

Evaluationsbasierte Transfergestaltung einer digitalen Schulinnovation

Ausgangslage

Veränderung von Schule stellt eine große Herausforderung dar (z. B. Gräsel & Parchmann, 2004). Insbesondere die Selbstreferentialität und die operative Geschlossenheit des sozialen Systems können Transfer und damit Wandel behindern (Gräsel, Jäger & Willke, 2006; Luhmann & Bednarz, 2005). Gerade das Durchsetzen fertiger Innovationen im Sinne eines top-down-Transfers ist daher häufig langfristig wenig wirksam (Snyder et al., 1992).

Vor diesem Hintergrund wurden am Standort Landau der Rheinland-Pfälzischen Technischen Universität Kaiserslautern-Landau im Ansatz einer symbiotischen Transferstrategie (Fußangel, Schellenbach-Zell & Gräsel, 2008) virtuelle Labore mit curricularer Verortung erstellt (Neff et al., 2021). Die digitalen Lernumgebungen wurden zur Vor- und Nachbereitung eines Experimentiersettings zum Thema Gewässeranalytik eingesetzt und simultan evaluiert. Mit dem Ziel einer Förderung der Implementation in der Schule wurden sowohl Lehrende als auch Lernende hinsichtlich der Nutzung des entwickelten didaktischen Konzeptes untersucht und transferrelevante kognitiv-affektive Variablen erhoben (Neff, 2023, im Druck).

Als zentrales summatives Ergebnis seitens der Lernenden konnte konstatiert werden, dass die virtuellen Labore von diesen kognitiv effektiv genutzt werden können. Zudem wurden mittels Logdaten die Bearbeitungsverläufe der Lernenden detailliert betrachtet. Hier zeigte sich eine große Varianz hinsichtlich der Navigationsstrategien innerhalb der digitalen Lernumgebungen. Zu deren Aufklärung wurden in einer weiteren Studie gezielt kognitive und nutzerbezogene Bedingungsvariablen in den Blick genommen. Im Fokus stand dabei die Frage, inwiefern die kognitive Flexibilität der Lernenden, also deren Fähigkeit auf wechselnde Anforderungen und Darstellungsformen adäquat zu reagieren (Braem & Egner, 2018), deren Nutzung der Lernumgebung beeinflusst. Im Beitrag werden die Persönlichkeitsmerkmale der Probanden sowie die Evaluation der Lernumgebung vorgestellt.

Methode

Die Erhebung erfolgte im Juli 2022 mit 26 Schüler:innen ($\text{♀} = 14$, $\text{♂} = 10$) eines Gymnasiums im Alter von 16 – 18 Jahren ($MW = 17.12$, $SD = 0.45$, $Med. = 17$). Die Lernenden zeigen im Bewertungssystem 0 – 15 Punkte der Oberstufe in Rheinland-Pfalz (Mainzer Studienstufe) gute fachliche Voraussetzungen in Chemie ($MW = 9.54$, $SD = 2.88$, $Med. = 11$, $n = 13$) und in Biologie ($MW = 10.08$, $SD = 3.28$, $Med. = 10$, $n = 24$). Die Vorkenntnisse im Fach Physik sind hingegen lediglich ausreichend bis befriedigend ($MW = 6.64$, $SD = 3.47$, $Med. = 8$, $n = 11$).

Die Teilnehmenden wurden während der Durchführung des digitalen Anteils der Lerneinheit hinsichtlich der Usability der virtuellen Labore (Brooke, 1996) und ihrer kognitiven Flexibilität (Cognitive Flexibility Inventory, Dennis & Vander Wal, 2010) befragt. Hierzu wurden jeweils deutschsprachige Übersetzungen der Original-Instrumente eingesetzt, die Items wurden entsprechend der Vorgaben der International Test Commission (ITC, 2017) unter Konsensfindung zwischen drei qualifizierten Übersetzungen durch Experten selektiert. Weiterhin wurde zur Validierung der Ergebnisse des CFI eine psychometrische Testung mittels Trailmaking-Tests (TMT, Teile A & B) durchgeführt (siehe z. B. Bowie & Harvey, 2006 für eine Beschreibung der Testung). Abweichend von der anzustrebenden Zeiterfassung durch den/ die

Testleiter:in wurde im vorliegenden Fall eine eigenständige Zeiterfassung durch die Teilnehmenden umgesetzt. Hierzu nutzten die Lernenden die Stoppuhr-Funktion ihrer zur Durchführung der Lerneinheit ohnehin erforderlichen Tablets.

