

Tanja Sczepanski, Johannes Hauser, Anna von Werthern &
Andreas Hendrich

Tutorenausbilder*innen als Multiplikatoren für gute Lehre – erste Evaluationsergebnisse des Qualifikationsprogramms TutorPlus

Zusammenfassung

TutorPlus ist ein zweijähriges multiplikatorenbasiertes Qualifikationsprogramm für Tutorenausbilder*innen. Fachdozent*innen aus unterschiedlichen Fakultäten werden dazu ausgebildet, ihrerseits Tutor*innen für spezifische Lehraufgaben zu qualifizieren. Um die beständige Qualität des Programms und damit eine nachhaltige Verbesserung der Lehre zu gewährleisten, ist die evaluative Begleitung eine tragende Komponente von TutorPlus. Mit einer Zusammenstellung von quantitativen und qualitativen Methoden werden zu diversen Messzeitpunkten Kompetenz- und Wissenszuwachs, die Einstellung zur Lehre (Dozierenden- und Studierendenorientierung), Zufriedenheit und Transfererfolg gemessen. Die bisherigen Ergebnisse zeigen unter anderem, dass die Studierendenorientierung im Laufe der Ausbildung signifikant zugenommen hat. Zudem ist ein signifikanter Zuwachs an Wissen über zentrale Konzepte der Tutorienarbeit sowie eine subjektiv empfundene Zunahme an methodisch-didaktischer Kompetenz – sowohl bei Tutorenausbilder*innen, als auch bei Tutor*innen – feststellbar. Hohe Mittelwerte in den Zufriedenheitsratings der drei Kohorten reflektieren die Akzeptanz der Ausbildung. Die Veranstaltungen werden durchgehend als nützlich und praxistauglich bewertet.

Schlüsselwörter

Tutorienarbeit; Programmevaluation; Lehrqualität; Lehreinstellung; Multiplikatoren

Tutor-trainers as multipliers for good teaching – first evaluative results for the qualification program TutorPlus

Abstract

TutorPlus is a two-year qualification program in which lecturers from various faculties are trained to qualify tutors within specific teaching responsibilities. In order to constantly ensure the quality of the program and thus guarantee a sustainable improvement of the teaching at universities, the evaluative support is a fundamental component of TutorPlus. With a compilation of both, quantitative and qualitative methods, the increase of competence and knowledge, the approaches to teaching (student-focused, as well teacher-focused approaches), satisfaction and the success of transfer are measured at different points in time. The previous results show that the student-focused approaches to teaching have been increased during the period of tuition. Also, a significant growth of knowledge concerning central concepts of the seminars taught by tutors, as well as a subjectively felt increase of methodological-didactic competence – within both parties, tutor-trainers and tutors themselves – is ascertainable. A high average within the ratings of satisfaction covering three cohorts, reflect the acceptance of the tuition. The concept is continuously evaluated as practical and useful.

Key words

tutoring; program evaluation; teaching quality; approaches to teaching; multipliers

1 Struktur & Inhalte der Ausbildung

In der deutschen Hochschullandschaft gibt es ein breites Spektrum an Tutorien. Tutor*innen werden in der Regel nur sehr marginal auf ihre Tätigkeit vorbereitet, es fehlt daher oft vor allem an methodischem und didaktischem Wissen darüber, wie man ein Tutorium plant, durchführt und nachbereitet. Um dem entgegenzuwirken, wurde im Rahmen des Qualitätspakt Lehre mit TutorPlus ein multiplikatorenbasiertes Programm aufgesetzt, welches eine Ausbildung für Lehrende aller Hierarchieebenen in Form einer hochschulischen Weiterbildung anbietet. Engagierte Lehrende aus unterschiedlichsten Fachbereichen werden zentral zu Tutorenausbilder*innen qualifiziert (siehe Antosch-Bardohn & Primus, 2015 für eine ausführliche Beschreibung). Das zentrale Ziel ist dabei, die Qualität von Tutorien durch eine strukturierte und homogenisierte Ausbildung der Tutor*innen langfristig zu verbessern und als genuinen Teil der Lehrkultur an der Universität zu verankern. In aufeinanderfolgenden Präsenz- und Transferphasen entwickeln die Tutorenausbilder*innen ein Schulungskonzept für Tutor*innen. Dieses wird im authentischen Setting (Tutorenschulung im Fachbereich) umgesetzt und von Experten- und Peer-coachings begleitet. Den angehenden Tutorenausbilder*innen stehen während der Transferphasen Expert*innen für individuelle Konzeptionsgespräche etc. zur Verfügung. Zudem können sie auf das umfassende hochschuldidaktische Kursangebot der Universität zugreifen.

Die Tutorenschulungen finden, dem multiplikativen Gedanken Folge leistend, fachbasiert und dezentral an den Fakultäten statt, da die Lehraufgaben von Tutor*innen sehr stark fach- bzw. bereichsabhängig sind. Damit die Ausbildung Erfolg hat, muss sie sich also an dem Bedarf der Fachbereiche und den Teilnehmenden orientieren. Deshalb wird vor Beginn der Ausbildung eine Bedarfsanalyse in den Fachbereichen durchgeführt, die es ermöglicht, die Inhalte der Ausbildung an die bestehenden Bedürfnisse und Vorkenntnisse anzupassen. Bereits vor Beginn der Ausbildung müssen die am Programm teilnehmenden Personen hochschuldidaktische Fortbildungen im Rahmen von 60 Arbeitseinheiten nachweisen, um zur Ausbildung zugelassen zu werden.

