

Jannis Zeller¹
Josef Riese¹

¹RWTH Aachen University

Datenbasierte Fähigkeitsprofile im Physikdidaktischen Wissen

Theoretischer Hintergrund und Motivation

Die Untersuchung des Professionswissens angehender Lehrkräfte steht bereits seit einigen Jahren im Fokus fachdidaktischer Forschung (Kaiser, Bremerich-Vos & König, 2020). Basierend auf Arbeiten von Shulman (1986) sowie Baumert und Kunter (2006) wird das Professionswissen dabei zumeist in (mindestens) die drei Domänen *Fachwissen* (FW), *Pädagogisches Wissen* (PW) und *Fachdidaktisches Wissen* (FDW) unterteilt. Auch für die Physik hat sich eine solche Unterteilung durchgesetzt (Riese, 2009) und wurde in unterschiedlichen Studien als theoretische Grundlage verwendet (z. B. Kleickmann et al., 2014; Vogelsang et al., 2019). Empirische Untersuchungen in Quer- und Längsschnitten wiesen auf erwartungskonforme Zusammenhänge (z. B. Reinhold et al., 2017) sowie zeitliche Entwicklung (z. B. Sorge et al., 2019; Kulgemeyer et al., 2020; Riese et al., 2022) dieser Domänen des Professionswissens hin. Ergebnisse zur Entwicklung des FDW liegen dabei bislang hauptsächlich in Form von globalen quantitativen Scores vor, die auf Kovariaten wie Studienfortschritt bezogen werden. Kriterienorientierte Beschreibungen von Ausprägungen der Wissensdomänen von angehenden Physiklehrkräften existieren für das FW (Woitkowski & Riese, 2017) sowie das PW (König, 2009), liegen für das FDW bisher allerdings hauptsächlich in Form von Scale-Anchoring-Modellierungen vor (Schiering et al., 2019; Zeller et al., 2022). Für diese Analysen werden üblicherweise Item-Response-Modellierungen verwendet, die sowohl für die Aufgabenschwierigkeiten als auch für der Personenfähigkeiten eine strikte Hierarchie voraussetzen und stark auf das jeweils zugrunde liegende Testinstrument bezogen sind. Das in diesem Beitrag beschriebene Projekt zielt darauf ab, authentische Sprachproduktionen der Proband:innen mit menschlichem Expertenwissen in Form von Scores und Aufgabenanalysen im Sinne einer Computational Grounded Theory (Nelson, 2020) zu verbinden. Anders als bisherige strikt hierarchischen Ansätze wird dabei a priori keine Hierarchie von Aufgaben und Personen vorausgesetzt. Dazu werden in einem ersten Schritt in diesem Beitrag Testdaten aus dem ProfiLe-P+ Projekt (Vogelsang et al., 2019) einer Reanalyse unterzogen. Dabei werden mithilfe einer Aufgabenkategorisierung und der Bildung von Clustern prototypische Personengruppen durch unsupervised Machine Learning identifiziert.

Stichprobe und Design

Bei den vorliegenden Daten handelt es sich um $N_0 = 846$ Bearbeitungen des Testinstruments zur Erfassung des FDW in vier Facetten von Gramzow (2015) durch Studierende. Das Testinstrument wird dabei als Power Test interpretiert, sodass fehlende Antworten (Missings) mit 0 Punkten kodiert werden. Personen mit mehr als 50 % Missings insgesamt oder mit über 25 % aufeinanderfolgenden Missings am Ende des Testinstruments wurden dabei ausgeschlossen. So blieben insgesamt $N = 778$ Bearbeitungen für die Analyse. Die Aufgaben des Testinstruments wurden zudem einer Anforderungsanalyse angelehnt an die Anforderungsdimensionen nach Anderson und Krathwohl (2001) unterzogen, bei der

