

Paul Unger¹
Karsten Rincke¹

¹Universität Regensburg

Vergleich hinführender und rückführender Strukturierungen im Physikunterricht

Die PISA-Studien zeigten, dass der in deutschen Schulen durchgeführte naturwissenschaftliche Unterricht im Vergleich zu Unterricht an finnischen Schulen unterlegen hinsichtlich der Leistung der Schüler/-innen wirkt (OECD, 2019; OECD, 2016; OECD, 2014). Im Projekt QuIP (Quality of Instruction in Physics) konnten mittels Videoanalysen länderspezifische Unterschiede unter anderem in der Strukturierung finnischen und deutschen Physikunterrichtes beobachtet werden (vgl. Geller et al., 2014). Dies zeichnet sich in der zeitlichen Gestaltung des Unterrichts ab: Bei der Zielsetzung des Erwerbes von Fachwissen wird in Deutschland viel Zeit auf eine hinführende Herleitung zum Fachinhalt verwendet, während an finnischen Schulen auf eine zeitintensive Herleitung zu Gunsten einer ausgeprägteren Konzeptisierung verzichtet wird (vgl. Wackermann, Trendel & Fischer, 2010; Geller, 2015 S. 110ff.).

Hinführende und rückführende Verknüpfung des Vorwissens

Deutscher Unterricht folgt typischerweise dem fragend-entwickelndem Unterrichtsstil. Ausgehend vom Vorwissen der Schüler/-innen wird dabei im Unterrichtsverlauf das neue Konzept hergeleitet. Dieses Vorgehen zur Verknüpfung des Vorwissens mit dem neuen Konzept wird im Folgenden als *hinführend* bezeichnet. Fragend-entwickelnden Unterricht lernwirksam durchzuführen ist eine anspruchsvolle und auch anstrengende Aufgabe für die Lehrkraft. Oft geht ein solcher Unterrichtsverlauf in ein lernunwirksames „Frage-Antwort-Spiel“ über (Grell & Grell, 1983). Alternativ kann das neue Konzept direkt und zu Beginn erklärt werden und anschließend die Verbindungen zum Vorwissen hervorgehoben werden. Die Verknüpfung mit dem Vorwissen geschieht *rückführend* (vgl. Unger & Rincke, 2022). Ziel dieses Promotionsvorhabens ist es, Erkenntnisse darüber zu gewinnen, ob die Verknüpfungsart des Vorwissens im Zuge hinführender oder rückführender Strukturierungen im Physikunterricht bei der Einführung neuer fachlicher Inhalte zu unterschiedlicher Lernleistung führt und sich Effekte bezüglich der Interessanztheit des Unterrichtes und der Überforderung der Lernenden zeigen.

Änderungen am Studiendesign und an den Erhebungsinstrumenten

Für einen Überblick zur Gestaltung der hinführend und rückführend strukturierten Unterrichtseinheiten wird auf den Tagungsbandbeitrag von Unger & Rincke (2022) verwiesen. Es wurde seither lediglich die kontextuelle Einbettung abgeändert: Anstatt das Konzept Transformator über den Aspekt der elektrischen Energieübertragung einzuführen (vgl. Muckenfuß, 2007), wurde der Kontext „Spannungsanpassung durch Netzteile“ als Einstieg in die Thematik gewählt.

Der in Unger & Rincke (2022) noch zur Abwägung gestellte Messzeitpunkt der kognitiven Belastung wurde fixiert: Die Erhebung wird mittels Items nach Thees et al. (2021) direkt nach der Hin- bzw. Rückführung durchgeführt. Die dadurch entstehende Unterbrechung des Unterrichtsverlaufes wird zugunsten einer eindeutigen Zuordnung möglicher Messunterschiede in

der kognitiven Belastung gerechtfertigt (Schmeck et al., 2015). Während dieser „Zwischen-erhebung“ wird als neue Variable die wahrgenommene Strukturierung des Unterrichtes¹ mittels Items nach Trepke, Seidel & Dalehefte (2003) erhoben. Verglichen zu Unger & Rincke (2022) ergaben sich nachfolgende Änderungen für die einleitende und abschließende Datenerhebung: Unter Berücksichtigung einer ökonomischen Datenerhebung werden auf den kognitive-Fähigkeiten Test sowie die Erhebung der Selbstwirksamkeitserwartung verzichtet. Testitems zu individuellem und situativem Interesse werden nach Habig (2017) anstatt Berger (2000) ausgewählt.

Ergebnisse der Pilotierung

Die Unterrichtseinheit wurde in vier 10ten Klassen als Wiederholungsstunde² erprobt. Vorwissenstest, Leistungstest und die Erhebung zur kognitiven Belastung wurden eingesetzt. Die Lernenden waren nach Einschätzung der jeweiligen Lehrkräfte innerhalb der Klassen bezüglich Geschlecht und Leistungsniveau heterogen verteilt. Die Stichprobe umfasst N=63 Lernende. Nach Aussortieren unbearbeiteter Evaluationsbögen konnten $N_{VW} = 62$ Evaluationsbögen für die Auswertung des Vorwissenstestes und $N_{CL} = 59$ für die Auswertung der Items zur kognitiven Belastung herangezogen werden. Im Folgenden werden die Ergebnisse zur Konstruktvalidität, Reliabilität und zur Trennschärfe des Vorwissenstestes und des Evaluationsbogens zur kognitiven Belastung gezeigt. Die Auswertung des Leistungstestes befindet sich derzeit noch in Arbeit.

