

Abstract

Motivationsanalysen in einem between-subjects ATI-Forschungsdesign

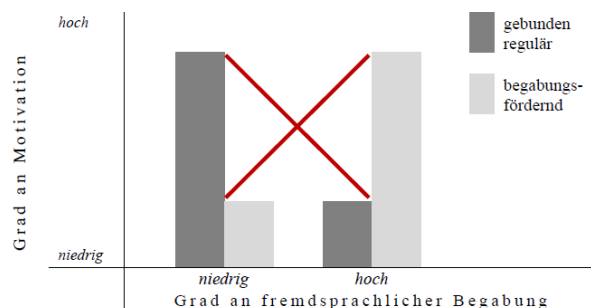
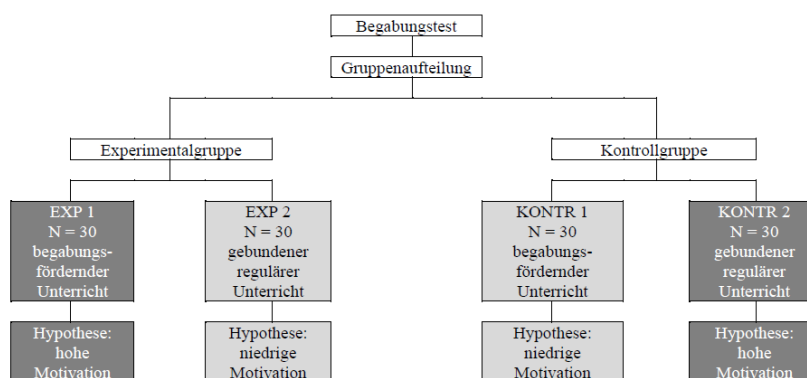
Thomas Wagner, PH ÖÖ

In dieser Präsentation geht es um ungelöste Fragestellungen hinsichtlich Messanordnung und Datenanalyse in einem quantitativen Forschungsdesign. Im Projekt FLAME-1 (*Foreign Language Aptitude & Motivation in English at Primary Schools, 2021-2023*) der drei Pädagogischen Hochschulen des Clusters Mitte geht es darum zu zeigen, ob vermeintlich begabungsfördernde schulische Interventionen fremdsprachlich begabte Schüler*innen der 4. Schulstufe tatsächlich mehr motivieren als durchschnittlich Begabte (Carreira, 2006; Schlak, 2008; Tellier & Roehr-Brackin, 2013). Hintergrund dieser Fragestellung ist die derzeitige Produktion von Lehrmaterialien mit begabungsförderndem Anspruch sowohl im Primar- als auch Sekundarstufenbereich trotz mangelnder Evidenz für deren motivationale Wirksamkeit.

Das Projekt folgt einem gemischten Design, mit einem quantitativen und einem qualitativen Teil. Letzterer bedient sich der Videographie und teilgesteuerten Leitfadeninterviews mit ausgewählten Schüler*innen und Lehrpersonen. Der quantitative Teil folgt voraussichtlich einem between-subjects aptitude-treatment-interaction Design, welches sich in der fremdsprachlichen Begabungsforschung nach wie vor großer Beliebtheit erfreut (Granena & Yilmaz, 2019; Yeh, 2012; <https://benjamins.com/catalog/jsls.2.2>). In diesem Design sind zwei Interventionskonditionen vorgesehen, ein gezielt begabungsfördernder Unterricht in Form einer Ressourcenecke / Lerninsel als Experimentalkondition sowie ein gebundener, "klassischer" Unterricht als Kontrollkondition. Das Sample wird anhand des Status hinsichtlich der fremdsprachlichen Begabung in durchschnittlich und überdurchschnittlich begabte Schüler*innen kategoriell aufgeteilt, jede Gruppe mit zwei Subgruppen für die Interaktion mit dem treatment; die Zuteilung in die jeweiligen Subgruppen erfolgt randomisiert. Messgröße im quantitativen Bereich des Projekts ist der Grad an Motivation (Selbstregulierung) nach der jeweiligen Intervention, erhoben mittels der für den Volksschulbereich adaptierten Selbstbestimmungsskala SRQ-A (*Fragebogen zum Lernen im Fach*, ursprünglich *Academic Self-Regulation Questionnaire*, mit den Subskalen extern, introjiziert, identifiziert, und intrinsisch, Müller et al., 2007; Ryan & Deci, 2000). Abbildung 1 illustriert dieses Design.

Abbildung 1.

