

## **Veränderung der Qualität von Sachunterricht im Vorbereitungsdienst**

### **Ausgangslage & theoretischer Hintergrund**

In diversen Studien konnte bereits gezeigt werden, dass sich in Praxisphasen der Lehrerbildung neben Wissen, Selbstbewusstsein und Überzeugungen auch allgemeine Lehrfähigkeiten entwickeln (Niemi, 2011; Schmidt et al., 2007). Zudem konnte gezeigt werden, dass einzelne Aspekte von Unterrichtsqualität signifikant über den Verlauf von Praxisphasen ansteigen (Baer et al., 2011). Der Anstieg ist dabei in einigen Bereichen, z. B. Organisation von Unterricht, stärker als in anderen, z. B. Einbau von Aspekten der Heterogenität (Baer et al., 2011). Dass es auch für erfahrene Lehrkräfte schwieriger ist, Merkmale auf der Individualebene der SchülerInnen umzusetzen als Merkmale der Strukturierung und Organisation, zeigen Analysen der Hamburger Schulinspektion (Pietsch, 2010).

All diese Studien nehmen vorrangig angehende Lehrkräfte aus der Sekundarstufe I zum Untersuchungsgegenstand oder beziehen sich auf Unterrichtsfächer wie Mathematik und Deutsch. Wie sich die Entwicklung von Sachunterrichtslehrkräften im Vorbereitungsdienst gestaltet ist hingegen ungeklärt. Das vorgestellte Teilprojekt zur Durchführung von Sachunterricht (SU) hat zum Ziel, diese ungeklärte Entwicklung nachzuzeichnen. Dabei wird vor allem in den Blick genommen, in welchem Maße die angehenden Lehrkräfte über die Ausbildungszeit hinweg in der Lage sind, qualitätsstiftende Maßnahmen im SU umzusetzen.

### **Forschungsfragen**

**F1:** Wie verändert sich die Durchführungsqualität von SU im Vorbereitungsdienst?

H1: Die Qualität der Unterrichtsdurchführung nimmt im Verlauf des Vorbereitungsdienstes zu (basierend auf z. B. Baer et al., 2011).

**F2:** Wie verändert sich die Durchführungsqualität von SU hinsichtlich einzelner Qualitätsmerkmale im Vorbereitungsdienst?

H2.1: Grundlegende Aspekte der Unterrichtsorganisation und -strukturierung (Klassenführung, Klarheit & Strukturiertheit, Angebotsvielfalt) entwickeln sich insbesondere in der ersten Hälfte des Vorbereitungsdienstes (basierend auf Pietsch, 2010).

H2.2: Aspekte, die die Individualebene der SuS einbeziehen (Umgang mit Heterogenität, Aktivierung, Lernförderliches Klima) entwickeln sich insbesondere in der zweiten Hälfte des Vorbereitungsdienstes (basierend auf Pietsch, 2010).

### **Methoden & Design**

Um die Entwicklung im Laufe der 1,5-jährigen Ausbildung in NRW nachzuzeichnen, wurden im Rahmen der verpflichtenden Unterrichtsbesuche drei Unterrichtsstunden nach standardisierenden Richtlinien videographiert. Die Aufnahmen erfolgten je einmal am Anfang, in der Mitte und am Ende des Vorbereitungsdienstes bei 12 LehramtsanwärterInnen LAA.

Für die Analyse der Videos wurde ein Kategoriensystem verwendet, welches guten Sachunterricht operationalisiert. Das Kategoriensystem basiert auf den Merkmalen guten Unterrichts (z. B. Helmke, 2009; Meyer, 2007). Zur fachspezifischen Ausdifferenzierung fließen sachunterrichtsdidaktische Aussagen in die Merkmalsbeschreibungen ein (vgl. Rau, Windt & Rumann, 2015). Das Kategoriensystem bildet die Merkmalsgruppen der Organi-

sation & Strukturierung (Merkmale Klassenführung, Klarheit & Strukturiertheit, Gestaltung des Lernangebots) sowie Individualebene der SchülerInnen (Merkmale Aktivierung, Lernförderliches Klima, Umgang mit Heterogenität) ab und wurde über ein Expertenrating validiert.