Der spezifische Mehrwert des multiplikatorenbasierten Ausbildungsprogramms ist darin zu sehen, dass über die Tutorenausbilder*innen als Multiplikatoren eine studierendenzentrierte und kompetenzorientierte Lehrkultur bottom-up, fakultätsübergreifend und nachhaltig implementiert werden kann. Durch die Ausbildung soll neben der Verbesserung der Tutorien auch das hochschuldidaktische Reflexionsniveau der Ausbilder*innen erhöht werden, was somit auch zur Verbesserung der eigenen Fachlehre beitragen soll. Das Motto *Gute Lehre wächst von unten!* kennzeichnet einen Generationswechsel in der Hochschuldidaktik – die Tutor*innen von heute sind die Lehrenden von morgen.

2 Ziele und Funktionen der Evaluation

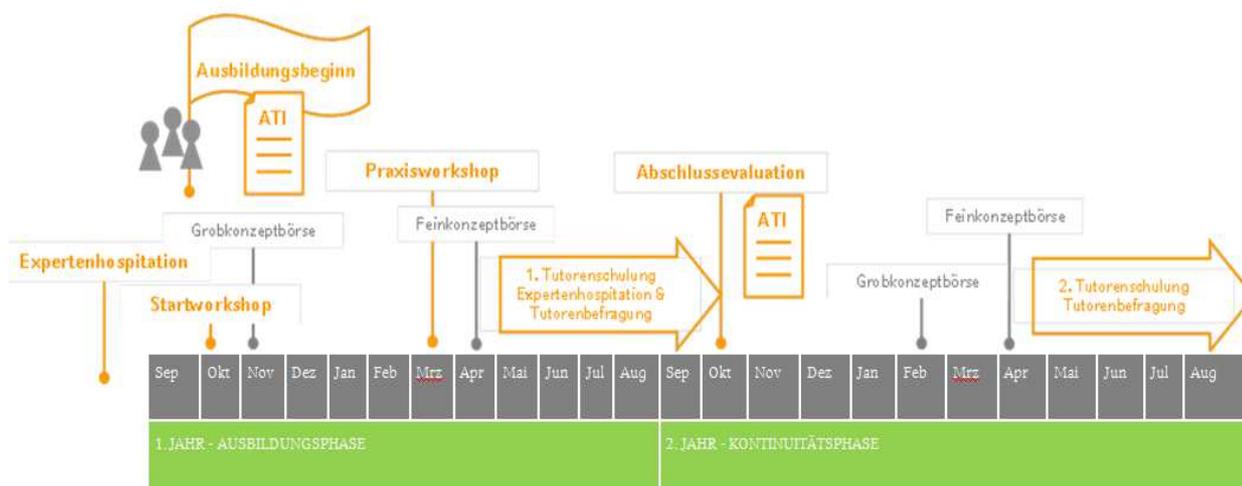
Um die beständige Qualität des Programms und damit eine nachhaltige Verbesserung der Lehre zu gewährleisten, ist die evaluative Begleitung eine tragende Komponente von TutorPlus. Stockmann (2004) folgend, kann eine Evaluation im Allgemeinen verschiedene Ziele und Funktionen haben: Erkenntnis, Lernen, Kontrolle und Legitimation. Die grundlegenden Fragen der Erkenntnis lauten bei TutorPlus: Hat unser Programm

überhaupt eine Wirkung und wenn ja, welche? Wie verändern sich die Teilnehmenden? Wie werden einzelne Maßnahmen bewertet? In der Konsequenz soll ein Lernprozess adressiert werden: Wie kann man das Programm modifizieren? Sollen bestimmte Inhalte verdichtet werden? Was sollte beibehalten werden und was nicht? Im Anschluss stellt sich die Frage, welche Kontrollmaßnahmen zur Steuerung des Programms getroffen werden. Die Legitimation des Programms entsteht schließlich durch eine empirisch fundierte Aussage darüber, ob das Programm im hochschuldidaktischen Kontext berechtigt und sinnvoll ist. Gemäß der unterschiedlichen Ziele und Funktionen umfasst die Evaluation in ihren einzelnen Maßnahmen formative wie auch summative Elemente. Im Folgenden wird vorwiegend der Frage nach Erkenntnis nachgegangen, da diese für TutorPlus die fundamentalste ist und sich Möglichkeiten des Lernens, der Kontrolle und der Legitimation zwangsläufig aus ihrer Beantwortung ergeben.

3 Evaluationsdesign

Methodisch sieht das Evaluationskonzept zunächst für jede Kohorte ein Ein-Gruppen-Design mit Pretest-Messung, Verlaufsmessungen und Posttest-Messung vor (Balzer, 2005). Weiterhin sollen die Kohorten auf verschiedenen Ebenen miteinander verglichen werden. Mit Hilfe von quantitativen und qualitativen Methoden werden zu diversen Messzeitpunkten Kompetenz- und Wissenszuwachs, Einstellungsveränderungen, Zufriedenheit und Transfererfolg mittels Selbst- und Fremdeinschätzung gemessen. Die Methodentriangulation dient dazu, die Ergebnisse zu verdichten und zu vertiefen. Auf dieser Basis werden dann die Kohorten miteinander verglichen. Das am häufigsten genutzte Medium der Evaluation ist die Befragung durch Fragebögen. Für jeden Messzeitpunkt wurde ein spezifischer Fragebogen mit geschlossenen und offenen Fragen konzipiert. Zusätzlich finden Hospitationen der Teilnehmenden durch die Trainer*innen statt. Um hier eine Vergleichbarkeit zwischen den Teilnehmenden sowie zwischen zwei Hospitationszeitpunkten zu erreichen, wurde ein standardisierter Beobachtungsbogen entwickelt. Die folgende Grafik zeigt die Messzeitpunkte der Begleitevaluation während der Ausbildung (siehe Abbildung 1).

