

# Messwiederholungspläne als Instrumente einer vergleichenden Evaluation von problemorientierter und traditioneller Lehre an der Medizinischen Fakultät der TU Dresden

*Bärbel Bergmann, Doreen Eisfeldt, Sabine Lanadio*

*Technische Universität Dresden, Institut für Allgemeine Psychologie, Biopsychologie und Methoden der Psychologie*

## 1. Einführung

Die zunehmende Beschleunigung der Wissensentwicklung macht traditionelle Auffassungen vom Lehren, als einem Transport für fertige Systeme von Wissen die von Lernenden zu rezipieren sind, obsolet. Mit der kürzer werdenden Halbwertszeit des Wissens wird diese Beschleunigung ausgedrückt und grob mit fünf Jahren quantifiziert (Giarini/Liedtke 1998). Für innovative Bereiche, zu denen Medizin, Medizintechnik und Pharmakologie gehören, wird die Halbwertszeit des Wissens mit weniger als zwei Jahren angegeben. Wissen wird zur wettbewerbsentscheidenden Ressource. Deshalb ist Wissensarbeit mit dem Integrieren, Bewerten und Bewahren von Wissen so zu erlernen, dass sie effizient funktioniert (Willke 1998). Diese Lernaufgabe sollte bereits während des Studiums in Angriff genommen werden. Die Beschleunigung der Wissensentwicklung hat zur Folge, dass von Seiten des Bildungssystems eine didaktische Aufbereitung relevanten Berufswissens nicht mehr zeitnah bewältigt werden kann, sondern dass aktuelles Wissen durch selbst organisiertes Lernen anzueignen ist. Aber nicht nur die in immer kürzer werdenden Zeitabschnitten erfolgende Wissensveränderung stellt traditionelle Lehr- und Lernformen in Frage. Auch die Berufspraxis verändert sich. Wissen ist in immer verschiedener werdenden Anwendungssituationen einzusetzen. Dazu muss es verändert und angepasst werden. Das geschieht in der Medizin wie auch in anderen Berufen mehrheitlich in kooperativer Arbeit, z.T. in interdisziplinärer Kooperation. In dieser erfolgt eine Aufbereitung von Wissen für aktuelle Probleme, aber auch eine Weiter- und Neuentwicklung von Wissen, für die durch die Ausbildung eine Vorbereitung erforderlich ist, die traditionelle Lehrkonzepte nur ungenügend leisten. Für viele neuartige Probleme ist Wissen beim Einzelnen nicht voll verfügbar. Relevantes Wissen kann verteilt in mehreren Personen existieren, zum Teil ist es aber auch erforderlich, es aktuell zu beschaffen oder zu erzeugen. In einer mehrheitlich als Gruppenarbeit erfolgenden Berufspraxis ist Wissen deshalb zu kommunizieren und in Gruppen auf neue Problemstellungen zu transferieren. Die Kompetenz zum Kooperieren in Gruppen muss erlernt werden und ebenso die im Berufsalltag wichtige

Kompetenz zum Wissensmanagement, welches die Befähigung zur strukturierten Kommunikation individuellen Wissens, zur Schaffung von Vertrauen als Basis für den individuellen Austausch von Wissen einschließt. Deshalb wird angestrebt, diese Fähigkeiten bereits während des Studiums zu trainieren.

Forderungen und Vorschläge für eine Verbesserung der Ausbildung von Medizinern sind zahlreich. Reuschler (2002) sieht in Modellen der Vereinigung von ärztlicher Tätigkeit und Lernen den Lösungsansatz und orientiert sich dabei an dem dualen Ausbildungssystem der deutschen Wirtschaft. Mit problemorientierten Lehrstrategien wird durch eine stärkere Verankerung der Lehre an konkreten praxisnahen Fallbeispielen versucht, die Kompetenz zu selbst organisiertem Lernen zu trainieren und für ihre Entwicklung zu motivieren. Die geschilderten Probleme der beschleunigten Wissensentwicklung und der Veränderung in der Berufspraxis haben dazu geführt, dass für die Medizinausbildung an verschiedenen Universitäten, auch an der Technischen Universität Dresden, eine Veränderung der Ausbildung angestrebt wird. Eine Reformierung der Ausbildung wird mit einer Evaluation des innovativen Ausbildungsprogramms verbunden. Eitel (2002: 2) verweist darauf, „dass eine rationelle und professionelle Gestaltung der Lehre im Allgemeinen und des Curriculums im Besonderen davon abhängt, inwieweit die Entscheidungen hierzu empirisch begründet bzw. mit Argumenten bzw. Erkenntnissen aus der Ausbildungsforschung unterlegt werden“. Das wichtigste Ziel dabei ist die Optimierung des neuen Programms. Die Entwicklung eines neuen Lehrprogramms ist eine sehr komplexe Gestaltungsaufgabe, die nicht auf Anhieb perfekt gelingt. Deshalb sind Stärken und vor allem noch bestehende Schwachpunkte festzustellen und dafür Verbesserungsmöglichkeiten aufzuzeigen.

