

Claudia Meinhardt
 Thorid Rabe
 Olaf Krey

Martin-Luther-Universität Halle-Wittenberg

Selbstwirksamkeitserwartungen (angehender) Physiklehrkräfte: Ausgewählte Ergebnisse einer Validierungsstudie

Einführung

Entwicklungs- als auch Validierungsschritte bzgl. des neuen Testinstrumentes zur Erhebung von Selbstwirksamkeitserwartungen (SWE) in physikdidaktischen Handlungsfeldern wurden bereits mehrfach auf GDCP-Jahrestagungen vorgestellt (Forschungsstand, Pilotstudie: Rabe, Krey & Meinhardt, 2013b; Skizze Gesamtprojekt: Rabe, Krey & Meinhardt, 2013a; Ergebnisse „qualitativer Validierungsschritte“: Meinhardt, Rabe & Krey, 2014; erste Ergebnisse „quantitativer Validierungsschritte“: Meinhardt, Rabe & Krey, 2015). Dieser Beitrag verzichtet daher auf eine eingehende theoretische Fundierung des Forschungsgegenstandes (Zur Definition des Konstruktes der SWE in physikdidaktischen Handlungsfeldern vergleiche z.B. Rabe, Meinhardt & Krey (2012)). Er fokussiert vielmehr auf eine Auswahl zentraler Ergebnisse der Hauptstudie. Für die Dimensionen „Planung“ und „Durchführung von Physikunterricht“ wurden jeweils Skalen in den Handlungsfeldern „Experimentieren“, „Umgang mit Schülervorstellungen“, „Umgang mit Aufgaben“ und „Elementarisieren“ entwickelt.

Forschungsfragen

Im Rahmen des Gesamtprojektes soll die übergeordnete Frage beantwortet werden, ob sich überhaupt Skalen auf dem gewählten Spezifitätsniveau konstruieren lassen, sodass für Messungen Validitäts- und Reliabilitätskriterien erfüllt sind. Die Ergebnisse jeder Teilstudie tragen zur Beantwortung dieser Frage bei, sodass am Ende des Projektes aus der Gesamtschau der Ergebnisse ein Validitätsargument abgeleitet werden kann (vgl. u.a. Kane, 1992; Messick, 1995). Die Ergebnisse der bisherigen Studien sind vielversprechend. Sie weisen darauf hin, dass mit den Skalen tatsächlich Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern gemessen werden können.

In der hier betrachteten Querschnittserhebung werden insbesondere Fragen nach der Eindimensionalität der Skalen, der Dimensionalität des Konstruktes, der Messinvarianz des Instruments und der Abgrenzung des Konstruktes zu benachbarten Konstrukten (allg. SWE, Lehrer-SWE, ...) betrachtet. Im Falle des Vorliegens von Messinvarianz bzgl. der drei interessierenden Gruppen (Physiklehramtsstudierende, Physikreferendare, Physiklehrkräfte) sollen aus exploratorischen Varianzanalysen Hinweise hinsichtlich der Interpretierbarkeit der Mittelwertstruktur gewonnen werden. Erste Ergebnisse konfirmatorischer Faktorenanalysen sprechen für die Eindimensionalität und Reliabilität der Skalen (Meinhardt et al., 2015).

Hypothesen

Zur Messinvarianz: Aufgrund der Eindrücke und Ergebnisse der Interviews mit Vertretern der drei Befragungsgruppen (vgl. Meinhardt et al., 2014) ist davon auszugehen, dass sich das Instrument als invariant bzgl. dieser Gruppen erweist.

Zur Mittelwertstruktur: Aufgrund des vielfach in der Literatur beschriebenen „Praxischocks“ (Abnahme der SWE nach Praxiserfahrungen, vgl. z.B. Woolfolk Hoy & Spero, 2005) wird davon ausgegangen, dass die Mittelwerte der Referendare auf den Skalen im Schnitt niedriger sind als die der Studierenden. Da im Sinne einer professionellen Entwicklung ein Anstieg der SWE plausibel ist, wird angenommen, dass die Mittelwerte der Skalen der Lehrkräfte auf allen Skalen höher sind als die der Studierenden bzw. Referendare.