Abbildung 1: Ausbildungsverlauf und Messzeitpunkte



Die Fragebogenitems können jeweils anhand fünfstufiger Likert-Skalen bewertet werden. Die Extrema der Skalen wurden einheitlich mit *trifft überhaupt nicht zu* und *trifft völlig zu* bezeichnet. Die Fragebögen wurden überwiegend im pen-&-paper-Format bearbeitet und mit EvaSys ausgewertet. Um eine anonyme Zuordnung von Fragebögen zu Teilnehmenden über den Ausbildungszeitraum hinweg und darüber hinaus möglich zu machen, wurden personenspezifische Codes etabliert und abgefragt.

Im Folgenden werden die unterschiedlichen Evaluationsinstrumente, die im Verlauf der Ausbildung eingesetzt werden, beschrieben. Dabei wird jeweils auf die aktuellste Form der Instrumente und das Vorgehen in den bisherigen drei Kohorten Bezug genommen.

3.1 Evaluation im Rahmen der Startworkshops

Zu Beginn der Ausbildung findet ein mehrtägiger Startworkshop statt. Dieser umfasst das gegenseitige Kennenlernen der Kursteilnehmer*innen und der Trainer*innen, eine erste Einführung in das Thema und den Ablauf der Ausbildung sowie eine grundlegende Wissensvermittlung über Tutorien. Außerdem werden hier tiefere Kenntnisse im Bereich Methodik, Didaktik, Umgang mit unterschiedlichen Teilnehmertypen, Moderation, Gruppenführung und Rolle der Tutorenausbilder*innen und Tutor*innen vermittelt. Die Evaluation zu Beginn der Ausbildung hat vier Themen zum Gegenstand: Erwartungen an den Startworkshop und die Ausbildung an sich, Erfassung des hochschuldidaktischen Vorwissens und der Einstellung zur Lehre sowie eine subjektive Einschätzung des Wissenszuwachses. Diese erste Evaluation hat zwei Messzeitpunkte (MZP): Zu Beginn des Startworkshops (MZP I) und am Ende des Startworkshops (MZP II). Zu MZP I werden zunächst die Erwartungen an die Ausbildung zur Tutorenausbilderin bzw. zum Tutorenausbilder und an den Startworkshop mittels offener Fragen abgefragt. Zu MZP II sollen die Teilnehmenden entsprechend angeben, ob bzw. inwiefern sich die vorher formulierten Erwartungen erfüllt oder nicht erfüllt haben. Dieser Erwartungsabgleich wird u.a. zur Optimierung des Startworkshops herangezogen. Um auf die spezifischen Bedürfnisse der Kohorten gezielt eingehen zu können, sollen die Teilnehmenden zu MZP I ihr hochschuldidaktisches Vorwissen in den Bereichen *Präsentationstechniken* (z.B. *Ich bin mir der Stärken meines stimmlichen Ausdrucks bewusst*), *Gesprächsführung* (z.B. *Ich kann Metakommunikation zur bewussten Gesprächsführung nutzen*) sowie *Didaktik & Methodik* (z.B. *Ich kann Ablaufpläne nach dem AVIVA-Schema konstruieren*) einschätzen. Durch die Erfassung des Vorwissens in den Bereichen Präsentationstechniken, Gesprächsführung, Didaktik & Methodik sowie der Erwartungen zu Beginn der Ausbildung kann das Programm kohortenspezifisch angepasst werden, um den Bedürfnissen der Teilnehmer*innen optimal gerecht zu werden.

Zusätzlich werden die Teilnehmenden aufgefordert, zu den unterschiedlichen Messzeitpunkten des Startworkshops ihr Wissen über im Startworkshop vermittelte Inhalte einzuschätzen. Dies diene der Erfolgskontrolle: Haben die Teilnehmenden über den Workshop hinweg theoretisches Wissen erworben? Entsprechend waren die Items zu beiden Messzeitpunkten die gleichen; beispielsweise *Ich kann auswählen, welches didaktische Wissen für meine Tutoren relevant ist.* und *Ich kenne Einflussmöglichkeiten auf gruppendynamische Prozesse.*

Ein weiteres Ziel der Evaluation des Startworkshops ist die Erfassung der Lehrorientierung mittels des *Approaches to Teaching Inventory* (ATI) nach Trigwell und Prosser (2004). Trigwell & Prosser entwickelten mit dem ATI eine Möglichkeit, zwei Facetten von Lehrorientierung zu messen (ihnen zu Folge sind studierendenzentrierte und dozierendenzentrierte Lehreinstellungen unterscheidbar), wobei die Reliabilität der beiden Subskalen als akzeptabel angesehen werden kann (Prosser & Trigwell, 2006).

Während die Konstrukte zunächst als unabhängig angesehen wurden, zeigte sich empirisch, dass eine leicht negative Korrelation zwischen diesen besteht (Prosser & Trigwell, 2006). Dabei wird die studierendenfokussierte Lehrorientierung als die effektivere (im Sinne des studentischen Kompetenzerwerbs) angesehen, da man davon ausgeht, dass Dozierende, die Wissenserwerb als Ergebnis eines aktiven Konstruktionsprozesses verstehen (studierendenfokussierte Lehrorientierung), Lernprozesse wirkungsvoller unterstützen können, als Lehrende, die Wissenserwerb als Ergebnis des Transports von Wissen vom Lehrenden zum Lernenden verstehen (dozierendenfokussierte Lehrorientierung; Braun & Hannover, 2008). Des Weiteren scheint sich eine studierendenfokussierte Lehrorientierung stärker in einem fachlichen und überfachlichen Kompetenzzuwachs von Studierenden niederzuschlagen als eine dozierendenfokussierte (Braun & Hannover, 2008).