Eine Lehrevaluation erfolgt überwiegend in Form der Bewertung von Lehrveranstaltungen durch Studierende (Spiel/Gössler 2001: 10). Dieser methodische Weg wird jedoch auch sehr kritisch gesehen. Kromrey (1994) verweist darauf, dass Akzeptanz- und Zufriedenheitsurteile zweifelhafte Indikatoren für gute Lehre sind. Die Antworten hängen in starkem Maße von der Beliebtheit oder Unbeliebtheit des behandelten Stoffes ab und sie werden durch die Zusammensetzung der Lehrveranstaltungsteilnehmer (freiwillig, unfreiwillig) beeinflusst. Allerdings ist diese Kritik auch von der Art des verwendeten Fragebogens abhängig. Es ist durchaus möglich, bei der Erhebung bewertender Aussagen zwischen Akzeptanz- und Lernerfolgsabschätzungen zu verschiedenen Facetten zu unterscheiden. Aber Verzerrungsmöglichkeiten bestehen bei dem methodischen Zugang über subjektive Einschätzungen. Studenten sind Alltagsevaluatoren. In einer späteren Arbeit beschreibt Kromrey die mit der Evaluation in Form der Erhebung studentischer Bewertungen verbundenen Fehlschlüsse wie folgt (Kromrey 2001: 42): „Jeder einzelne von ihnen bewertet *irgendetwas* (was er mit dem in der Frage angesprochenen Sachverhalt ad hoc assoziiert) *irgendwie* („alles in allem“ oder „aus aktueller Erfahrung“ oder „und Blick auf das Wesentliche“ oder ...) unter *irgendwelchen* Gesichtspunkten (Nutzen für sein Studium oder vermuteter Nutzen für den angestrebten Beruf oder aktuelles persönliches Interesse oder abstrakt-verallgemeinertes Interesse *der* Studierenden oder ...). Die Bedeutung der im standardisierten Erhebungsbogen gegebenen Antworten ist so nicht mehr rekonstruierbar.“

Damit sind unterschiedliche Bezugssysteme und Bewertungskriterien bei der Generierung bewertender Urteile über Lehrveranstaltungen als eine Quelle für mögliche Fehlschlüsse aus so gewonnenen Aussagen benannt.

Diese Probleme lassen sich mit speziellen methodischen Anordnungen bei der Durchführung von Evaluationen verkleinern.

Ein erster Schritt besteht im gleichzeitigen Einholen von bewertenden Urteilen zu einem Vergleichsprogramm. Ein innovatives Lehrprogramm kann mit dem gut implementierten traditionellen Lehrprogrammen verglichen werden. Damit erhalten Bewertungsaussagen zu einem innovativen Programm einen Vergleichsmaßstab. Es wird sichtbar, in welchen Punkten das innovative Programm zu besseren Bewertungen führt und wo das innovative Programm in der Wahrnehmung der Programmteilnehmer hinter dem traditionellen Programm zurückbleibt.

Ein zweiter Schritt kann sich dem Problem der Kontrolle personeller Störvariablen in der Form der Varianz in der Bewertung zugrunde liegenden subjektiven Kriterien widmen. Dafür können Within-Designs gewählt werden, bei denen die meist nicht klar explizierbaren subjektiven Bewertungskriterien für die zu vergleichenden Varianten konstant gehalten werden, indem die gleichen Personen das innovative und das traditionelle Programm bewerten. In Designs mit wiederholten Messungen wird deshalb ein Weg gesehen, die Interpretationsprobleme bewertender Urteile der Studierenden zu verkleinern. Am Beispiel der Evaluation neuer Lehrstrategien in der Mediziner Ausbildung an der TU Dresden sollen Messwiederholungspläne als eine Strategie zur Evaluation innovativer Lehrprogramme mit dem Ziel der Optimierung dieser Programme vorgestellt werden.

## 2. Ziel der reformierten Ausbildung

Die Medizinische Fakultät Carl Gustav Carus der Technischen Universität führt seit 1999 neue Lehrmethoden in der Mediziner Ausbildung in Form des Dresdner integrativen problem-/praxisorientierten Lernens (DIPOL) ein (Dieter 2001; Bergmann/Dieter/Nitsche 2002). Das Ziel besteht darin, die bisher durch die einzelnen Fachdisziplinen getragenen Studieninhalte vertikal und horizontal miteinander zu verzahnen. Als Partner wurde die Harvard Medical School (Boston, USA) gewonnen, die seit 1985 problemorientiertes Lernen realisiert. In Dresden wird jedoch nicht das Lehrmodell der Harvard Medical School kopiert sondern ein Hybridmodell realisiert. Das ist eine Integration von problemorientierter und traditioneller Lehre, die so durchgeführt wird, dass inhaltlich auf Tutorien als dem Kernstück problemorientierter Lehre abgestimmte Vorlesungen, Seminare und Praktika zusammen einen Kurs bilden.

Das innovative Lehrprogramm DIPOL ist damit ein sehr komplexer Evaluationsgegenstand, der allgemein als ein neues Curriculum für den vorklinischen und den klinischen Studienabschnitt beschrieben werden kann. Das Curriculum ist in inhaltliche Lehrgebiete gegliedert. Die Lehrinhalte sind in Kursen des Problemorientierten Lernens (POL-Kurse) organisiert. Innerhalb der POL-Kurse sind die Tutorien das innovative Element. Tutorien sind Kleingruppen von acht bis zehn Studenten, die sich mit der Unterstützung eines Tutors anhand schriftlich präsentierter Patientenfälle zu den Inhalten des Kurses Wissen und Fähigkeiten erarbeiten. Der Tutor unterstützt die Planung und Durchführung der Lernaktivität und strukturiert den Gruppenprozess. Er ist kein Seminarleiter sondern ein Moderator. Die Tutoren werden durch eine Schulung auf die Unterstützung des problemorientierten Lernens

vorbereitet. In den Tutorien stellen die Lehrfälle, an denen problemorientiertes Lernen realisiert wird, wichtige Bausteine dar. Eine formative Evaluation mit dem Ziel der Optimierung eines innovativen Programms betrifft alle diese Ebenen, d.h. die Ebene der Lehrfälle, die Ebene der POL-Kurse und die Ebene des Curriculums.