Der Vorwissenstest umfasst elf Items mit jeweils vier Antwortmöglichkeiten. Zugrundeliegende Konstrukte sind *Induktion* (Items VW-1 bis VW-6) und *elektrische Geräte* (Items VW-7 bis VW-11). Der Evaluationsbogen zur kognitiven Belastung umfasst zwölf Items, die jeweils auf einer sechs-stufigen Ordinalskala bewertet wurden. Zugrundeliegende Konstrukte sind *intrinsische* (Items CL-1 bis CL-4), *extrinsische* (Items CL-5 bis CL-8) und *lernbezogene kognitive Belastung* (Items CL-9 bis CL-12) (vgl. Thees et al., 2021). Die Untersuchung der Konstruktvalidität beider Messinstrumente wird mithilfe einer explorativen Faktoranalyse durchgeführt. Diese wird ebenfalls zur systematischen Reduktion der Anzahl verwendeter Items herangezogen. Aus dem Scree-Plot ergeben sich für die erhobenen Messwerte des Vorwissenstests zwei³ Faktoren und für die Messwerte des Evaluationsbogens zur kognitiven Belastung drei Faktoren. Zur Abschätzung der Reliabilität der einzelnen Subskalen beider Evaluationsbögen wird das Ordinale-Alpha α_O genutzt⁴. Die Trennschärfe der Items innerhalb der Subskalen wird mit Berücksichtigung einer Part-Hole Korrektur abgeschätzt. *Tabelle 1*

¹ Nach Clausen (2002) ist die Wahrnehmung von Schüler/-innen über Unterricht insbesondere dann relevant, wenn Zusammenhänge bezüglich Leistung oder dem Interesse betrachtet werden.

² Im Schuljahr 2021/2022 war das Thema Transformator aufgrund des Lehrplanwechsels in Bayern kein Inhalt des Unterrichtsverlaufsplans aller Jahrgangsstufen. In den betrachteten Klassen wurde das Thema Transformator im vergangenen Schuljahr 2020/2021 unterrichtet.

³ Nach Kaiser-Guttman-Kriterium (KG) lassen sich für den Vorwissenstest bis zu vier Faktoren rechtfertigen. Bortz & Schuster (2010, S. 415ff.) empfehlen eine Abschätzung über den „Knick“ im Scree-Plot, da das Kaiser-Guttman-Kriterium zur Überschätzung bedeutsamer Faktoren neigt.

⁴ Nach Gadermann et al. (2012) eignet sich das Ordinale-Alpha deutlich besser zur Reliabilitätsabschätzung für ordinale Skalen als Cronbachs-Alpha.

fasst die Ergebnisse der Analyse zusammen. Faktorladungen werden über einem Wert von .40 dargestellt.

Tabelle 1: statistische Analyse der Evaluationsbögen Vorwissen und kognitive Belastung.

	Skala	Item	Faktorladungen	h^2	Trennschärfe	Reliabilität
Vorwissen	Induktion	VW-1	.67	.66	.44 < s < .72	$\alpha_0 = .69$
		VW-2	.81	.68	.33 < s < .68	
		VW-5	.60	.50	.28 < s < .77	
	El. Geräte	VW-8	.75	.67	.38 < s < .68	$\alpha_0 = .70$
		VW-9	.71	.66	.43 < s < .80	
		VW-10	.72	.62	.42 < s < .75	
Kognitive Belastung	intrinsische	CL-1	.80	.66	.76 < s < .89	$\alpha_0 = .91$
		CL-2	.89	.82	.80 < s < .92	
		CL-3	.85	.74	.78 < s < .90	
	extrinsische	CL-5	.89	.80	.70 < s < .90	$\alpha_0 = .83$
		CL-6	.49	.25	.46 < s < .69	
		CL-7	.63	.63	.65 < s < .85	
	lernbezogene	CL-10	.79	.74	.78 < s < .90	$\alpha_0 = .90$
		CL-11	.85	.64	.72 < s < .86	
		CL-12	.77	.74	.76 < s < .91	

Die Items VW-6, VW-7, VW-11, CL-4, CL-8 und CL-9 wurden im Zuge der explorativen Faktoranalyse aussortiert, die Items VW-3 und VW-4 konnten keinem Faktor zugeordnet werden. Die Faktorladungen sowie Kommunalitäten (h^2) legen nahe, dass die zugrundeliegenden Konstrukte abgebildet werden. Insbesondere kann eine Differenzierung der kognitiven Belastung in die drei Subskalen intrinsische, extrinsische und lernbezogene Belastung bekräftigt werden, was im Einklang mit den Ergebnissen von Thees et al. (2021) steht. Lediglich das Item CL-6 klärt nur 25% der errechneten Varianz innerhalb des Faktors auf und weist eine schwache Faktorladung auf. Während in den übrigen Items speziell nach physikalischen Konzepten, Darstellungsarten oder allgemeiner nach der Qualität der Erklärung gefragt wird, wird durch CL-6 explizit nach *sprachlichen* Unklarheiten gefragt. Mit einem Akzeptanzintervall für die Reliabilität von $.55 < \alpha < .95$ (vgl. Rost, 2013, S. 178f.) befinden sich die Ergebnisse im akzeptablen Bereich.