Bei der Evaluation der TutorPlus-Ausbildung wird die Lehreinstellung über eine deutsche Version des ATI erfasst; die Übersetzung erfolgte im Rahmen eines wissenschaftlichen Projekts an der LMU München. Da erwartet wurde, dass sich die Lehreinstellung über den gesamten Ausbildungsverlauf hinweg verändert, wird der ATI zu Beginn (MZP I, Startworkshop) und gegen Ende der Ausbildung (MZP II, Abschlussevaluation) eingesetzt.

3.2 Evaluation im Rahmen der Praxisworkshops

Der Praxisworkshop umfasst drei Tage und beschäftigt sich inhaltlich vorrangig mit der eigenständigen Durchführung von Bausteinen der individuell erarbeiteten Konzepte. Die Ziele waren dabei der Erwerb von Handlungskompetenz, die Transfersicherung, das gegenseitige Geben von Feedback sowie der soziale Austausch.

Im Rahmen der Evaluation des Praxisworkshops sollen die Teilnehmenden angeben, ob sie sich *vorbereitet [fühlen], die Schulung von Tutor*innen durchzuführen* und ob das Team sie noch in irgendeiner Weise unterstützen kann. Eine weitere Einschätzung soll dazu abgegeben werden, inwiefern sich ihrer Kompetenz im Bereich *Feedback geben und nehmen* vergrößert hat. Mit offenen Fragen wird erfragt, was sich die Teilnehmenden *noch an Inhalten gewünscht* hätten und welche *drei relevanten Erkenntnisse* sie aus dem Workshop mitnehmen. Ergänzend können die Teilnehmenden ihre Zufriedenheit mit dem Praxisworkshop und der bisherigen Gesamtausbildung mittels mehrerer skaliertes Items mitteilen.

3.3 Abschlussevaluation

Die Abschlussevaluation markiert in jeder Kohorte den Abschluss des ersten Ausbildungsjahres. In der anschließenden Kontinuitätsphase agieren die Tutorenausbilder*innen selbständig und weitgehend unabhängig von TutorPlus. In der Abschlussevaluation sol-

len die Tutorenausbilder*innen zunächst einschätzen, ob sie in den verschiedenen Ausbildungsbereichen einen bedeutsamen Kompetenzzuwachs erleben (*Ich schätze meine Kompetenzen in [Kompetenzbereich] nach der Ausbildung bedeutend höher ein als davor*). Die definierten Kompetenzbereiche waren dabei *Didaktik, Rhetorik & Präsentation, Methodik, Gruppenführung, Diskussionsleitung* und *Train-the-Trainer*.

Zudem wird die Zufriedenheit mit den Inhalten (z.B. *Ich habe mich aufgrund der in der Ausbildung vermittelten Inhalte gut vorbereitet gefühlt, die Schulung von Tutorinnen und Tutoren durchzuführen.*), der Organisation (z.B. *Zum allgemeinen zeitlichen und organisatorischen Ablauf der Ausbildung habe ich mich stets gut informiert gefühlt.*) und den Trainer*innen (z.B. *Die Ausbilderinnen und Ausbilder wirkten auf mich fachlich kompetent.*) mit je drei Items abgefragt. Zur Abschlussevaluation werden die Teilnehmer*innen mithilfe des ATI ein weiteres Mal bezüglich ihrer Lehreinstellung befragt. Ergänzt wird die Abschlussevaluation mit offenen Fragen, die Hinweise darauf geben sollen, wie man die Ausbildung im weiteren Verlauf optimieren kann. So werden sie zum Beispiel gefragt, welches Lehrformat für sie den größten Mehrwert hatte, welche Ausbildungsinhalte sie als besonders hilfreich empfanden oder welche Inhalte ihnen gefehlt haben.

3.4 Evaluation im Rahmen der Hospitationen

Vor Beginn der Ausbildung sowie bei der Durchführung der Tutorenschulungen findet stets eine Hospitation durch die Trainer*innen statt. Hierzu wurde ein differenzierter Beobachtungsbogen entwickelt. Die Hospitation soll zusätzliche Informationen aus Expertenperspektive liefern. Durch einen Vorher-Nachher-Vergleich (es findet jeweils eine Hospitation vor und nach Teilnahme an TutorPlus statt), soll erfasst werden, ob sich der Kompetenzzuwachs auch in der Umsetzung feststellen lässt. Der standardisierte Beobachtungsbogen hilft dabei, die Beobachtungen strukturierter zu gestalten. Außerdem dient er als Einstiegshilfe in das Rückmeldegespräch. Aufgeteilt ist der Bogen in die sechs Kompetenzfelder *Rhetorik und Präsentation, Kommunikation und Interaktion, Gruppenführung und Diskussion, Train-the-Trainer-Perspektive, Didaktik und Methodik*. Zusätzlich wird zum Ende des Bogens der Gesamteindruck bewertet. Eingeschätzt werden die einzelnen Verhaltensweisen auf einer fünfstufigen Skala, zusätzlich besteht die Möglichkeit, konkrete Beobachtungen anzumerken.

3.5 Evaluation im Rahmen der Tutorenschulungen

Zwischen Praxisworkshop und Abschlussevaluation führen die Teilnehmenden eine erste Tutorenschulung durch, in der Kontinuitätsphase konzipieren sie eine weitere Schulung und führen diese ebenfalls durch. Die Tutorenausbilder*innen sollen hier Sicherheit in der Konzeption und Durchführung von Tutorenschulungen bekommen, damit sie während ihrer Lehrtätigkeit auch ohne Unterstützung selbständig Tutor*innen schulen können. Die Tutorenschulung ist die konkrete Realisierung der vorher erstellten Fein- und Grobkonzepte. Die Evaluation der Tutorenschulungen zeichnet sich dadurch aus, dass hier zum ersten Mal der Transfer gemessen wird. Es wird überprüft, ob das Gelernte in den Tutorenschulungen angewendet, umgesetzt und vermittelt wird und so bei den Tutor*innen Wirkung zeigen kann. Dazu bewerten die Tutor*innen hier ihren subjektiven Kompetenzzuwachs während *ihrer* Ausbildung analog zur Abschlussevaluation: *Ich schätze meine*

Kompetenzen im Bereich [Didaktik / Rhetorik & Präsentation / Methodik / Gruppenführung] nach der Tutorenschulung bedeutend höher ein als davor. wird die Zufriedenheit mit der Schulung und der jeweiligen ausbildenden Person erfasst. Die Items des Fragebogens wurden an den Fragebogen der Abschlussevaluation angepasst, um eine Vergleichbarkeit zu ermöglichen.