Zur Abgrenzung der Konstrukte: Aufgrund der Spezifität des Instrumentes ist davon auszugehen, dass die Skalen lediglich im mittleren Bereich mit Skalen, die ein allgemeineres SWE-Niveau abbilden, korrelieren. Aufgrund der größeren inhaltlichen Nähe der Skalen zur Lehrer-SWE von Schmitz & Schwarzer (2000) wird eine höhere Korrelation zu dieser Skala angenommen als zur Skala bzgl. allgemeiner SWE (Schwarzer & Jerusalem, 1999).

Methode und Stichprobe

Basierend auf einer Analyse der Mittelwert- und Kovarianzstruktur des Datensatzes (MACS-Ansatz, vgl. z.B. Byrne, 2012) wird im Rahmen eines Mehrgruppenvergleichs die Messinvarianz des Instrumentes getestet. Dazu wird die Software Mplus (Muthén & Muthén, 2012) genutzt und auf ein robustes Maximum-Likelihood-Schätzverfahren zurückgegriffen. Die genesteten Modelle werden mittels χ^2 -Differenzentests (Satorra-Bentler skalierter χ^2 -Wert) verglichen. Im Falle der positiven Testung auf metrische Invarianz können Mittelwertdifferenzen zu einer gewählten Referenzgruppe (deren Mittelwert auf Null fixiert ist) ausgelesen und somit ggf. existierende Gruppenunterschiede identifiziert werden. Da die Skalenwerte nicht normalverteilt und streng genommen ordinalskaliert sind, wird für die Zusammenhangsanalysen der etwas robustere Koeffizient Spearman's ρ berichtet. Insgesamt wurden rund 900 Personen befragt. Für eine detailliertere Beschreibung der Stichprobe siehe Meinhardt et al. (2015).

Ergebnisse

Wie exemplarisch für die Skala SWE-SV-D (Handlungsfeld: Umgang mit Schülervorstellungen, Dimension: Durchführung von Physikunterricht) gezeigt (vgl. Tab. 1), kann für jede Skala mindestens partielle Messinvarianz konstatiert werden. Die Mittelwerte der Referendare sind für vier Skalen signifikant niedriger als die Mittelwerte der Studierenden. Die Mittelwerte der Lehrkräfte sind für jede Skala signifikant größer als die Mittelwerte der Studierenden (vgl. Abb.1). Die Mittelwertdifferenzen zwischen Referendaren und Lehrkräften sind ebenfalls signifikant. Die Skalen der allg. SWE und der Lehrer-SWE korrelieren im mittleren Bereich ($\rho=.47^{**}$). Ebenfalls im mittleren Bereich korrelieren die neu entwickelten SWE-Skalen mit der Lehrer-SWE-Skala ($.44^{**} \leq \rho \leq .48^{**}$). Die Korrelationen der neu entwickelten Skalen zur allg. SWE-Skala sind jeweils geringer ($.23^{**} \leq \rho \leq .38^{**}$). Die Angaben in Klammern beziehen sich auf die Studierendenkohorte.

Modell	#	χ^2	df	p	CFI	RMSEA	SRMR	SB- χ^2	Δ df	p	$ \Delta$ CFI	
konfigural	1	52.67	27	.00	.977	.055	.030					
metrisch	2	66.94	39	.00	.975	.048	.080	1 vs. 2	13.93	12	.31	.002
skalar	3	97.20	49	.00	.957	.056	.095	2 vs. 3	32.44	10	.00	.018
frei [svd2]*	4	77.08	47	.00	.973	.045	.078	3 vs. 4	24.42	2	.00	.016
								4 vs. 1	23.48	20	.27	.004

Tab. 1: Tests auf Messinvarianz für die Skala SWE-SV-D sowie Teststatistik der χ^2 -Differenzentests; * Intercept des Items svd2 frei geschätzt

Diskussion

Die Analysen bestätigen im Wesentlichen die Hypothesen. Es kann davon ausgegangen werden, dass für die untersuchten Gruppen Ausprägungen des gleichen Konstruktes erhoben werden konnten. Die Mittelwertstruktur ist mit Blick auf die Literatur durch den sogenannten Praxisschock interpretierbar. Darüber hinaus scheinen die durch die Skalen repräsentierten Konstrukte hinreichend zu benachbarten Konstrukten abgrenzbar zu sein. Mit Bezug zur übergeordneten Forschungsfrage können somit weitere Validitätsargumente angeführt werden.