Nächste Schritte

Die in diesem Beitrag vorgestellten Items werden in die Erhebungsbögen der Hauptstudie übernommen. Das Item CL-6 wird trotz schlechter Kommunalität und Faktorladung beibehalten, da eine Einschätzung der sprachlichen Verständlichkeit auch ohne eine mögliche Verknüpfung zur extrinsischen kognitiven Belastung aufschlussreich wirkt. Werte zu Konstruktvalidität, Reliabilität und Trennschärfe der Erhebungitems zum Fach- und Sachinteresse sowie zur Wahrnehmung der Lernenden über die Struktur des Unterrichtes und über das Auftreten der Lehrkraft werden aus den Daten der Hauptstudie bestimmt und mit den Werten vergangener Studien verglichen, in denen die Items ebenfalls eingesetzt wurden (vgl. Habig, 2017; Maurer 2016; van Vorst, 2013; Fechner, 2009; Trepke, Seidel & Dalehefte, 2003). Eine inhaltliche und fachdidaktische Überarbeitung des Leistungstestes ist derzeit in Arbeit. Der Beginn der Datenerhebung zur Hauptstudie ist für November 2022 angesetzt.

Literatur

- Berger, R. (2000). *Moderne bildgebende Verfahren der medizinischen Diagnostik—Ein Weg zu interessanterem Physikunterricht*. Logos.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. (7. Aufl.). Springer.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Waxmann.
- Fechner, S. (2009). Effects of context-oriented learning on student interest and achievement in chemistry education. *Logos*.
- Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17(3), 13.
- Geller, C. (2015). *Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb: Eine Videostudie in Finnland, Deutschland und der Schweiz*. Logos.
- Geller, C., Neumann, K., Boone, W. J., & Fischer, H. E. (2014). What Makes the Finnish Different in Science? Assessing and Comparing Students' Science Learning in Three Countries. *International Journal of Science Education*, 36(18), 3042–3066. <https://doi.org/10.1080/09500693.2014.950185>
- Grell, J., & Grell, M. (1983). *Unterrichtsrezepte* (12 (2010)). Beltz Verlag. <https://nbn-resolving.org/urn:nbn:de:101:1-2014040413776>
- Habig, S. (2017). *Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren*. Logos Verlag Berlin GmbH.
- Maurer, C. (2016). *Strukturierung von Lehr-Lern-Sequenzen*. Logos.
- Muckenfuß, H. (2007). Vom Induktionsgesetz zum Transformator. *Unterrichtsgang zur Bedeutung des Transformators für die elektrische Energieübertragung*. In *Naturwissenschaften im Unterricht. Physik* (Bd. 18, Nummer 102, S. 25–29).
- OECD (2014). *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264208780-en>.
- OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/9789264266490-en>.
- OECD (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/5f07c754-en>.
- Rost, D. H. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien: Eine Einführung* (3., vollst. überarb. und erw. Aufl). Klinkhardt.
- Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2015). Measuring cognitive load with subjective rating scales during problem solving: Differences between immediate and delayed ratings. *Instructional Science*, 43(1), 93–114. <https://doi.org/10.1007/s11251-014-9328-3>
- Thees, M., Kapp, S., Altmeyer, K., Malone, S., Brünken, R., & Kuhn, J. (2021). Comparing Two Subjective Rating Scales Assessing Cognitive Load During Technology-Enhanced STEM Laboratory Courses. *Frontiers in Education*, 6, 705551. <https://doi.org/10.3389/educ.2021.705551>
- Treppe, C., Seidel, T., & Dalehefte, I. M. (2003). Zielorientierung im Physikunterricht. In T. Seidel, M. Prenzel, R. Duit, & M. Lehrke (Hrsg.), *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“* (S. 201–228). IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften.
- Unger, P., Karsten, R. (2022) Vergleich analysierender und erarbeitender Strukturierungen im Physikunterricht. In S. Habig (Hrsg.) *Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen* (S. 736-739). GDCP-Tagungsband
- Vorst, H. van. (2013). *Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie*. Logos.
- Wackermann, R., Trendel, G., & Fischer, H. E. (2010). Evaluation of a Theory of Instructional Sequences for Physics Instruction. *International Journal of Science Education*, 32(7), 963–985. <https://doi.org/10.1080/09500690902984792>