Für den in diesem Beitrag vorgenommenen Vergleich traditioneller und problemorientierter Lehre werden als Analyseebene die inhaltlichen Lehrgebiete gewählt. In POL-Kursen organisierte Lehrinhalte werden mit traditionell gelehrteten inhaltlichen Gebieten verglichen. Mit der Implementierung von DIPOL sind ehrgeizige Ziele verbunden. Konstruktivistische und situierte Lerntheorien bilden das theoretische Fundament problemorientierten Lernens (vgl. Reinmann-Rothmeier/Mandl 1997; Gräsel 1997). Sie begründen, dass auf ihnen basierende Lehrstrategien zu einer höheren Befähigung zum Umgang mit Fachwissen, also bei Medizinern zu einer höheren Befähigung bei der selbständigen Entwicklung von Diagnose- und Behandlungsmethoden und zu Verbesserungen im Wissenstransfer führen. Durch ihre Umsetzung in problemorientiertes Lernen soll gleichzeitig ein Training in kooperativer Arbeit und damit auch im Managen verteilten Wissens erreicht werden. So soll die Überbetonung systematischer und theoretischer Wissensvermittlung in der traditionellen Ausbildung überwunden und gleichzeitig eine Schulung im deduktiven Denken und praktischer Fähigkeiten und Fertigkeiten durch das Organisieren von Gelegenheiten zu Übungsmöglichkeiten erreicht werden. Durch die problemorientierte Lehre wird zusätzlich angestrebt, mittels des Weckens von Neugier, von Spaß und der Organisation des Lernens so, dass subjektiv ein Lernfluss erlebt wird, intrinsische Motivation als wichtige Voraussetzung für die angestrebte Befähigung zu selbst organisiertem Lernen zu erreichen. Der Umfang an systematischer Wissensvermittlung ist bei problemorientiertem Lernen zugunsten größerer Spielräume für die selbständige Aneignung und Integration von Wissen reduziert. In zusammenfassender Form werden bei der Einführung problemorientierter Lehre (POL) im Vergleich zu traditioneller Lehre die folgenden Lernresultate erwartet:

– Aneignung von Fachwissen	POL = TRAD
– Aneignung von Methoden des Umgangs mit Fachwissen	POL > TRAD
– Erfahrungen mit kooperativer Arbeit	POL > TRAD
– Motivation zu selbständigem Lernen	POL > TRAD
– Befähigung zum Transferieren	POL > TRAD

### 3. Evaluationsstrategie

#### 3.1 Evaluationskonzept

Die Evaluation des reformierten Medizinstudiums orientierte sich an dem Vier-Ebenen-Modell von Kirkpatrick (1998). Die vier Ebenen sind:

- *Reaktionen bzw. Akzeptanz*  
Die Zufriedenheit bzw. Akzeptanz der Teilnehmer mit verschiedenen Aspekten der Ausbildungsgestaltung wird erhoben.
- *Lernerfolg*  
Der Lernerfolg in Form angeeigneten Wissens, verbesserter Fähigkeiten bei der

Diagnose und Entwicklung von Behandlungskonzepten sowie bei der Realisierung kooperativer Arbeit und in Form von Einstellungen, z.B. in Form der Motivation zu selbständigem Lernen werden erfasst.

– *Transfer*

In dieser Ebene wird die Übertragbarkeit des Gelernten auf Berufsaufgaben festgestellt.

– *Resultate oder ultimate outcomes*

Die Messung der Ergebnisse von Ausbildungsprogrammen wird auf der Ebene der Organisation in für diese üblichen Kennziffern, die eine Weiterverarbeitung in Form von Kosten Nutzen-Analysen ermöglichen, vorgenommen.

Diese Ebene kann bei der Evaluation von DIPOL nicht realisiert werden.

### 3.2 Evaluationsdesign

Als Königsweg der Feststellung der Wirksamkeit von innovativen Programmen gelten randomisierte Versuchs-Kontrollgruppenpläne mit der vergleichenden Feststellung von Effekten des innovativen und eines Standardprogramms (Metzler/Krause 1997). Dieses ideale Design ist bei der Evaluation von DIPOL nicht realisierbar, weil die POL-Kurse jahrgangsweise eingeführt werden und es deshalb unmöglich ist, die Studenten per Zufall der problemorientierten und der traditionellen Lehrform für vergleichbare Lehrinhalte zuzuweisen.

In der Phase der Einführung problemorientierter Lehre ist aber nicht das gesamte Studium auf die neue Lehrform umgestellt. Traditionelle und problemorientierte Lehrphasen wechseln einander ab. Dieser Tatbestand wurde genutzt für die Kontrolle personeller Variablen, die als Varianz in subjektiven Bezugswerten bei der Abgabe von bewertenden Aussagen zu den Lehrformen wirksam werden. Messwiederholungspläne wurden realisiert, in denen Aussagen zu Reaktionen, zu Einschätzungen des Lernerfolgs und der Transfererwartung bei problemorientierter und traditioneller Lehre bei denselben Studenten erhoben wurden. Allerdings besteht bei diesem Vorgehen ein Problem. Mit der Lehrform variiert auch der Lehrinhalt. Diese Konfundierung von Lehrform und Lehrinhalt kann Befunde, mit denen Aussagen über die Lehrform angestrebt werden, verzerren. Eine Lösung dieses Problems wird darin gesehen, für beide Lehrformen Messungen bei verschiedenen Inhalten vorzunehmen und so durch eine Aggregation der Resultate über verschiedene Lehrinhalte die inhaltsbedingten Effekte auszumitteln. In allgemeiner Form gibt die Tabelle 1 den Untersuchungsplan an.

