

Facetten der mentalen Struktur beim Wechsel zwischen molekularen Repräsentationen

Theoretischer Hintergrund

Externe Repräsentationen spielen in der Chemie für das Verständnis und die Kommunikation chemischer Phänomene eine bedeutende Rolle (Gilbert, 2005; Kozma & Russell, 1997, 2005; Taskin, Bernholt & Parchmann, 2015; Wu & Shah, 2004, Mammino, 2008). Besonders die molekulare Ebene wird durch verschiedene Darstellungen wie Summen- und Strukturformeln oder Kugel-Stab-Modelle repräsentiert und in bestehendem Lehr- und Lernmaterial verwendet, um den Zugang zu nicht sichtbaren Prozessen zu ermöglichen (Hoffmann & Laszlo, 1991; Wu & Shah 2004). Chemische Formeln und Visualisierungen bilden zugleich einen Bestandteil der chemischen Fachsprache (Bernholt et al. 2012; Taskin & Bernholt, 2014, Hoffmann & Laszlo, 1991, Wu & Shah, 2004) und lassen sich aufgrund ihrer strukturellen Merkmale in drei Kategorien gliedern: symbolisch, hybrid und ikonisch (s. Grottke & Tiemann, 2020). Grundlage dieser 3-Gliederung bildet die Einteilung in symbolhafte und bildhafte Repräsentationen (Schnotz & Bannert, 2003; Schnotz, 2002) sowie die in diesem Projekt zu hybriden Darstellungen konkludierten Moleküldarstellungen mit symbolischen als auch ikonischen Merkmalen (Talanquer, 2011). Der Umgang mit diesen vielfältigen Repräsentationen stellt hohe Anforderungen an Lernende (Copolo & Hounshell, 1995; Fleischer, 2017; Kozma & Russell, 1997), wobei das in Beziehung setzen unterschiedlicher externer Repräsentationen und deren Überführung ineinander einen entscheidenden Bestandteil darstellt und als Translationsfähigkeit definiert wird (Kozma & Russell, 1997, 2005; Keig & Rubba, 1993). Bestehende Studien zeigen, dass Lernende Schwierigkeiten beim Übersetzen zwischen den verschiedenen Darstellungsweisen der molekularen Ebene aufwiesen. Dabei wurden die Übersetzungen zwischen chemischen Formeln, Elektronenkonfigurationen und Kugel-Stab-Modellen (Keig & Rubba, 1993; Furio et al. 2000), die Ermittlung der Moleküldarstellung aus Summenformeln (Furio et al. 2000) und die Strukturauflärung der Translationsfähigkeit bei Studierenden mit Fokus auf der Übersetzung von bildhaft in symbolisch (Fleischer, 2017) näher untersucht.

Forschungsinteresse und Forschungsfragen

Daraus entsteht das Forschungsdesiderat, die Translationsfähigkeit in einem quantitativen Setting für Lernende des schulischen Bereiches mit Betrachtung von Hin- und Rückübersetzungen sowie der Identifikation möglicher Einflussfaktoren zu untersuchen. Als Einflussgrößen auf die Lösungswahrscheinlichkeit bei Translationsaufgaben werden itemspezifische Aspekte, wie die Kombinationen aus Ausgangs- und Zielkategorien der externen Repräsentationen (Übersetzungswege) im chemischen Kontext sowie die in den Aufgaben verwendeten Stoffklassen und personenspezifische Eigenschaften betrachtet, welche zu folgenden Forschungsfragen führen:

FF1: Welche Itemcharakteristika (a: Übersetzungswege, b: Stoffklassen) zeigen einen empirischen Einfluss auf die Itemschwierigkeit?

FF2: Welche Zusammenhänge bestehen zwischen der Translationsfähigkeit im chemischen Kontext und ausgewählten personenspezifischen Eigenschaften?

Design und Methode

Zur Erfassung des latenten Konstruktes der Translationsfähigkeit bei Lernenden des schulischen Bereiches wurde ein technologiebasierter Multiple-Choice-Test konzipiert.