induktiv und deduktiv die neun Dimensionen *Fachwissen*, *Beispiel*, *Unterrichtssituation*, *Erinnern*, *Verstehen*, *Anwenden*, *Analysieren*, *Evaluieren* und *Kreieren* ermittelt wurden. Dabei kann eine Aufgabe mehreren Anforderungsdimensionen zugeordnet werden. Nach einer kommunikativen Validierung lag die Beurteilerübereinstimmung der Zuordnungen bei einem mittleren $\kappa = 0.58$, wobei sich die Differenzierung zwischen den Dimensionen *Anwenden* und *Analysieren* als am problematischsten erwies. Das Ziel dieser Analyse war es, die Dimensionalität des Datensatzes derart zu reduzieren und somit inhaltlich zu verdichten, dass eine Interpretierbarkeit (Nelson, 2020) der explorativen Analyse ermöglicht wird. Für diesen Zweck ist die erzielte Übereinstimmung insgesamt zufriedenstellend. Die Daten wurden anschließend personenweise aggregiert, sodass anschließend für jede:n Proband:in ein Score in jeder der neun Dimensionen vorlag.

Methodik

Zur Ermittlung des prototypischen Antwortverhaltens von Personenclustern wurde das k -Means Verfahren gewählt, da es ermöglicht, Zentrumsvektoren und somit typische Ausprägungen der Scores in den Anforderungsdimensionen für die Cluster zu extrahieren. Dieses Verfahren profitiert von möglichst gleichmäßig skalierten Daten, weshalb in einem Zwischenschritt die Scores in den Dimensionen auf den jeweils maximal erreichten Score (d. h. den Score der besten Person in der jeweiligen Dimension) normiert wurden. Mit diesem Vorgehen werden zudem Personenunterschiede am besten sichtbar gemacht. Um eine gute Abdeckung der Varianz bei gleichzeitig angemessener inhaltlicher Verdichtung zu erreichen, wurde ein Modell mit $k = 5$ Clustern gewählt und an die Daten angepasst.

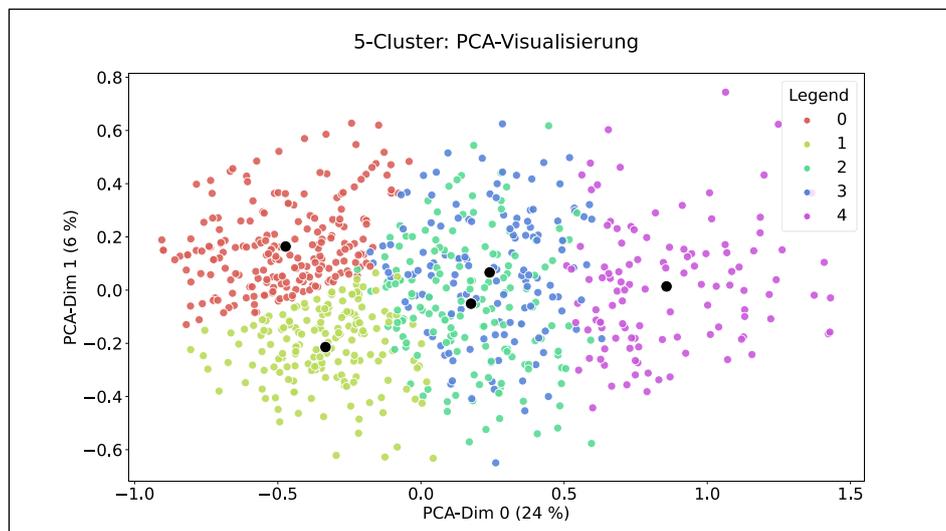


Abb. 1: Darstellung der 9-dim. Cluster in einem 2-dim. Unterraum, der durch PCA ermittelt wurde. Die PCA-Dimension auf der x-Achse klärt dabei 24 % und die PCA-Dimension auf der y-Achse klärt weitere 5 % der Varianz auf. Die schwarzen Punkte repräsentieren die Zentren der jeweiligen Cluster.