*Tabelle 1:* Allgemeines Schema eines Messwiederholungsdesigns für den Vergleich von traditioneller und problemorientierter Lehrstrategie

Teilnehmer	Lehrform					
	Traditionelle Lehre			Problemorientierte Lehre		
	Abschnitt A, Abschnitt B, Abschnitt C			Abschnitt A, Abschnitt B, Abschnitt C		
Studierender 1	Kriterien für – Reaktionsebene – Lernebene analog analog – Transfer			Kriterien für – Reaktionsebene – Lernebene analog analog – Transfer		
Studierender 2	„	„	„	„	„	„
•						
•						
Studierender n	„	„	„	„	„	„

In den vorzustellenden Untersuchungen wurde dieses Designschema mehrfach realisiert und zwar für Studenten der Zahnmedizin im 7. Fachsemester (Eisfeldt 2002). Hier wurde die traditionelle Lehre des vorausgehenden Fachsemesters sowie die im 7. Semester stattfindende traditionelle Lehre mit der problemorientierten Lehre im Kurs „Zahnheilkunde 1“ und dem Kurs „Zahn-, Mund-, Kieferheilkunde“ verglichen. Den Untersuchungsplan enthält Tabelle 2.

*Tabelle 2:* Messwiederholungsdesign für den Vergleich traditioneller und problemorientierter Lehre bei Zahnmedizinern im 7. Fachsemester

Studierende	Lehrform			
	Traditionelle Lehre		Problemorientierte Lehre	
	vorausgegangenes Fachsemester (Allg. Chirurgie, Biomathematik, Kieferorthopädie u.a.)	Gleiches Fachsemester (Arbeitswissenschaft, Zahnerhaltungskunde, Kieferchirurgie u.a.)	POL-Kurs „Zahnheilkunde“	POL-Kurs „Zahn-, Mund-, Kieferheilkunde“
1	Lernen Transfer	Lernen Transfer	Lernen Transfer	Lernen Transfer
2	„	„	„	„
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
44	„	„	„	„

In dieser Studie wurden nur Kriterien zum Lernerfolg und zur Transfererwartung erhoben.

Eine weitere Studie wurde bei Medizinstudenten des 8. Fachsemesters realisiert (Lanadio 2002). Die traditionelle Lehre in den Fächern Chirurgie und Innere Medizin wurde mit dem POL-Kurs Onkologie verglichen. Den Versuchsplan gibt die Tabelle 3 an.

*Tabelle 3:* Messwiederholungsdesign für den Vergleich von traditioneller Lehre (Chirurgie, Innere Medizin) und problemorientierter Lehre im Kurs Onkologie für Medizinstudenten im 8. Fachsemester

Studierende	Lehrform	
	Traditionelle Lehre Chirurgie, Innere Medizin	POL-Kurs Onkologie
1	Reaktionen Lernen Transfer	Reaktionen Lernen Transfer
2	"	"
.	.	.
.	.	.
.	.	.
62	"	"

### 3.3 Evaluationskriterien

Die vergleichende Evaluation orientiert sich in dieser Arbeit an dem Kirkpatrick'schen Evaluationsmodell, wobei die vierte Ebene nicht realisiert werden kann. Es sind Kriterien zu Reaktionen auf die Lehrformen, zum Lernerfolg und zum Transfer zu erheben. Als methodischer Zugang wird die Befragung Studierender gewählt. Für den Lernerfolg sind Ergebnisse der Leistungsprüfungen in den klinischen Lehrabschnitten im Prinzip ebenfalls zu nutzende Kriterien. In dieser Arbeit wurde auf diesen methodischen Zugang verzichtet, weil die Prüfungsinhalte nicht mit dem Umfang in dieser Arbeit verglichener Lehrabschnitte kompatibel sind. Es wurden standardisierte Fragebögen genutzt. Sie sind gemäß des verwendeten Evaluationskonzepts und der mit der innovativen Lehrform angestrebten Ziele rational konstruiert, faktorenanalytisch überprüft und es sind psychometrische Gütekriterien bestimmt (Nitsche 2003). Für die einzelnen Evaluationsebenen sind die folgenden Skalen entwickelt:

#### *Reaktionsebene*

- Akzeptanz der organisatorischen Rahmenbedingungen
- Allgemeine Einstellung zu problemorientiertem Lernen

#### *Lernebene*

- Aneignung von Wissen
- Aneignung von Fähigkeiten
- Motivation zu selbständigem Lernen
- Erfahrungen in kooperativer Arbeit