4 Ergebnisse

4.1 Wissenszuwachs (Startworkshop)

In allen drei Kohorten zeigt sich ein signifikanter Wissenszuwachs im Verlauf des Startworkshops (siehe Tabelle 1). Die entsprechenden Effekte sind als sehr groß einzustufen (Cohen, 1992). Man kann also davon ausgehen, dass es gelungen ist, zentrale theoretische Konzepte zu vermitteln.

Tabelle 1: Wissenszuwachs

	MZP I (Freitag)		MZP II (Sonntag)		t-Wert	d ^a
	M	SD	M	SD		
1. Kohorte	2.78	.53	3.67	.36	3.89 ^b	2.00
2. Kohorte	3.04	.77	4.04	.44	4.76 ^c	1.58
3. Kohorte	2.77	.58	3.91	.78	4.88 ^d	1.66

Anmerkungen:

Jeweils Skala von 1 „stimme überhaupt nicht zu“ bis 5 „stimme völlig zu“

^a Zur besseren Vergleichbarkeit wurde Cohens d für unabhängige Stichproben berechnet

^b Unterschied ist auf einem Bonferroni-Holm-korrigierten Signifikanzniveau von $\alpha = .017$ signifikant

^c Unterschied ist auf einem Bonferroni-Holm-korrigierten Signifikanzniveau von $\alpha = .025$ signifikant

^d Unterschied ist auf einem Bonferroni-Holm-korrigierten Signifikanzniveau von $\alpha = .05$ signifikant

4.2 Zufriedenheit (Abschlussevaluation)

Eine deskriptive Datenanalyse zeigt, dass die Zufriedenheit mit der Ausbildung gegen Ende der Ausbildung in den Kohorten 1 bis 3 hoch ausgeprägt war (siehe Tabelle 2). Für die Variablen *Zufriedenheit mit Inhalten*, *Zufriedenheit mit Organisation* und *Fazit zur Ausbildung* konnten keine signifikanten Unterschiede zwischen den Kohorten festgestellt werden (einfaktorielle ANOVA; alle $p > .05$). Für die Variable *Zufriedenheit mit TrainerInnen* werden die Varianzen in den verschiedenen Kohorten als heterogen angenommen (Levene-Test; $F(2,24) = 20.73, p < .001$). Die Voraussetzungen für das Durchführen einer ANOVA wurden somit verletzt. Der Kruskal-Wallis-Test, der als nichtparametrische Alternative gewählt wurde, signalisiert für diese Variable einen signifikanten Unterschied zwischen den Kohorten ($p = .045$). Der Post-hoc-Test nach Games-Howell identifiziert dabei keine signifikanten Einzelvergleiche (alle $p > .05$). Ein Grund für die Signifikanz des Kruskal-Wallis-Tests mag sein, dass von den Teilnehmer*innen der zweiten Kohorte ohne

jegliche Varianz eine „5“ vergeben wurde. Insgesamt kann aus den Daten geschlossen werden, dass die Zufriedenheit in den verschiedenen Bereichen in allen Kohorten auf einem hohen Niveau stabil geblieben ist.

Tabelle 2: Zufriedenheit mit der Ausbildung

	1. Kohorte		2. Kohorte		3. Kohorte	
	M	SD	M	SD	M	SD
Zufriedenheit mit Inhalten	4.59	.36	4.71	.28	4.57	.47
Zufriedenheit mit Organisation	4.61	.42	4.44	.50	4.65	.53
Zufriedenheit mit TrainerInnen	4.93	.15	5.00	.00	4.73	.31
Fazit zur Ausbildung	4.63	.69	4.81	.26	4.60	.52

Anmerkung:

Jeweils Skala von 1 „stimme überhaupt nicht zu“ bis 5 „stimme völlig zu“

4.3 Lehreinstellung (Startworkshop – Abschlussevaluation)

Die Veränderung der Lehreinstellung zwischen Startworkshop und Abschlussevaluation wurde für die zweite und dritte Kohorte mittels einer zweifaktoriellen, gemischten MANOVA (Typ III) berechnet. Der between-Faktor war dabei die Kohorte, der within-Faktor der Messzeitpunkt. Die abhängigen Variablen waren die mittleren Summenscores der über den ATI gemessenen Studierenden- und Dozierendenorientierung. Es zeigten sich ein signifikanter Haupteffekt in der Veränderung der Studierendenorientierung ($F(1,14) = 14.39, p = .002, \eta^2 = .19, \eta_p^2 = .51$), ein signifikanter Interaktionseffekt zwischen der Veränderung der Studierendenorientierung und der Kohorte ($F(1,14) = 6.01, p = .028, \eta^2 = .08, \eta_p^2 = .30$) und ein signifikanter Interaktionseffekt zwischen den Veränderungen von Studierenden- und Dozierendenorientierung ($F(1,14) = 8.14, p = .013, \eta^2 = .07, \eta_p^2 = .30$). Die Unterschiede in der Dozierendenorientierung und die Interaktion zwischen Dozierendenorientierung und Kohorte sind dabei nicht signifikant ($F(1,14) = .03, p = .87, \eta^2 = .001, \eta_p^2 = .001$ bzw. $F(1,14) = .44, p = .52, \eta^2 = .03, \eta_p^2 = .03$). Ebenfalls nicht signifikant ist der Haupteffekt der Kohorte ($F(1,14) = 3.36, p = .09, \eta^2 = .04, \eta_p^2 = .19$). In beiden Kohorten bezieht sich der signifikante Unterschied in der Studierendenorientierung dabei auf einen Anstieg und der nicht-signifikante Unterschied in der Dozierendenorientierung auf ein Absinken im Laufe der Ausbildung (siehe Tabelle 3). Die signifikante Interaktion zwischen Studierendenorientierung und Kohorte kann in dem Kontext so interpretiert werden, dass die Studierendenorientierung in der dritten Kohorte noch stärker angestiegen ist als in der zweiten. Die signifikante Interaktion zwischen den Veränderungen von Studierenden- und Dozierendenorientierung ist vor dem empirischen und theoretischen Hintergrund verständlich – die Konstrukte korrelieren leicht negativ (Prosser & Trigwell, 2006). Besondere Bedeutung hat hierbei der gefundene signifikante Haupteffekt einer erhöhten Studierendenorientierung über den Verlauf der Ausbildung hinweg. Der Vergleich der

