.
- Leuders, T. (2023). Aufgaben in Forschung und Praxis. In R. Bruder, A. Büchter, H. Gasteiger, B. Schmidt-Thieme & H.-G. Weigand (Hrsg.), *Handbuch der Mathematikdidaktik* (2. Auflage, S. 493–521). Springer Spektrum.
- Sawatzky, A. (2020). *Messung der statistischen Kompetenz in der Hochschulausbildung am Beispiel des Statistical Reasoning Assessment* [Dissertation]. Universität zu Köln.
- Schüller, K., Koch, H. & Rampelt, F. (2021). *Data-Literacy-Charta*. Stifterverband. <https://www.stifterverband.org/charta-data-literacy>
- Watson, J. M. (1997). Assessing statistical thinking using the media. *The assessment challenge in statistics education*, 107–121