*Transfer*

## – Transfererwartung

Die Erhebung der Skala „Allgemeine Einstellung zum problemorientierten Lernen“ ist zunächst deskriptiv für die POL-Kurse interessant. Zusätzlich kann diese Skala genutzt werden, um das Ausmaß der Verzerrung bei den Einschätzungen verschiedener Kriterien des Lernerfolgs durch diese Personenvariable zu kontrollieren. Für die Datenerhebung wurde ein sechsstufiges Antwortformat gewählt, bei welchem Grade der Zielerreichung in drei Stufen erhoben wurden und auch eine fehlende Zielerreichung abgestuft erfasst wurde. Zur Veranschaulichung der vorzunehmenden Bewertung wurden den Antwortkategorien negative Zahlen zugeordnet, wenn eine fehlende Zielerreichung anzugeben war und positive Zahlen, wenn es um eine Bestätigung der Zielerreichung ging. Der Verzicht auf eine neutrale Mittelkategorie wird bewusst realisiert, um eine Akzentuierung der Evaluationsergebnisse zu sichern. Für die Auswertung erfolgte eine Umkodierung in die Ziffern 1 bis 6, so dass größere Zahlen eine bessere Zielerreichung abbilden. Die Art der Erhebung und die Kodierung für die Auswertung veranschaulicht das folgende Beispiel:

*Tabelle 4:* Antwortformat der Erhebung der Evaluationskriterien

Item	trifft überhaupt nicht zu	trifft überwiegend nicht zu	trifft eher nicht zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu
Ich habe umfassende Kenntnisse erworben, so dass ich sagen kann, ich beherrsche das Gelernte.	- 3	- 2	- 1	1	2	3
	0	0	0	0	0	0
	1	2	3	4	5	6

Umkodierung für die Verrechnung (Je größer die Zahl, desto besser die Zielerreichung)

Die Erhebungen fanden jeweils zum Abschluss des jeweiligen Lehrabschnitts statt.

### 3.4 Datenauswertung

Die Deskription der Evaluationsergebnisse erfolgte über die Berechnung von Mittelwerten und Standardabweichungen. Gemäß des verwendeten 6 stufigen Antwortformats weisen Mittelwerte größer als 3,5 auf eine tendenzielle Zielerreichung hin und geben umso bessere Zielerreichungen an je größer die Mittelwerte sind. Für den Vergleich der beiden Lehrformen werden für die Kriterien der Lernebene und die Transfererwartung multivariate Varianzanalysen mit Messwiederholung gerechnet und bei Vorhandensein eines signifikanten Effekts wurden die Unterschiede in den einzelnen Kriterien durch die Berechnung von Effektstärken quantifiziert. In der Studie der Medizinstudenten erfolgte zusätzlich eine Prüfung der Abhängigkeit der Selbsteinschätzungen in den verschiedenen Lernkriterien von der Allgemeinen

Einstellung zum problemorientierten Lernen. Dazu wurden Korrelationskoeffizienten berechnet.

## 4. Ergebnisse

### 4.1 Ergebnisse der Studie bei den Zahnmedizinstudenten

Die erhobenen Einschätzungen für die einzelnen Skalen bei den verschiedenen Abschnitten traditioneller Lehre wurden aggregiert. Ebenso wurden die Einschätzungen beider POL-Kurse zusammengefasst, so dass die gebildeten Werte für die jeweilige Lehrform stehen. Die Tabelle 4 enthält die Resultate.

*Tabelle 5:* Ergebnisse der Selbsteinschätzungen des Lernerfolgs und der Transfererwartung der Studenten der Zahnmedizin bei den verglichenen Abschnitten traditioneller und problemorientierter Lehre

Kriterien	Lehrform			
	Traditionelle Lehre		Problemorientierte Lehre	
	M	SD	M	SD
<b>Lernebene</b>				
– Wissen	4,27	0,43	4,30	0,81
– Fähigkeiten	3,90	0,57	4,24	0,72
– Motivation zu selbständigem Lernen	3,64	0,55	4,10	0,57
– Erfahrungen mit kooperativer Arbeit	4,01	0,49	4,29	0,58
<b>Transfer</b>				
– Transfererwartung	4,44	0,49	4,12	0,90

Die Werte bewegen sich um den Wert 4 für die traditionelle Lehre und etwas oberhalb von 4 für die problemorientierte Lehre. Das beschreibt eine tendenzielle Zielerreichung, zeigt aber auch noch erheblichen Verbesserungsspielraum an und drückt aus, dass die Studenten unmittelbar nach Absolvieren der Lehre noch keine gute Zielerreichung bestätigen. Die zusätzliche Betrachtung der Streuungen ist interessant. Sie sind durchgängig bei der problemorientierten Lehre größer. Das weist auf eine ambivalente Beurteilung dieser Lehrform hin, d.h. ein Teil der Studierenden bewertet die problemorientierte Lehre sehr gut, ein anderer Teil aber auch sehr kritisch. Die Beurteilungen reichen in Werte innerhalb von 3,5 hinein, bei denen auch eine tendenzielle Zielerreichung nicht mehr bestätigt werden kann. Ein Teil der Studierenden bestätigt das Erreichen der Lernziele nicht und das weist auf deutlichen Verbesserungsbedarf hin.

Die multivariate Varianzanalyse weist signifikante Unterschiede aus (Pillai Spur: 0,625;  $p=0,006$ ). Damit belegen die Resultate eine leichte Überlegenheit der problemorientierten Lehre gegenüber der traditionellen in den Selbsteinschätzungen des Lernerfolgs. Die Berechnung von Effektstärken (s. Tabelle 6) quantifiziert die Unterschiede für die Aneignung von Fähigkeiten als kleine Effekte, für die Befähigung zur Kooperation in Gruppen als mittlere, für die Motivation zu selbständigem Lernen als große Effekte. Für die Transfererwartung liegt ein hypothesenkonträrer

mittlerer Effekt vor. Hier haben die Studenten für die traditionelle Lehre die Transfererwartung besser beurteilt.