ATI Forschungsdesign



Im Idealfall lässt sich aus den quantitativen Daten eine disordinale Interaktion, also die gekreuzt komplementären Motivationswerte wie in Abbildung 1 herausarbeiten, was, zusammen mit kongruierenden Daten aus dem qualitativen Bereich, die gewünschte empirische Evidenz im Sinne des ATI darstellen würde. Folgende Fragen stellen sich bezüglich des Designs und der Analyse der quantitativen Daten in diesem Projekt:

- (a) Braucht es ein crossed within-subject Design, um die Vergleichbarkeit der Gruppen zu erhöhen („eigene Kontrollgruppe“)?
- (b) Braucht es eine Baseline-Messung von Motivation in Form eines Pretests (Kovariate) und kann man bedenkenlos dieselbe Motivationskala für Messwiederholungen heranziehen?
- (c) Wie kann ein sinnvoller Grenzwert bzw. eine Grobklassifizierung des Summen-Scores des Begabungstests a priori konzipiert bzw. aus den Daten abgeleitet werden, sodass ein binäre kategorielle Einteilung des Samples möglich wird?
- (d) Welche Berechnungsmodelle sind sinnvoll für eine Poweranalyse für eine 2 x 2 ANOVA?
- (e) Welchen Skalenstatus schreibt man aktuell Likert-Skalen zu, und wie verbreitet und akzeptabel sind parametrische Verfahren zu deren Analyse (Agresti, 2002; Bidgood et al., 2021; Carifio & Perla, 2007; Gibson et al., 2011; Jamieson, 2004; Norman, 2010; Sprouse & Almeida, 2017)?
- (f) Gibt es Alternativen bei der Datenanalyse zum Mittelwertvergleich nach Gruppe (ANOVA) und den damit verbundenen rigiden Anforderungen an die Datenlage?
 - a. Lineare gemischte Modelle mit subject und item als random effects (Cunnings, 2012; Gibson et al., 2011; Kizach, 2014; Norman, 2010)?
 - b. Sollten die verschiedenen Schulen als random blocking variable modelliert werden?
 - c. Eignen sich generalisierte kumulative gemischte Modelle mit subject und item als random effects (Proportional Odds Model)?

Literatur

- Agresti, A. (2002). *Categorical data analysis*. John Wiley & Sons.
- Bidgood, A., Pine, J., Rowland, C., Sala, G., Freudenthal, D. T., & Ambridge, B. (2021). Verb argument structure overgeneralisations for the English intransitive and transitive constructions: Grammaticality judgments, production priming and a meta-analytic synthesis. *Language and Cognition*, Online, August 2021.
- Carifio, J. & Perla, R. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 2, 106-116.
- Carreira, J. (2006). Motivation for learning English as a foreign language in Japanese elementary schools. *Journal of the Japan Association for Language Teaching*, 28(2), 135-157.
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369-382.
- Gibson, E., Piantadosi, S., & Fedorenko, K. (2011). Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass*, 5(8), 509-524.
- Granena, G., & Yilmaz, Y. (2019). Aptitude-treatment interaction in L2 learning: A research synthesis. *Studies in English Education*, 23, 803-830.
- Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical Education*, 38, 1212-1218.
- Kizach, J. (2014). Analyzing Likert-scale data with mixed-effects linear models: a simulation study. Poster session presented at Linguistic Evidence 2014, Tübingen, Germany.
- Müller, F. H., Hanfstingl, B., & Andreitz, I. (2007). Skalen zur motivationalen Regulation beim Lernen von Schülerinnen und Schülern. Adaptierte und ergänzte Version des Academic Self-Regulation Questionnaire (SRQ-A) nach Ryan & Connell. *Wissenschaftliche Beiträge aus dem Institut für Unterrichts- und Schulentwicklung (IUS)*, 1, 1-17.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625-632.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Schlak, T. (2008). Fremdsprachenlernneigung: Tabuthema oder Forschungslücke? Zum Zusammenhang von Fremdsprachenlernneigung, Fremdsprachenlernen und Fremdsprachenvermittlung. *Zeitschrift für Fremdsprachenforschung*, 19(1), 3-30.
- Sprouse, J., & Almeida, D. (2017). Design sensitivity and statistical power in acceptability judgment experiments. *Glossa*, 2(1), 1-32.
- Tellier, A., & Roehr-Brackin, K. (2013). The development of language learning aptitude and metalinguistic awareness in primary-school children: A classroom study. *Essex Research Reports in Linguistics*, 62(1), 1-28.
- Yeh, Y. (2012). Aptitude-treatment interaction. In: Seel, N.M. (Hrsg.), *Encyclopedia of the Sciences of Learning* (S. 295-299). Boston, MA: Springer.