Dabei wurde aus der Kreuzung der drei Moleküldarstellungs-Kategorien eine 3x3 Untersuchungsmatrix aufgespannt (s. Grottko & Tiemann, 2020). Die Kombination aus einer Ausgangs- und einer Zielkategorie stellt je einen Übersetzungsweg dar, welcher durch 10 Aufgaben repräsentiert wird, sodass für die Hauptstudie 90 Multiple-Choice Aufgaben eingesetzt wurden. Die Aufgaben wurden auf Basis von Konstruktionsmanualen entwickelt und in einem Think-Aloud-Setting (N = 10) sowie in einer quantitativen Vorstudie validiert (N = 225). Aus testökonomischen Gründen wurden die Aufgaben in einem unvollständigen Testheftdesign angeordnet (YSDs nach Frey et al. 2009), sodass jeder Proband nur 18 Aufgaben aus dem Aufgabenpool bearbeitete. Die Datenerhebung erfolgte über das Umfragetool LimeSurvey mit 629 Schüler:innen der 10. und 11. Jahrgangsstufe Berliner Schulen. Zudem wurden personenbezogene Eigenschaften, wie Fachwissen oder fluide Intelligenz für die Bearbeitung der FF2 erhoben.

Der vorliegende Beitrag gibt einen Einblick in die Ergebnisse zur FF1, bei deren Analyse das Linear-Logistische Test-Modell (LLTM; Fischer, 1973) zu Grunde gelegt wurde. Im Gegensatz zu mehrdimensionalen Testmodellen, die mehrere latente Fähigkeiten der Personen postulieren, werden im LLTM-Modell für unterschiedliche Itemgruppen gemeinsame Verschiebungen auf der Logit-Skala modelliert. Die Items wurden dabei gemäß der zuvor aufgestellten Modellpläne (Abb. 1) mit Informationen zu ihren Itemeigenschaften kodiert (Zuordnung der Aufgabe zu einem Übersetzungsweg oder einer Stoffklasse). Die zuvor angesprochene 3x3 Untersuchungsmatrix bildet die Basis für die Analyse der Übersetzungswege (FF1a), wobei vier Modelle aufgestellt wurden. Das Stoffklassenmodell (FF1b) differenziert die in den Aufgaben verwendeten Stoffklassen.

Modell 1: neun Üw.	Modell 2: sechs Üw.	Modell 5: Stoffklassen T1 Alkan, Alken, Alkin, cycl. KW T2 Alkohol T3 Aldehyd T4 Keton T5 Sonstige T6 Carbonsäure, Fettsäure T7 Aminosäure T8 Ester T9 Ether
S-S S-H S-I	S-S S-H S-I	
H-S H-H H-I	H-S H-H H-I	
I-S I-H I-I	I-S I-H I-I	
Modell 3: zwei Üw.	Modell 4: keine Üw.	
S-S S-H S-I	S-S S-H S-I	
H-S H-H H-I	H-S H-H H-I	
I-S I-H I-I	I-S I-H I-I	

Abb. 1: Modellpläne
links (FF1a): Übersetzungswege (Üw.), rechts (FF1b): Stoffklassen

Umgesetzt wird das Modellfitting mittels Bayesianischer Item-Response-Theory über das R-Paket *brms* von Bürkner (2021). Dies ermöglichte eine flexible Modellspezifikation mit Fokus auf die Itemeigenschaften (Übersetzungsweg, Stoffklassen).

Ergebnisse

Der Modellvergleich der in *brms* (Bürkner, 2021) gefitteten Modelle erfolgt über das PSIS-LOO (Pareto-smoothed importance sampling cross-validation) Kriterium (Vehtari et al, 2017) mittels marginaler Log-Likelihood (Merkle et al., 2019) über eine externe Funktion (*marginal_loglik*) aus dem Paket *birtms* von Schäfer (2021), da in *brms* lediglich die konditionale Log-Likelihood für den Modellvergleich herangezogen wird.

Alle Modelle erfüllen die Kriterien der Modellkonvergenz (Rhat, MCMC-plots, ESS). Über alle Modelle hinweg ließ die 2pl-Skalierung im Vergleich zu der 1pl-Skalierung generalisierbarere Aussagen zu, sodass die Ergebnisse der 2pl-Skalierung präsentiert werden. Der Modellvergleich zeigt die LOO-ELPD-Differenzen der zu vergleichenden Modelle, wobei das Modell mit dem geringsten Wert jeweils die Referenz darstellt. Bei beiden Modellvergleichen (Tab. 1: Übersetzungswege, Stoffklassen) zeigen sich LOO-ELPD-

Differenzen < 4 , sodass alle Modelle eine ähnliche Vorhersagekraft zu haben scheinen und sich daher nicht unterscheiden lassen. Es können keine Unterschiede in der Lösungswahrscheinlichkeit bei den Translationsaufgaben in Abhängigkeit eines bestimmten Weges oder der Darstellungsart gefunden werden. Auch das Stoffklassen-Modell erweist sich nicht als bedeutsam prädiktiver. Sowohl Übersetzungsweg als auch die Stoffklasse des dargestellten Moleküls scheinen keinen, alle Kategorien betreffenden, Einfluss auf die Lösungswahrscheinlichkeit zu haben.