Zum Datenmanagement wurden die manuell erzeugten Ergebnisse des TMT digitalisiert und mit den bereits digital vorliegenden Fragebogendaten zusammengeführt. Aufgrund fehlender Angaben konnten einzelne Nutzende nicht eindeutig zugeordnet werden; diese wurden für die weiteren Analysen ausgeschlossen. Die Datenauswertung erfolgte mit dem Statistikprogramm RStudio (2022.12.0 Build 353, R Version 4.2.2). Nach Eliminierung der fehlenden Daten im Sinne einer Available Case Analysis wurden zunächst deskriptive Kenndaten der (Sub-)Skalen ermittelt. Weiterhin wurden im Rahmen einer Itemanalyse die interne Konsistenz (Cronbach's Alpha), korrigierte Itemtrennschärfen sowie die Itemschwierigkeit nach Dahl ermittelt. Weiterhin wurden die Instrumente mit Bezug zur kognitiven Flexibilität durch Zusammenhangsanalysen in Bezug gesetzt.

Ergebnisse

Die ermittelte System Usability ist mit 43.02 von 100 maximal möglichen Punkten wenig zufriedenstellend, in der Zuordnung nach Bangor, Kortum & Miller (2008) ist dieser Wert als mangelhaft beziehungsweise nicht akzeptabel zu bezeichnen.

Das siebenstufige Cognitive Flexibility Inventory zeigt einen Mittelwert von 4.36 ($SD = 0.65$, $Med. = 4.25$). Bei Aufteilung der Skala zeigt die Subskala „Alternativen“ eine mittlere Ausprägung von 4.62 ($SD = 1.01$, $Med. = 4.77$), die Subskala „Kontrolle“ weist marginal geringere Werte auf ($MW = 4.48$, $SD = 1.02$, $Med. = 4.5$). In der Einzelfallbetrachtung zeigen sich größere Schwankungsbreiten (Abbildung 1 zeigt die Varianz als Boxplotdarstellung).

Die Zeitmessungen zur Bearbeitung des Trailmaking-Tests wurden, getrennt nach Teil A und Teil B, als Scores ausgewertet. Tabelle 2 stellt die deskriptiven Kennwerte literaturbasierten Referenzwerten gegenüber (siehe auch Abbildung 2).

Tabelle 1: Itemkennwerte nach paarweisem Ausschluss fehlender Daten. Fett gedruckte Kennwerte verweisen auf möglicherweise problematische Eigenschaften der Skalen.

Konstrukte	r_{it} (>.3)	p_{Dahl} (20 < p < 80)	α (>.7)	n
CFI Gesamtskala (4 Items, Skalierung 1-7)	-0.04 – 0.78	34.67 – 74.00	.77	26
CFI Alternativen (20 Items, Skalierung 1-7)	0.26 – 0.71	50.00 – 74.00	.63	26
CFI Kontrolle (7 Items, Skalierung 1 – 7)	-0.24 – 0.16	34.67 – 66.67	.63	26
Usability (10 Items, Skalierung 1– 5)	.25* - .88	2.17 – 41.30	.86	23

* 1 Item mit $r_{it} = .25$, übrige $> .55$

Tabelle 2: Kenndaten des TMT in Relation zu Referenzwerten für 18 – 24-Jährige mit 10 – 15 Bildungsjahren nach Tombaugh (2004).

	Kennwerte der Stichprobe			Referenzwerte (Tombaugh, 2004)		
	MW	SD	Med.	MW	SD	Med.
TMT-A	23.75	7.74	21.74	22.93	6.87	21.70
TMT-B	53.98	15.13	55.22	48.97	12.69	47.00

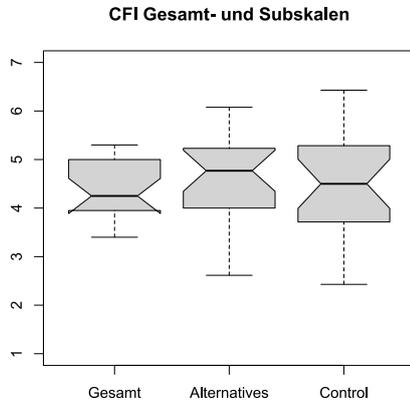


Abbildung 1: Boxplotdarstellung der Gesamt- und Subskalen des CFI.

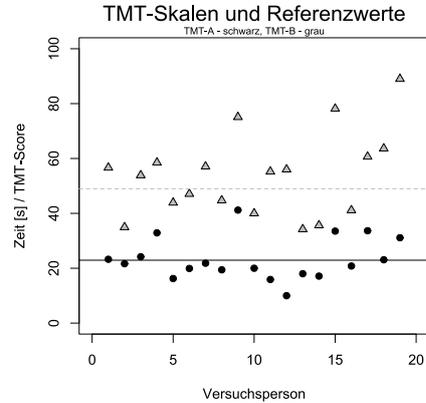


Abbildung 2: TMT-Scores und Literaturwerte.