*Tabelle 6:* Darstellung der Effektstärkeunterschiede in den Selbsteinschätzungen verschiedener Kriterien zwischen problemorientierter und traditioneller Lehre in der Studie bei Zahnmedizinstudenten

	Asymptotische Signifikanz	Effektstärke
– Wissen	0,863	–
– Fähigkeiten	0,036	0,15
– Motivation zu selbständigem Lernen	0,002	1,04
– Erfahrungen mit kooperativer Arbeit	0,022	0,68
– Transfererwartung	0,036	-0,66

## 4.2 Ergebnisse der Studie bei den Medizinstudenten

Die Resultate der Einschätzungen der traditionellen Lehre (Chirurgie und Innere Medizin) und des POL-Kurses Onkologie durch Studenten des 8. Fachsemesters enthält die Tabelle 7. Hier sind auch die Befunde der Reaktionsebene dargestellt. Die Skala „Allgemeine Einstellung zu POL“ wurde auch bei den Erhebungen zur traditionellen Lehre eingesetzt, um die Beeinflussung der Urteile durch diese Einstellungsvariable nachträglich prüfen zu können.

*Tabelle 7:* Ergebnisse der Selbsteinschätzungen zu den verschiedenen Kriterien zu traditioneller und problemorientierter Lehre bei den vergleichenden Lehrabschnitten durch Medizinstudenten

Kriterien	Traditionelle Lehre		POL- Kurs Onkologie	
	M	SD	M	SD
<b>Reaktionsebene</b>				
– Akzeptanz der organisatorischen Rahmenbedingungen	4,17	0,98	3,43	0,95
– Allgemeine Einstellung zu POL	3,96	1,14	3,81	1,04
<b>Lernebene</b>				
– Wissen	3,99	0,80	3,44	0,95
– Fähigkeiten	3,90	0,70	3,71	0,99
– Motivation zu selbständigem Lernen	3,31	1,33	3,13	1,05
– Erfahrungen mit kooperativer Arbeit	3,92	0,84	4,08	0,70
<b>Transfer</b>				
– Transfererwartung	3,98	0,84	3,49	1,00

In dieser Studie sind die Einschätzungen zu den Kriterien generell nicht günstig. Die Mittelwerte liegen für die traditionelle Lehre mit einer Ausnahme (Motivation zu selbständigem Lernen) knapp über dem Wert von 3,5 und zeigen somit eine tendenzielle Zielerreichung an. Für den POL-Kurs Onkologie unterschreiten sogar die Mittelwerte bei vier Kriterien den Wert von 3,5. Schon die Art der organisatorischen Implementierung dieses Kurses erreicht wenig Akzeptanz. Die zusätzliche

Berücksichtigung der Streuungen macht klar, dass ein großer Teil der Studenten bei vier Kriterien mehr als die Hälfte die Zielerreichung nicht bestätigt. Damit ist drastischer Verbesserungsbedarf angezeigt.

Die multivariate Varianzanalyse weist signifikante Unterschiede zugunsten der traditionellen Lehrform aus. Die Effektunterschiede sind aber klein und in den Einzelvergleichen nur für die Skala Wissen ( $p=0,001$ , Eta Quadrat 0,181) und die Transfererwartung ( $p=0,004$ , Eta Quadrat 0,133) signifikant. Diese Studie zeigt, dass die mit dem POL-Kurs Onkologie angestrebten Ziele in der evaluierten Version noch nicht erreicht sind und erheblicher Verbesserungsbedarf besteht.

In dieser Studie wurde auch der Frage nach Zusammenhängen zwischen der Allgemeinen Einstellung zum problemorientierten Lernen und den Selbsteinschätzungen des Lernerfolgs und der Transfererwartung nachgegangen. Tabelle 8 enthält die Korrelationskoeffizienten sowohl für die traditionelle Lehre in den Fächern Chirurgie und Innere Medizin als auch für den POL-Kurs Onkologie. Für die traditionelle Lehre sind die Korrelationskoeffizienten bis auf zwei Ausnahmen nicht signifikant. Für den Kurs des problemorientierten Lernens sind die Korrelationskoeffizienten durchgängig signifikant und sie drücken mehrheitlich einen starken Zusammenhang aus.

*Tabelle 8: Zusammenhänge zwischen der Skala „Allgemeine Einstellung zum problemorientierten Lernen“ und den Selbsteinschätzungen des Lernerfolgs und der Transfererwartung (Lanadio 2002)*

Kriterien	Traditionelle Lehre (Chirurgie, Innere Medizin)		POL- Kurs Onkologie	
	Korrelations- koeffizient	p	Korrelations- koeffizient	p
– Wissen	0,13	0,18	0,72	0,00
– Fähigkeiten	0,31	0,00	0,78	0,00
– Motivation zu selbständigem Lernen	-0,04	0,68	0,77	0,00
– Erfahrungen mit kooperativer Arbeit	0,01	0,89	0,29	0,00
– Transfererwartung	0,41	0,00	0,80	0,00

Wird der Einfluss der personellen Variable „Allgemeine Einstellung zum problemorientierten Lernen“ mittels Kovarianzanalyse bei der Unterschiedsprüfung zwischen den beiden Lehrformen kontrolliert, so bleiben die Unterschiede signifikant. Sie sind bezüglich der Ziele, die mit der Einführung des problemorientierten Lernens angestrebt werden, hypothesenkonträr. Die traditionelle Lehre führt zu etwas günstigeren Werten in den Zielerreichungskriterien.