gefundenen Effektstärken η^2 bzw. η_p^2 weist darauf hin, dass es sich hierbei um einen substantiellen Effekt handelt.

Tabelle 3: Studierenden- und Dozierendenorientierung

	MZP I		MZP II	
	(Startworkshop)		(Abschlussevaluation)	
	M	SD	M	SD
Studierendenorientierung				
2. Kohorte	4.11	.37	4.39	.23
3. Kohorte	3.32	.48	3.83	.67
Dozierendenorientierung				
2. Kohorte	3.27	.54	3.20	.73
3. Kohorte	3.65	.43	3.06	.72

Anmerkung:

Jeweils Skala von 1 „stimme überhaupt nicht zu“ bis 5 „stimme völlig zu“

4.4 Kompetenzzuwachs (Abschlussevaluation)

In allen drei Kohorten haben die Teilnehmer*innen nach eigener Einschätzung im Mittel einen bedeutsamen Kompetenzzuwachs in den Bereichen *Didaktik, Rhetorik & Präsentation, Methodik, Gruppenführung, Diskussionsleitung* und *Train-the-Trainer* über die Ausbildung hinweg erfahren (siehe Tabelle 4). Analysen der Unterschiede zwischen den Kohorten deckten dabei keine signifikanten Unterschiede auf (einfaktorielle ANOVA; alle $p > .05$).

Tabelle 4: Subjektive Kompetenzsteigerung

	1. Kohorte		2. Kohorte		3. Kohorte	
	M	SD	M	SD	M	SD
Didaktik	4.22	0.67	4.25	0.89	4.30	0.95
Rhetorik & Präsentation	3.44	0.89	4.00	0.53	4.10	0.99
Methodik	4.11	1.05	4.50	0.76	4.30	0.67
Gruppenführung	3.78	0.67	4.13	0.83	4.00	1.15
Diskussionsleitung	3.44	0.89	4.00	0.76	3.20	1.03
Train-the-Trainer	4.33	1.00	4.50	0.53	4.40	0.70

Anmerkung:

Jeweils Skala von 1 „stimme überhaupt nicht zu“ bis 5 „stimme völlig zu“

4.5 Fremdeinschätzung Performanz (Hospitationen)

Die Ergebnisse der Hospitationen sind trotz standardisiertem Beobachtungsbogen zu spezifisch und unsystematisch, um allgemeine Aussagen über das Programm machen zu können. Im Kontext der Evaluation diente die Hospitation vor Beginn der Ausbildung zunächst dazu, die bereits vorhandenen Kompetenzen der Teilnehmer*innen einzuschätzen und die Ausbildungsinhalte entsprechend anpassen zu können. Des Weiteren wurden die Bögen von den Trainer*innen dazu genutzt, um den Teilnehmenden ein individuelles Feedback zu geben. Durch die unterschiedlichen Settings der Beobachtungen (vor der Ausbildung in einer Lehrveranstaltung vs. während der Ausbildung in einer Tutorenschulung) lassen sich zudem keine objektiven Vergleiche zwischen den Messzeitpunkten ziehen. Die offenen Antworten aus der Abschlussevaluation zeigen aber deutlich, wie wertvoll dieses Instrument für die Teilnehmer*innen selbst ist. Mehrfach wird die Experten Hospitation als *(Lehr-)Format mit dem größten Mehrwert hinsichtlich der Durchführung eigenen Schulungen* genannt. Durch die konkreten Hospitationsrückmeldungen kann die eigene Lehre kritisch reflektiert und in Folge dessen verbessert werden. Im Rahmen einer weiteren Überarbeitung des Hospitationssystems wird versucht, dem Spagat zwischen individuellem Feedback und objektivierbarer Beobachtung der Performanz möglichst gerecht zu werden.

4.6 Transfererfolg (Tutorenschulungen)

Um den Transfererfolg zu messen, wurden die aggregierten Evaluationsdaten von 14 Tutorenschulungen mit insgesamt 96 abgegebenen Fragebögen ausgewertet. Es zeigt sich, dass die geschulten Tutor*innen von einer im Mittel hoch ausgeprägten Kompetenzsteigerung, v.a. in den Bereichen *Didaktik* und *Methodik*, berichten (siehe Tabelle 3). In den Bereichen *Rhetorik & Präsentation* und *Gruppenführung* ist immer noch eine mittlere subjektive Kompetenzsteigerung messbar, wobei diese nicht allzu stark ausgeprägt ist. Die allgemeine Zufriedenheit mit der Organisation und den Inhalten der Schulung sowie mit den Ausbildenden ist dabei sehr hoch.