Die Analysen der videographierten Unterrichtsstunden erfolgten standardisiert nach den Vorgaben zur qualitativen Inhaltsanalyse nach Mayring & Fenzl (2014). Die Auswertungseinheit teilt sich in vier Analyseschritte:

- (1) Unterrichtsvideo insgesamt ansehen; Beobachtungsnotizen anfertigen
- (2) Sichtung des Anhangs mit Fotos aus dem Klassenraum und kopiertem Arbeitsmaterial
- (3) Unterrichtsvideo in 5-minütigen Sequenzen ansehen und kodieren
- (4) Finale Kodierungen unter Berücksichtigung der Schritte 1-3

Die Kontexteinheit als größte kodierbare Einheit ist in diesem Fall die gesamte Unterrichtsstunde. Zum einen werden in Form eines Overall-Samplings alle niedrig-inferenten Facetten (z. B. Lehrer-Schüler-Zentrierung) nach Schritt (1) eingeschätzt; zum anderen werden alle hoch-inferenten Facetten (z. B. positive Fehlerkultur) in Form eines Overall-Samplings in Schritt (4) eingeschätzt.

Zwischen diesen Kontextkodierungen (1) und (4) liegen die Kodiereinheiten als kleinste kodierbare Einheit. In dieser Studie sind die Kodiereinheiten entweder einzelne Arbeitsmittel bzw. Fotos oder einzelne Events innerhalb der 5-minütigen Sequenzen. Einzelne Facetten werden auf Basis den jeweiligen Events kodiert (z. B. Umgang mit Störungen).

Trotz des komplexen Verfahrens konnte mit zwei geschulten Kodierern an zunächst 6 doppelt kodieren Unterrichtsstunden eine durchaus zufriedenstellende Beurteilerübereinstimmung von  $.571 < \kappa \leq 1.0$  für die einzelnen Facetten erreicht werden (Wirtz & Caspar, 2002). Weitere Analysen stehen noch aus.

## Ergebnisse

Die Veränderung der Durchführungsqualität von Sachunterricht verläuft auf der deskriptiven Ebene positiv (Abb. 1, H1), wird jedoch statistisch nicht signifikant ( $F(2,22) = 2.160$ ,  $p = .143$ , part.  $\eta^2 = .164$ ). Ähnliche Ergebnisse findet man auch für die Merkmale der Organisation & Strukturierung (Abb. 1, H2.1) mit  $F(2,22) = 0.930$ ,  $p = .396$ , part.  $\eta^2 = .078$  sowie für die Merkmale der Individualebene der SchülerInnen (Abb. 1, H2.2) mit  $F(2,22) = 2.778$ ,  $p = .084$ , part.  $\eta^2 = .202$ . Aufgrund fehlender Normalverteilung der Daten wurde stets zusätzlich der non-parametrische Friedman Test gerechnet, der allerdings nicht zu grundsätzlich anderen Ergebnissen führt.

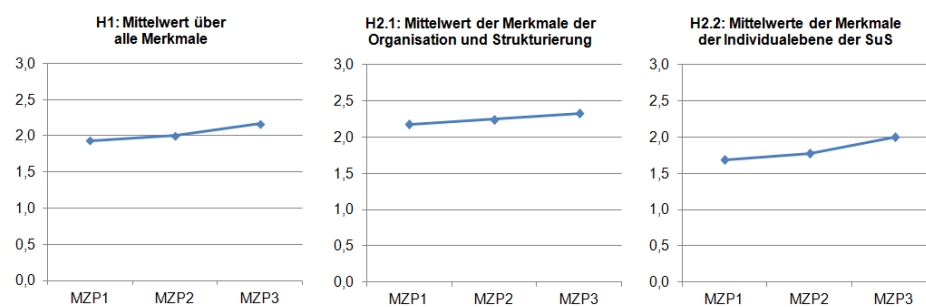


Abbildung 1

Auch die in Hypothese H2.1 und H2.2 postulierten Unterschiede im Entwicklungsverlauf der Merkmale der Organisation & Strukturierung bzw. der Merkmale der Individualebene der SchülerInnen. Diese zeigen sich nicht in einem statistisch bedeutsamen Ausmaß. Allerdings werden die Merkmale der Organisation & Strukturierung zu allen drei

Messzeitpunkten in höherem Maße umgesetzt als die Merkmale der Individualebene. Der Unterschied zwischen den Merkmalsgruppen wird im Wilcoxon-Test für abhängige Stichproben zu allen drei Messzeitpunkten signifikant (MZP 1:  $p = .002$ , MZP 2:  $p = .003$ ; MZP 3:  $p = .008$ ).