## 5. Diskussion

Für die Evaluation problemorientierter Lehre ist es nicht immer realistisch, kontrollstarke Randomisierungspläne einzusetzen. In der Phase der schrittweisen Einführung problemorientierter Lehrmodule können Messwiederholungspläne für eine vergleichende Evaluation problemorientierter und traditioneller Lehre genutzt werden. Mit ihnen kann ausgewiesen werden, in welchem Maße die angestrebten Ziele erreicht sind, wie groß der Verbesserungsspielraum ist und ob sich der Grad der Zielerrei-

chung der problemorientierten Lehre gegenüber der traditionellen Lehre unterscheidet. Dieser Variantenvergleich ist nicht nur eine Strategie für den Wirksamkeitsnachweis eines neuen Programms. Sie liefert – zeitig durchgeführt – für Programmgestalter wichtige Informationen für die Optimierung des neuen Programms. Für das Ziel der formativen Evaluation und der Optimierung der problemorientierten Lehre ist die hier beschriebene methodische Strategie selbstverständlich nicht die einzige. Sie ist zu ergänzen um Erhebungen zu Stärken und Schwächen der realisierten Programmversion und um das Einholen von Verbesserungsvorschlägen sowohl von den Programmteilnehmern als auch von den Personen, die das Programm durchführen, insbesondere von den Tutoren. Auf diese Weise können quantitativ beschriebene Schwachstellen inhaltlich verständlich gemacht werden und Verbesserungsmöglichkeiten werden dann leichter akzeptiert. Das Verknüpfen solcher „begleitenden“ Erhebungen von Schwachstellen und Gestaltungsvorschlägen mit einem methodisch kontrollierten quantitativen Vergleich der problemorientierten Lehre mit dem Standardprogramm „Traditionelle Lehre“ hat Vorteile. Die Bewertung des innovativen Programms wird erleichtert, weil diese Bewertungen durch das Vergleichsprogramm quasi verständlich verankert werden. Den Programm-Akteuren, d.h. den Verantwortlichen für die Entwicklung und Ausarbeitung des neuen Programms und den Verantwortlichen für die Implementierung und Durchführung des Programms wird die Interpretation von Schwachstellen erleichtert, weil sie durch den Vergleich konkretisiert werden und damit werden auch die Akzeptanz von Optimierungsvorschlägen und die aktive Mitwirkung an Verbesserungen erleichtert.

Die vorgestellten Befunde zeigen: Es ist nicht leicht, eine neue Lehrform schon in der Planungsphase hinsichtlich vieler Aspekte so zu konzipieren, dass die gewünschten Ziele gut erreicht werden. Die zwei Beispiele in diesem Beitrag zeigen heterogene Befunde. In der Studie bei den Zahnmedizinern konnten die mit der problemorientierten Lehrform angestrebten Ziele tendenziell erreicht werden. In mehreren Kriterien wurden die Lernresultate bei problemorientierter Lehre besser beurteilt als bei traditioneller Lehre. Im zweiten Beispiel ist der Befund umgekehrt. Das belegt auch die Schwierigkeit, die es bereitet, eine eingeführte und damit auch gut erprobte Lehrform zu verbessern. Eine Phase der Optimierung der problemorientierten Lehre ist deshalb einer Beurteilung ihrer Wirksamkeit voranzustellen. Befunde aus unserer Gruppe zeigen, dass mit Hilfe sehr konkret durchgeführter formativer Evaluation Verbesserungen der POL-Kurse in Bezug auf die erreichten Ziele gelingen (Nitsche 2003). Die Gestaltungsaufgabe ist so komplex und erfordert ein koordiniertes Zusammenwirken vieler Akteure, so dass es realistischer ist, sie als Prozess einer schrittweisen Optimierung zu begreifen. Dieser Prozess ist auf Rückkopplungen angewiesen, die durch Evaluationen gegeben werden.

Die Befunde des Beitrags machen auch deutlich, dass eine Mehr-Ebenen-Evaluation wichtig ist. Die zweite Studie mit dem Beispiel der Mediziner Ausbildung markiert nicht nur für den neuen POL-Kurs eine unbefriedigende Zielerreichung in den Kriterien des Lernerfolgs. Sie gibt auch an, dass die Akzeptanz der organisatorischen Rahmenbedingungen für den POL-Kurs Onkologie schlechter ist als für Abschnitte traditioneller Lehre im gleichen Fachsemester. Zufriedenheit und Akzeptanz eines neuen Programms sind, wie Kirkpatrick (1998) nachdrücklich argumentiert, eine Voraussetzung für Lernerfolg. Wenn sie nicht gegeben sind, kommt es zu Reibungsverlusten. Im Beispiel konnten solche als Verärgerung über die zeit-

liche Implementierung des POL-Kurses konkretisiert werden, die es den Studierenden nur ungenügend ermöglichte, sich auf die Tutorien vorzubereiten, und auf nicht ausreichendes Material für das im problemorientierten Lernen geforderte Selbststudium (zu wenig Computerarbeitsplätze, zu restriktive Bibliotheksöffnungszeiten). Die Nachbesserung an solchen Punkten ist erforderlich, um eine theoriegerechte Implementierung des problemorientierten Lernens, d.h. in der Sprache der Evaluation eine hinreichende Treatmentreinheit zu erreichen. Bei Mängeln in der Implementierung müssen die Bedingungen für das Programm zunächst so organisiert werden, dass eine ordnungsgemäße Durchführung des Programms erreicht wird.