Tab. 1: Modellvergleiche über LOO-ELPD, PSIS, marginale Log-Likelihood

Modellvergleich: Übersetzungswege				Modellvergleich: Stoffklassen			
Nr.	Modell	elpd-diff	se-diff	Nr.	Modell	elpd-diff	se-diff
2	sechs Üw. (2pl)	0.0	0.0	5	Stoffklassen (2pl)	0.0	0.0
1	neun Üw. (2pl)	-0.3	2.2	2	sechs Üw. (2pl)	-2.8	2.6
3	zwei Üw. (2pl)	-0.8	2.4	1	neun Üw. (2pl)	-3.1	2.8
4	keine Üw. (2pl)	-0.9	2.7	3	zwei Üw. (2pl)	-3.6	2.4
-	-	-	-	4	keine Üw. (2pl)	-3.6	2.9

Anmerkung: Üw. = Übersetzungswege

Die detaillierte Analyse der Itemschwierigkeits-Parameterverteilungen der komplexesten Modelle (9-Üw.-Modell und Stoffklassen-Modell) zeigte aber für einzelne Kategorien einen modellierbaren Einfluss. Grundlage für diese Vergleiche bilden die 4000 Posterior-Samples der Itemgruppenparameter (akzeptierte Parameterschätzungen aus dem MCMC-Verfahren). In dem Modell 1 (9-Üw.) zeigt sich, dass der Übersetzungsweg hybrid in hybrid (H-H) als schwerer einzustufen ist als der Übersetzungsweg symbolisch in ikonisch (S-I). Die anderen Übersetzungswege weisen unter Annahme eines Credible-Intervalls von .95 keine systematische Abweichung voneinander auf. Somit zeigt sich nur bezüglich des Vergleichs der Übersetzungswege H-H und S-I ein über die Modellschätzungen konsistenter Unterschied bezüglich der mittleren Itemgruppenschwierigkeiten. Folglich zeigen sich auch keine systematischen Unterschiede (CI = .95) bei Hin- und Rückübersetzungen in den Itemschwierigkeiten. Im Stoffklassen-Modell lassen sich Unterschiede in der Lösungswahrscheinlichkeit bei den Translationsaufgaben zwischen mehreren Stoffklassen feststellen. Dabei zeigen sich systematische Unterschiede bei der Übersetzung von Alkanen, Alkenen und Alkinen (T1) im Vergleich zu anderen Stoffklassen, welche überwiegend schwerer zu übersetzen scheinen als T1. Ether-Verbindungen (T9) scheinen hingegen leichter in der Übersetzung zu sein als Ester- (T8), Carbonsäure- (T6) oder Aldehyd-Verbindungen (T3). Alkohol- (T2), Aldehyd- (T3), Keton- (T4), Carbonsäure- (T6) und Ester-Verbindungen (T8) weisen hingegen untereinander keine systematischen Unterschiede auf und sind demnach annähernd gleich schwer bzw. leicht zu übersetzen.

Deutung

Unter Annahme einer gemeinsamen latenten Translationsfähigkeit konnten empirisch keine Unterschiede bezüglich der Schwierigkeit in den Übersetzungswegen zwischen verschiedenen Moleküldarstellungen identifiziert werden (FF1a). Für die Überprüfung eines mehrdimensionalen Mechanismus des Übersetzens zwischen diesen bedürfte es einer Folgestudie, in der Schüler:innen mehr Fragen je Zelle bearbeiten. Erst dann können Aussagen über eine differenzierte kognitive Verarbeitung getroffen und die Repräsentationsunabhängigkeit geprüft werden. Die Unterschiede in der Lösungswahrscheinlichkeit bei Aufgaben verschiedener Stoffklassen könnte auf die curriculare Reihenfolge und somit Vertrautheit mit diesen Stoffklassen zurückgeführt werden, aber auch auf die Komplexität der Darstellungen aufgrund ihrer funktionellen Gruppe (FF1b).