Diskussion

Abweichend zu den Befunden der Erhebungen im Rahmen einer Studie unter 146 Lernenden (Neff, 2023) zeigt sich in der vorliegenden Auswertung eine Abweichung der Usability um rund 25 Punkte nach unten. Dieser Diskrepanz kann die Variation der genutzten Endgeräte zu Grunde liegen. Die vorliegend eingesetzten Tablets erlaubten trotz der oberflächlich vereinheitlichten, browserbasierten Darstellungsweise keine kontinuierliche Anzeige der Navigationsstruktur der Lernumgebung. Diese konnte lediglich bedarfsgerecht via Touchbefehl eingeblendet werden. Ähnlich abweichend stellten sich sogenannte Mouse-Over-Funktionen dar. Die mit Hilfe des CFI erhobene kognitive Flexibilität liegt in der oberen Hälfte des Skalenbereichs und ist somit gleichermaßen für Altersgruppe und Bildungsniveau erwartungskonform wie auch zufriedenstellend. Zusammenhangsanalysen konnten darüber hinaus starke negative Korrelationen der CFI-Scores mit den TMT aufdecken. Anhaltspunkte für eine Validität der eingesetzten Instrumente im vorliegenden spezifischen Anwendungsfall sind somit ersichtlich.

Im hier vorherrschenden Kontext erfolgt keine Interpretation der Daten der Trailmaking-Tests in Hinblick auf diagnostische oder psychopathologische Fragestellungen. Nach Mitrushina, Boone, Razani & D'Elia (2005) muss der demographische Hintergrund der Proband:innen im Zuge der Interpretation mitberücksichtigt werden, da die Performanz zumindest hinsichtlich der TMT in Abhängigkeit von Alter und Bildungsgrad stark variiert. Dahingehend stratifizierte Normdaten bietet beispielsweise Tombaugh (2004; eine Übersicht und Einordnung zahlreicher Normierungsstudien bieten Mitrushina, Boone, Razani & D'Elia, 2005). Im direkten Vergleich bleibt die Versuchsgruppe geringfügig hinter den Referenzwerten zurück. Aus methodischer Perspektive ist hierbei die, aus forschungspragmatischen Gründen implementierte, eigenständige Zeiterfassung der Proband:innen als mögliche Fehlerquelle zu identifizieren. Weitergehende Erkenntnisse auf Personenebene werden durch eine fokussierte Auswertung unter Triangulation der personenbezogenen kognitiven Flexibilität mit dem konkreten Verhalten der Nutzer:innen – abgebildet durch Logfiles und Screencast-Daten – zeitnah publiziert.

Literatur

- Bangor, A., Kortum, P. T. & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574-594.
- Bowie, C. R. & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature protocols*, 1(5), 2277-2281.
- Braem, S. & Egner, T. (2018). Getting a grip on cognitive flexibility. *Current directions in psychological science*, 27(6), 470-476.
- Brooke, J. (1996). SUS - A quick and dirty usability scale. *Usability evaluation in industry*, 189(194), 4-7.
- Dennis, J. P. & Vander Wal, J. S. (2010). The Cognitive Flexibility Inventory: Instrument Development and Estimates of Reliability and Validity. *Cognitive Therapy and Research*, 34(3), 241-253.
- Fußangel, K., Schellenbach-Zell, J. & Gräsel, C. (2008). Die Verbreitung von Chemie im Kontext: Entwicklung der symbiotischen Implementationsstrategie. In R. Demuth, C. Gräsel, I. Parchmann & B. Ralle (Hrsg.), *Chemie im Kontext: Von der Innovation zur nachhaltigen Verbreitung eines Unterrichtskonzepts* (S. 49-82). Waxmann.
- Gräsel, C. & Parchmann, I. (2004). Implementationsforschung - oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32(3), 196-214.
- Gräsel, C., Jäger, M. & Willke, H. (2006). Konzeption einer übergreifenden Transferforschung und Einbeziehung des internationalen Forschungsstandes: Expertise II zum Transferforschungsprogramm. In R. Nickolaus & C. Gräsel (Hrsg.), *Innovation und Transfer: Expertisen zur Transferforschung* (S. 445-566). Schneider-Verl. Hohengehren.
- International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests (Second Edition)*. www.InTestCom.org
- Luhmann, N. & Bednarz, J. (2005). *Social systems. Writing science*. Stanford Univ. Press.
- Mitrushina, M., Boone, K. B., Razani, J. & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2. ed.). Oxford Univ. Press.
- Neff, S. (2023). Transfer digitaler Innovationen in die Schulpraxis - Eine explorative Untersuchung zur Förderung der Implementation. Dissertation. In: M. Hopf & M. Ropohl (Hrsg.): Studien zum Physik- und Chemielernen. Logos Verlag Berlin.
- Neff, S., Gierl, K., Engl, A., Decker, B., Roth, T., Becker, J., Patzke, K., Winterholler, B., Kauertz, A. & Risch, B. (2021). Virtuelle Labore für den MINT-Unterricht - Transferprozess einer hochschulischen Innovation in den Schulkontext. In U. Schmidt & K. Schönheim (Hrsg.), *Transfer von Innovation und Wissen* (S. 75-101). Springer Fachmedien Wiesbaden.
- Snyder, J., Bolin, F. & Zumwalt, K. (1992). Curriculum Implementation. In P. W. Jackson (Hrsg.), *Handbook of research on curriculum: A project of the American Educational Research Association* (S. 699-703). Macmillan.
- Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 19(2), 203-214.