Das zweite Beispiel zeigt auch, es ist nicht nur wichtig, die Akzeptanz mit den Implementierungsbedingungen für das neue Programm zu erheben sondern auch die Personenvariable „Allgemeine Einstellung zum problemorientierten Lernen“. Bei den Medizinstudenten im Beispiel wurde sie zweimal erhoben und jedes Mal mit einer sehr großen Streuung. Das zeigt, es gibt bei den Studierenden eine ambivalente Haltung zum problemorientierten Lernen. Ein Teil von ihnen beurteilt diese Lehrform positiv und ist froh, sie zu erleben. Ein anderer Teil hat aber nicht nur eine sehr kritische Haltung sondern zeigt auch Reaktanz. Die Verzerrung der Lernerfolgseinschätzungen durch diese Personvariable wurde durch die Berechnung der Korrelationskoeffizienten mit den Lernerfolgskriterien belegt. Das hat Konsequenzen für die Interpretation von Evaluationsbefunden und das Ableiten von Schlussfolgerungen aus ihnen. Das Ziel einer Evaluation besteht aber nicht darin, Verzerrungen von Messresultaten zu begründen sondern in der Optimierung neuer Programme und im Ausweisen ihrer Potenzen und Grenzen. Die große Varianz in der Akzeptanz des problemorientierten Lernens macht darauf aufmerksam, dass die Ziele und Mechanismen des problemorientierten Lernens nicht allen Studierenden verständlich sind. Das weist darauf hin, dass für eine neue Lehrform auch Aufklärungsarbeit zu leisten ist.

## 6. Literatur

- Bergmann, Bärbel/Dieter, Peter/Nitsche, Ines (2002): Evaluation des Dresdner Modells DIPOL (Dresdner integratives problemorientiertes Lernen) des reformierten Medizinstudiums. Medizinische Ausbildung. Supplement der Zeitschrift „Das Gesundheitswesen“, 19, S. 44-49.
- Dieter, Peter (2001): Reformcurriculum der Medizin nach Harvard – Modell an der Medizinischen Fakultät Carl Gustav Carus der TU Dresden. Wissenschaftliche Zeitschrift TU Dresden, 50 (4), S. 37-42.
- Eisfeldt, Doreen (2002): Evaluation des problemorientierten Lernens im Studiengang Zahnmedizin der Medizinischen Fakultät der TU Dresden. Diplomarbeit TU Dresden (unveröff.).
- Eitel, Florian H. (2002): Die neue Approbationsordnung verlangt tiefgreifende Änderungen in der Lehrorganisation. Medizinische Ausbildung, 19, S. 1-2.
- Giarini, Orio/Liedtke, Patrick M. (1998): Wie wir arbeiten werden. Der neue Bericht an den Club of Rome. Hamburg: Hoffmann und Campe Verlag.
- Gräsel, Cornelia (1997): Problemorientiertes Lernen. Göttingen: Hogrefe.
- Kirkpatrick, Donald L. (1998): Evaluation training programs. The four levels. 2. Ausgabe. San Francisco: Berrell-Koehler Publishers.
- Kromrey, Helmut (1994): Wie erkennt man „gute Lehre“? Was studentische Vorlesungsbefragungen (nicht) aussagen. Empirische Pädagogik. Zeitschrift zu Theorie und Praxis erziehungswissenschaftlicher Forschung, 8 (2), S. 153-168.

- Kromrey, Helmut (2001): Evaluation von Lehre und Studium – Anforderung an Methodik und Design. In: Spiel, C. (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck. Münster: Waxmann, S. 21-60.
- Lanadio, Sabine (2002): Summative Evaluation problemorientierter Lehrstrategien an der Medizinischen Fakultät Carl Gustav Carus der TU Dresden am Beispiel des Vergleichs der problemorientierten Lehre mit der traditionellen Lehre. Diplomarbeit TU Dresden (unveröff.).
- Metzler, Peter/Krause, Bodo (1997): Methodischer Standard bei Studien zur Therapieevaluation. *Methods of Psychological Research Online*, 2 (1). URL: <http://www.pabst-publishers.de/mpri>.
- Nitsche, Ines (2003): Evaluationskonzept zur Optimierung komplexer, problemorientierter Curricula am Beispiel des Medizinstudiums. Dissertationsschrift TU Dresden.
- Reinmann-Rothmeier, Gabriele/Mandl, Heinz (1997): Lehren im Erwachsenenalter. Auffassungen vom Lehren und Lernen, Prinzipien und Methoden. *Enzyklopädie der Psychologie, Psychologie der Erwachsenenbildung*. In: *Pädagogische Psychologie*, 4, S. 355-403.
- Reuschler, Hans (2002): Erfahrungsbasiertes Medizinstudium. Medizinische Ausbildung. Supplement der Zeitschrift „Das Gesundheitswesen“, 19, S. 13-19.
- Spiel, Christiane/Gössler, Martin (2001): Zwischen Selbstzweck und Qualitätsmanagement – Quo vadis, evaluatione? In Spiel, C. (Hg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck. Münster: Waxmann, S. 9-20.
- Willke, Hans (1998): Organisierte Wissensarbeit. *Zeitschrift für Soziologie*, 27 (3), S. 161-177.