Literatur

- Bernholt, S., Fischer, I., Heuer, S., Taskin, V., Martens, J., & Parchmann, I. (2012). Die chemische Formelsprache – (un-) vermeidbare Hürden auf dem Weg zu einer Verständnissentwicklung? *CHEMKON*, 19(4), 171-178. doi:10.1002/ckon.201210183
- Bürkner, P.-C. (2021). *Brms: Bayesian Regression Models using „stan“*. R package version 2.15.0. <https://CRAN.R-project.org/package=brms>
- Copolo, C. E., & Hounshell, P. B. (1995). Using Three-Dimensional Models to Teach Molecular Structures in High School Chemistry. *Journal of Science Education and Technology*, 4(4), 295-305. doi:10.1007/BF02211261
- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychol.* 37, 359–374. doi: 10.1016/0001-6918(73)90003-6
- Fleischer, T. (2017). *Untersuchung der chemischen Fachsprache unter besonderer Berücksichtigung chemischer Repräsentationen*. Berlin: Logos.
- Frey, A., Hartig, J., & Rupp, A. A. (2009). An NCME Instructional Module on Booklet Designs in Large-Scale Assessments of Student Achievement: Theory and Practice. *Educational Measurement: Issues and Practice*, 28(3), 39-53. doi:10.1111/j.1745-3992.2009.00154.x
- Furió, C., Calatayud, M. L., Bárcenas, S. L., & Padilla, O. M. (2000). Functional fixedness and functional reduction as common sense reasonings in chemical equilibrium and in geometry and polarity of molecules. *Science Education*, 84(5), 545-565. doi:10.1002/1098-237X(200009)84:5<545::AID-SCE1>3.0.CO;2-1
- Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization in Science Education* (pp. 9-27). Dordrecht: Springer.
- Grottko, T. & Tiemann, R. (2020). Repräsentationswechsel zwischen molekularen Darstellungsformen. In: S. Habig (Ed.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Wien 2019. (pp. 574-577). Universität Duisburg-Essen.
- Hoffmann, R., & Laszlo, P. (1991). Representation in Chemistry. *Angewandte Chemie International Edition in English*, 30(1), 1-16. doi:10.1002/anie.199100013
- Keig, P. F., & Rubba, P. A. (1993). Translation of Representations of the Structure of Matter and its Relationship to Reasoning, Gender, Spatial Reasoning, and Specific Prior Knowledge. *Journal of Research in Science Teaching*, 30(8), 883-903. doi:10.1002/tea.3660300807
- Kozma, R., & Russell, J. (1997). Multimedia and Understanding: Expert and Novice Responses to Different Representations of Chemical Phenomena. *Journal of Research in Science Teaching*, 34(9), 949-968. doi:10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U
- Kozma, R., & Russell, J. (2005). Students Becoming Chemists: Developing Representational Competence. In J. K. Gilbert (Ed.), *Visualization in Science Education* (pp. 121-145). Dordrecht: Springer.
- Mammino, L. (2008). Teaching Chemistry with and Without External Representations in Professional Environments with Limited Resources. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and Practice in Science Education* (Vol. 3, pp. 155-185). Dordrecht: Springer.
- Merkle, E. C., Furr, D., & Rabe-Hesketh, S. (2019). Bayesian comparison of latent variable models: Conditional versus marginal likelihoods. *Psychometrika*, 84(3), 802–829. <https://doi.org/10.1007/s11336-019-09679-0>
- Schäfer (2021). Marginal Likelihood. <https://birtms.netlify.app/post/marginal-likelihood/#fnref2> In: Schäfer (2021). *birtms: Bayesian item response theory modeling with visual model evaluation* <https://github.com/Famondir/birtms>
- Schnotz, W. (2002). Commentary: Towards an Integrated View of Learning from Text and Visual Displays. *Educational Psychology Review*, 14(1), 101-120. doi:10.1023/A:1013136727916
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and instruction*, 13(2), 141-156. doi:10.1016/S0959-4752(02)00017-8
- Taskin, V., & Bernholt, S. (2014). Students' Understanding of Chemical Formulae: A review of empirical research. *International Journal of Science Education*, 36(1), 157-185. doi:10.1080/09500693.2012.744492
- Taskin, V., Bernholt, S., & Parchmann, I. (2015). An inventory for measuring student teachers' knowledge of chemical representations: design, validation, and psychometric analysis. *Chemistry Education Research and Practice*, 16(3), 460-477. doi:10.1039/C4RP00214H
- Talanquer, V. (2011). Macro, Submicro, and Symbolic: The many faces of the chemistry “triplet”. *International Journal of Science Education*, 33(2), 179-195. doi:10.1080/09500690903386435
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Wu, H.-K., & Shah, P. (2004). Exploring Visuospatial Thinking in Chemistry Learning. *Science Education*, 88(3), 465-492. doi:10.1002/sce.10126