Zur tiefergehenden Datenanalyse wurden die Profillinien aller Probanden optisch inspiziert. Dabei konnten drei Gruppen identifiziert werden: Positive Veränderung, Stagnation und negative Veränderung über die MZP 1 und 3.

Im Rahmen einer polythetischen Typenbildung, welche aus dem vorliegenden Material erfolgt, wird im Weiteren geprüft, ob den identifizierten Veränderungsgruppen auch unterschiedliche (Sub-)Typen zu Grunde liegen. Hierfür werden die LAA vor allem in Bezug auf ihre Merkmalsausprägungen zu den jeweiligen Messzeitpunkten miteinander verglichen. Darüber hinaus werden auch vorliegende personenbezogene Daten (z. B. Studium, Erfahrung) in die Typenbildung einbezogen.

### Diskussion

Aufgrund fehlender statistischer Signifikanz müssen alle drei Hypothesen zurückgewiesen werden. Die gefundenen mittleren bis großen Effekte deuten aber auf eine Bedeutsamkeit des Vorbereitungsdienstes hinsichtlich der Qualität der Unterrichtsdurchführung hin. Die mangelnde Signifikanz könnte durch die geringe Stichprobengröße begründet sein.

Darüber hinaus beruhen die vorliegenden Ergebnisse auf einer Gelegenheitsstichprobe und spiegeln möglicherweise eine Positivauswahl wieder. Indizien dafür zeigen sich in der hohen Ausprägung der Qualitätsmerkmale über alle Messzeitpunkte (Abb. 1). Auch der Umstand, dass Unterrichtsbesuche als Datengrundlage dienen, kann dazu führen, dass die Mittelwerte besonders hoch ausfallen. Viele LAA investieren dort besonders viel Energie (Storr, 2006).

Für LAA, die zu Beginn eher wenige qualitätsstiftenden Maßnahmen zeigen, scheint es viele Lerngelegenheiten zu geben. Für LAA, die bereits zu Beginn der Studie viele qualitätsstiftenden Maßnahmen im SU umsetzen, scheint es hingegen entweder zu wenige Lerngelegenheiten zu geben (Stagnierer) oder der Fall zu sein, dass sie insbesondere am Anfang sehr viel Unterstützung durch Mentoren o. ä. erhalten (negative Veränderung).

### Literatur

- Baer, M., Kocher, M., Wyss, C., Guldimann, T., Larcher, S. & Dörr, G. (2011). Lehrerbildung und Praxiserfahrung im ersten Berufsjahr und ihre Wirkung auf die Unterrichtskompetenzen von Studierenden und jungen Lehrpersonen im Berufseinstieg. *Zeitschrift für Erziehungswissenschaft*, 14(1), 85–117.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett Kallmeyer.
- Mayring, P., & Fenzl, T. (2014). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.). *Handbuch Methoden der empirischen Sozialforschung* (543-556). Wiesbaden: Springer.
- Meyer, H. (2007). *Was ist guter Unterricht?* Frankfurt am Main: Cornelsen Scriptor.
- Niemi, H. (2011). Educating student teachers to become high quality professionals - a Finnish case. *CEPS Journal*, 1(1), 43–66.
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 121–148
- Rau, S., Windt, A. & Rumann, S. (2015). Entwicklung von Sachunterricht in der zweiten Phase der Lehrerbildung. In: S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Bremen 2014* (S. 603-605). Kiel: IPN.
- Schmidt et al. (2007). *The Preparation Gap: Teacher Education for Middle School Mathematics in Six Countries (MT21 Report)*. Abgerufen von <http://usteds.msu.edu/MT21Report.pdf>
- Storr, B. (2006). „In der Lehrprobe da machst du `ne Show“ – Das Referendariat als Gegenstand rekonstruktiver Sozialforschung. Berlin: Logos.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.