Die Daten wurden zu Visualisierungszwecken mithilfe einer Principal Component Analyse (PCA) vom ursprünglichen neundimensionalen Parameterraum auf einen zweidimensionalen Parameterraum projiziert (Abb. 1). Dabei korrespondiert die erste Dimension des reduzierten Parameterraums (x-Achse in Abb. 1) stark mit dem Gesamtscore bzw. der Personenfähigkeit einer Item-Response-Modellierung.

Ergebnisse und Ausblick

Das zentrale Ergebnis der Analyse sind die Ausprägungen der Dimensionen der Clusterzentren (Darstellung in einem Netzdiagramm in Abb. 2). Sie repräsentieren die Ausprägung der Scores in prototypischen Personengruppen. Cluster 0 (Fachsemester Physik: $M = 2.16$, $SD = 1.96$) und Cluster 4 (Fachsemester Physik: $M = 7.15$, $SD = 4.08$) stellen die Gruppen der eher leistungsschwachen und leistungsstarken Studierenden dar, wie man es bei Betrachtung des Studienfortschritts erwarten würde. Das Cluster 1 (Fachsemester Physik: $M = 3.32$, $SD = 2.67$) repräsentiert Studierende, die in den Dimensionen *Verstehen* und *Erinnern* bereits eine höhere Ausprägung erreicht haben, in den anderen Dimensionen aber noch nicht. Auffällig ist insbesondere die Ausprägung der Cluster 2 (Fachsemester Physik: $M = 5.02$, $SD = 3.23$) und 3 (Fachsemester Physik: $M = 5.23$, $SD = 3.77$) welche sich in Bezug auf Studienfortschritt und Gesamtscore kaum unterscheiden. Trotzdem scheint hier die Kompetenz, Elemente von Unterricht zu *kreieren*, zumindest für einen Teil der Studierenden nicht synchron zur Kompetenz, Elemente von Unterricht zu *evaluieren*, ausgebildet zu werden.

In anschließenden Analysen sollen die authentischen Sprachprodukte der Proband:innen mithilfe von Natural Language Processing und insbesondere Topic Models auf die gefundenen Cluster bezogen werden. Weiterhin werden zeitliche Entwicklung der Clusterzuordnung der Proband:innen sowie Zusammenhänge mit konkreten Lehrveranstaltungen untersucht.

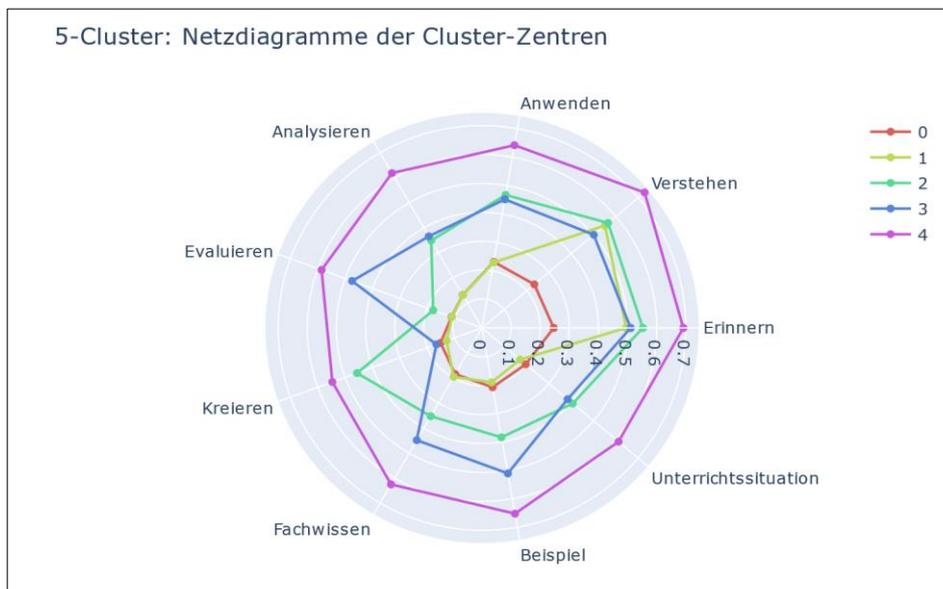


Abb. 2: Netzdiagramme der Clusterzentren.