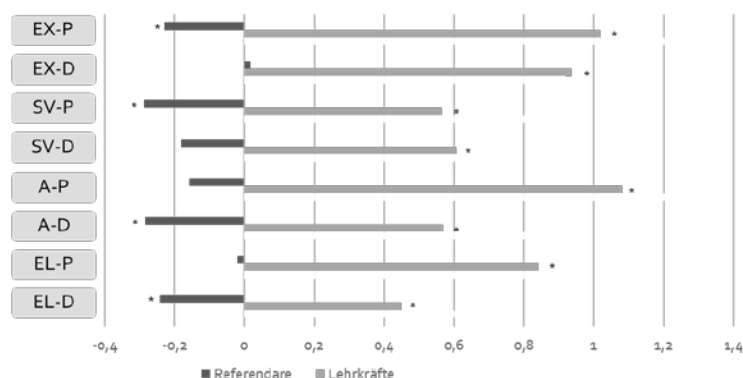


Abb. 1: Mittelwertdifferenzen der Referendare bzw. Lehrkräfte im Vergleich zum Referenzgruppenmittelwert der Studierenden ($M=0$), * signifikante Differenz

Ausblick

Detailliertere Analysen zur Messinvarianz z.B für verschiedene Subgruppen sowie Analysen der Zusammenhangs- und Mittelwertstruktur werden im Rahmen des Promotionsprojektes durchgeführt. Ein umfassendes Skalenhandbuch wird veröffentlicht werden.

Literatur

- Byrne, B. M. (2012). *Structural Equation Modeling with Mplus. Basic Concepts, Applications, and Programming*. New York & London: Routledge.
- Kane, M. T. (1992). An Argument-Based Approach to Validity. *Quantitative Methods in Psychology*, 112(3), 527–535.
- Meinhardt, C., Rabe, T. & Krey, O. (2014). Qualitative Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen (SWE) in physikdidaktischen Handlungsfeldern. In S. Bernholt (Ed.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in München 2013* (S. 558–560). Kiel: IPN.
- Meinhardt, C., Rabe, T. & Krey, O. (2015). Quantitative Validierung eines Testinstruments zu Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern – Erste Ergebnisse. In S. Bernholt (Ed.), *Heterogenität & Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014* (S. 283–285). Kiel: IPN.
- Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, 50(9), 741–749.
- Muthén, L. K. & Muthén, B. O. (2012). *Mplus. User's Guide* (7th Ed.). Los Angeles, CA: Muthén & Muthén.
- Rabe, T., Krey, O. & Meinhardt, C. (2013a). Physikdidaktische Selbstwirksamkeitserwartungen (zukünftiger) Physiklehrkräfte II – eine Projektskizze. In S. Bernholt (Ed.), *Inquiry-based Learning - Forschendes Lernen. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Hannover 2012* (S. 638–640). Münster: LIT.
- Rabe, T., Krey, O. & Meinhardt, C. (2013b). Physikdidaktische Selbstwirksamkeitserwartungen zukünftiger Physiklehrkräfte I. In S. Bernholt (Ed.), *Inquiry-based Learning - Forschendes Lernen. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Hannover 2012* (S. 635–637). Münster: LIT.
- Rabe, T., Meinhardt, C. & Krey, O. (2012). Entwicklung eines Instruments zur Erhebung von Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern. *Zeitschrift Für Didaktik der Naturwissenschaften*, 18, 293–315.
- Schmitz, G. S. & Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnittbefunde mit einem neuen Instrument. *Zeitschrift für Pädagogische Psychologie*, 14(1), 12–25.
- Schwarzer, R., & Jerusalem, M. (1999). Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Berlin.
- Woolfolk Hoy, A. & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, 21(4), 343–356.