Tabelle 3: Transfererfolg: Kompetenzsteigerung und Zufriedenheit

	M	SD
Kompetenzsteigerung „Didaktik“	4.31	.72
Kompetenzsteigerung „Rhetorik & Präsentation“	3.67	.97
Kompetenzsteigerung „Methodik“	4.56	.68
Kompetenzsteigerung „Gruppenführung“	3.67	.93
Zufriedenheit mit Inhalten und Organisation	4.58	.49
Zufriedenheit mit Ausbildenden	4.78	.39

Anmerkung:

Jeweils Skala von 1 „stimme überhaupt nicht zu“ bis 5 „stimme völlig zu“

Darüber hinaus hatte die Tutorenschulung offenbar auch einen Einfluss auf die Motivation der Tutor*innen: Die berichtete Motivation, selbst Tutorien durchzuführen, war nach der Schulung signifikant höher als zuvor ($M = 4.69$ bzw. $M = 4.03$; $t(86) = 7.46$, $p < .001$, t-Test für abhängige Stichproben). Der entsprechende Effekt ist als groß einzustufen (Cohens $d = .87$; Cohen, 1992). Zur besseren Vergleichbarkeit wurde die Effektstärke für unabhängige Stichproben berechnet.

4.7 Zusammenfassung

Es lassen sich folgende, grundlegende Erkenntnisse aus den Ergebnissen der Evaluation ableiten: Es zeigt sich für die zweite und dritte Kohorte ein signifikantes Ansteigen der Studierendenorientierung zwischen Startworkshop und Abschlussevaluation. Außerdem ist ein signifikanter Zuwachs an Wissen über zentrale Konzepte der Tutorienarbeit sowie eine subjektiv empfundene Zunahme an hochschuldidaktischer Kompetenz feststellbar. Diese Ergebnisse sprechen zusammengenommen für relevante intrapsychische Veränderungen, die sich im Laufe der Ausbildung ergeben haben. Zusätzlich reflektieren hohe Mittelwerte in den Zufriedenheitsratings der drei Kohorten die Akzeptanz der Ausbildung. Jenseits der Veränderung der Tutorenausbilder*innen an sich zeigen sich auch Transfererfolge bei den Tutor*innen – vorwiegend in den didaktischen und methodischen Kompetenzbereichen. Der multiplikative Effekt der Ausbildung wird deutlich, wenn man die Zahl der geschulten Tutoren und Tutorinnen betrachtet: Durch die Tutorenausbilder*innen der ersten drei Kohorten wurden bereits über 600 Tutoren und Tutorinnen ausgebildet.

5 Limitationen

Die Aussagekraft der Programmevaluation von TutorPlus wird jedoch auch von einigen Faktoren limitiert. Grundsätzlich geraten die klassischen Evaluationsansätze bei innovativen Programmen im Bildungsbereich an ihre Grenzen, da es schwierig ist, oftmals bewusst allgemein gehaltene Zielstellungen von Bildungsprogrammen zu operationalisieren und Ursache-Wirkungszusammenhänge in komplexen und dynamischen Zusammenhangsgefügen abseits von (quasi-)experimentellen Designs nachzuweisen. Das bisherige Vorgehen, durch bewusst explorativ angelegte Erhebungen Wirkannahmen darüber, welche Veränderungen wie durch TutorPlus hervorgerufen werden, zu untersuchen, soll v.a. zwei Dinge ermöglichen: Ad hoc Modifizierungen am Programm vornehmen zu können sowie eine aussagekräftige Datengrundlage für die Entwicklung einer Programmtheorie (vgl. hierzu im Überblick Giel, 2013) bereitzustellen.

In der konkreten Umsetzung der bisherigen Evaluation zeigen sich wiederum auch folgende Limitationen: Zum einen ist es nicht einfach, in jeder Kohorte alle Teilnehmenden zum Ausfüllen der mittlerweile recht umfangreichen Evaluationsbögen zu motivieren. Die Rücklaufquote lag im pen- & paper-Format dabei zwar immer bei 100% der Anwesenden, bei Online-Umfragen jedoch eher zwischen 65% und 85%, was im Vergleich mit üblichen Rücklaufquoten aber immer noch akzeptabel erscheint. Zumindest gibt es keine Hinweise dafür, dass ein selektiver drop-out vorliegt.

Diskutiert werden kann auch die grundsätzliche Erhebungsmethode des Fragebogens. Dieser basiert auf subjektiven Einschätzungen der Teilnehmenden – in der Evaluation kann folglich nur die subjektive Einschätzung des Programmes und seiner Wirkung abgebildet werden. Dass verschiedene Personen in den Erhebungen aber meist zu sehr ähnlichen Urteilen kommen, lässt auf einen intersubjektiven Konsens schließen. Im Kontext der Erfassung von Wissen lässt sich im Sinne einer Objektivierung über die Möglichkeit von Wissenstests nachdenken.

6 Ausblick

Die hohe Zufriedenheit mit der Ausbildung offenbart nur wenige Schwachstellen des Programms. Dennoch können durch die Evaluation fortlaufend Implikationen für die Verbesserung des Programms formuliert werden. Es zeigte sich beispielsweise, dass es in der Kommunikation zwischen Teilnehmerinnen und dem TutorPlus-Team noch Optimierungsbedarf gibt.

Die Evaluation deckte vor allem strukturelle Probleme auf: Die fakultätsbezogene Unterstützung wird von den Tutorenausbilder*innen bisher als zu gering wahrgenommen. Dabei stellt sich die Frage, wie man die institutionelle Verankerung sicherstellen kann, sodass die Tutorenausbilder*innen mehr Mittel, Zeit und Unterstützung von der Fakultät erhalten. Ein weiteres strukturelles Problem des Programms ist die Fluktuation von Tutorenausbilder*innen – einige TutorPlus-ausgebildete Personen sind inzwischen an anderen Universitäten bzw. in anderen Arbeitskontexten beschäftigt. Dass Mitarbeiter*innen der Universität meist nur befristete Verträge erhalten, ist eine übergeordnete Problematik, die maßgeblich dazu beiträgt.