Literatur

- Anderson, L. W. & Krathwohl, D. R. (Hrsg.). (2001). *A taxonomy for learning, teaching, and assessing A revision of Bloom's taxonomy of educational objectives* (4. Aufl.). New York: Longman.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Gramzow, Y. (2015). Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik: Modellierung und Testkonstruktion. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 181). Berlin: Logos Verlag.
- Kaiser, G., Bremerich-Vos, A., & König, J. (2020). Professionswissen. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 811–818). Bad Heilbrunn: Klinkhardt. <https://doi.org/10.35468/hblb2020-100>
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer-Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280–288.
- König, J. (2009). Zur Bildung von Kompetenzniveaus im Pädagogischen Wissen von Lehramtsstudierenden: Terminologie und Komplexität kognitiver Bearbeitungsprozesse als Anforderungsmerkmale von Testaufgaben? *Lehrerbildung auf dem Prüfstand*, 2(2), 244–262. <https://doi.org/10.25656/01:14703>
- Kulgemeyer, C., Borowski, A., Buschhüter, D., Enkrott, P., Kempin, M., Reinhold, P., Riese, J., Schecker, H., Schröder, J. & Vogelsang, C. (2020). Professional knowledge affects action-related skills: The development of preservice physics teachers' explaining skills during a field experience. *Journal of Research in Science Teaching*, 52(10), 1554–1582. <https://doi.org/10.1002/tea.21632>
- Nelson, L. K. (2020). Computational grounded theory: A methodological framework. *Sociological Methods & Research*, 49(1), 3–42. <https://doi.org/10.1177/0049124117729703>
- Reinhold, P., Riese, J. & Gramzow, Y. (2017). Fachdidaktisches Wissen im Lehramtsstudium Physik. In H. Fischler & E. Sumfleth (Hrsg.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik* (S. 39–56). Berlin: Logos Verlag.
- Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 97). Berlin: Logos Verlag.
- Riese, J., Vogelsang, C., Schröder, J., Borowski, A., Kulgemeyer, C., Reinhold, P. & Schecker, H. (2022). Entwicklung von Unterrichtsplanungsfähigkeit im Fach Physik: Welchen Einfluss hat Professionswissen? *Zeitschrift für Erziehungswissenschaft*, 25, 843–867. <https://doi.org/10.1007/s11618-022-01112-0>
- Schiering, D., Sorge, S., Petersen, S. & Neumann, K. (2019). Konstruktion eines qualitativen Niveaumodells im fachdidaktischen Wissen von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 25, 211–229. <https://doi.org/10.1007/s40573-019-00100-y>
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Sorge, S., Kröger, J., Petersen, S., & Neumann, K. (2019). Structure and development of pre-service physics teachers' professional knowledge. *International Journal of Science Education*, 41(7), 862–889. <https://doi.org/10.1080/09500693.2017.1346326>
- Vogelsang, C., Borowski, A., Buschhüter, D., Enkrott, P., Kempin, M., Kulgemeyer, C., Reinhold, P., Riese, J., Schecker, H. & Schröder, J. (2019). Entwicklung von Professionswissen und Unterrichtsperformanz im Lehramtsstudium Physik—Analysen zu valider Testwertinterpretation. *Zeitschrift für Pädagogik*, 65(4), 473–491. <https://doi.org/10.25656/01:23990>
- Woitkowski, D. & Riese, J. (2017). Kriterienorientierte Konstruktion eines Kompetenzniveaumodells im physikalischen Fachwissen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23, 39–52. <https://doi.org/10.1007/s40573-016-0054-z>
- Zeller, J., Jordans, M. & Riese, J. (2022). Ansätze zur Ermittlung von Kompetenzniveaus im Fachdidaktischen Wissen. In S. Habig (Hrsg.), *Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen, Tagungsband der GDCP Jahrestagung 2021*. Essen: Universität Duisburg-Essen.