Im Rahmen einer stetigen Weiterentwicklung der Evaluation wurden und werden in einem iterativen Prozess stets Änderungen und Ergänzungen vorgenommen. Beispielsweise wurde die oben genannte zentralisierte Evaluation von Tutorenschulungen eingeführt, die die selbständig entworfenen Evaluationen der Tutorenausbilder ablöste. Zudem wurde der ATI zur Messung der Lehreinstellung integriert und der Beobachtungsbogen für Hospitationen wird aktuell, wie an anderer Stelle beschrieben, überarbeitet. Weiterhin wurden im Rahmen der *Success Case Method* (Brinkerhoff, 2003) leitfadengestützte Interviews geführt, deren Auswertungen sowohl in die Entwicklung der bereits erwähnten Programmtheorie einfließen, als auch Aufschluss hinsichtlich der Nachhaltigkeit des Programms geben sollen. Zu diesem Zwecke wurde auch eine schriftliche Follow-Up-Erhebung durchgeführt, die aktuell ausgewertet wird. Grundsätzlich werden die einzelnen Erhebungsinstrumente laufend und in enger Absprache mit den Trainer*innen des Teams weiterentwickelt. Um den Erfolg der Ausbildung weiter zu überprüfen, ist ein Tutorienvergleich geplant. Das heißt, Tutorien von Tutor*innen, die von TutorPlus-Teilnehmer*innen geschult wurden, mit Tutorien von Tutor*innen, die nicht von TutorPlus-Teilnehmer*innen geschult wurden, zu vergleichen. Nur so lässt sich letztendlich (zumindest annäherungsweise) feststellen, ob die angestrebte Verbesserung der Lehre bei der Zielgruppe der Studierenden ankommt.

Unter Bezugnahme auf Berliner (2002), der die Forschung im Bereich der Erziehungs- und Bildungswissenschaften als „The Hardest Science of All“ (S. 18) bezeichnet, sieht sich auch das Evaluationskonzept von TutorPlus vielen Herausforderungen, welche die Evaluation von Bildungsprogrammen mit sich bringt, gegenüber. Die klassischen Zugänge über die Zielstellungen oder Vergleichs- und Kontrollgruppen sind für die Evaluation von TutorPlus wenig geeignet. Dem soll mittelfristig mit der Entwicklung eines belastbaren theoriebasierten Evaluationskonzept, ergo einer Programmtheorie, begegnet werden – es soll nicht nur eine Aussage darüber getroffen werden, ob das Programm funktioniert, sondern es ist vielmehr von Interesse, wie und warum es für wen funktioniert oder nicht funktioniert.

Literatur

- Antosch-Bardohn, J. & Primus, N. (2015). Gute Lehre wächst von unten! Eine multiplikatorenbasierte Tutorenausbilder-Ausbildung. In B. Berendt, A. Fleischmann, N. Schaper, B. Szczyrba & J. Wildt (Hrsg.), *Neues Handbuch Hochschullehre* (139-150). Berlin: Raabe.
- Balzer, L. (2005). Wie werden Evaluationsprojekte erfolgreich. *Ein integrierender theoretischer Ansatz und eine empirische Studie zum Evaluationsprozess*. Landau: Verlag Empirische Pädagogik.
- Berliner, D.C. (2002). Educational Research: The Hardest Science of All. In: *Educational Researcher*, 31(8), 18–20.
- Braun, E. & Hannover, B. (2008). Zum Zusammenhang von Lehr-Orientierung und Lehrgestaltung von Hochschuldozierenden und subjektivem Kompetenzzuwachs bei Studierenden. *Zeitschrift für Erziehungswissenschaft, Sonderheft 9*, 277-291.
- Brinkerhoff, R. (2003). *The success case method: Find out quickly what's working and what's not*. San Francisco: Berrett-Koehler Publishers.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Giel, S. (2013) *Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen*. Münster (u.a.): Waxmann.
- Prosser, M. & Trigwell, K. (2006). Confirmatory factor analysis of the Approaches to Teaching Inventory. *British Journal of Educational Psychology*, 76, 405-419.
- Stockmann, Reinhard (2002). Was ist eine gute Evaluation (CEval-Arbeitspapiere; 9). Abgerufen auf der Website des Centrum für Evaluation der Universität des Saarlands: http://www.ceval.de/modx/fileadmin/user_upload/PDFs/workpaper9.pdf
- Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review*, 16(4), 409-424.

Autor*innen

Tanja, Sczepanski, Universität Hamburg, Projekt ProfaLe Qualitätsoffensive Lehrerbildung, Hamburg, Deutschland; Email: tanja.szczepanski@uni-hamburg.de

Johannes, Hauser, LMU München, Programm PROFIL, München, Deutschland; Email: hauser@profil.lmu.de

Anna, von Werthern, LMU München, Programm PROFIL, München, Deutschland; Email: vonwerthern@profil.lmu.de

Dr., Andreas, Hendrich, LMU München, Programm PROFIL, München, Deutschland; Email: andreas.hendrich@lmu.de



Zitiervorschlag: Sczepanski, T., Hauser, J., von Werthern, A. & Hendrich, A. (2016). TutorenausbilderInnen als Multiplikatoren für gute Lehre – erste Evaluationsergebnisse des Qualifikationsprogramms TutorPlus. *die hochschullehre*, Jahrgang 2/2016, online unter: www.hochschullehre.org