



Validierung und Weiterentwicklung der Pädagogisch qualifizierten Fahrschulüberwachung (PQFÜ)

Susann Mörl, Daniel Kasper & Dietmar Sturzbecher



IMPRESSUM

Titel: **Validierung und Weiterentwicklung der Pädagogisch qualifizierten Fahrschulüberwachung**

Herausgeber: IPV GmbH

Autoren: Mörl, S., Kasper, D. & Sturzbecher, D.

Anschriften: IPV GmbH
Zum Alten Amtshaus 3
16727 Oberkrämer
Tel.: 0151-12222280
E-Mail: info@ipv-ok.com

Das diesem Bericht zugrunde liegende Forschungsprojekt wurde mit Mitteln des Ministeriums für Infrastruktur und Raumordnung des Landes Brandenburg (MIR) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Alle Rechte vorbehalten. Jede Verwertung ohne Zustimmung des Herausgebers verstößt gegen das Urheberrecht und wird gerichtlich verfolgt. Das gilt insbesondere für Vervielfältigungen jeder Art, Übersetzungen, Mikroverfilmung und die Einspeicherung in elektronische Systeme einschließlich Weiterverarbeitung.

Zitiervorschlag: Mörl, S., Kasper, D. & Sturzbecher, D. (2008). Validierung und Weiterentwicklung der Pädagogisch qualifizierten Fahrschulüberwachung. Oberkrämer: IPV.

1. Auflage, 2008

© IPV GmbH Oberkrämer

Inhalt

1	Problem- und Zielstellung.....	5
2	Ergebnisse der Begleituntersuchung und Schlussfolgerungen	7
3	Ergebnisse der Validierungsstudie und Schlussfolgerungen	9
4	Fazit und Empfehlungen für die Weiterentwicklung der PQFÜ.....	19
	Literatur	21

1 Problem- und Zielstellung

Das Risiko, einen Verkehrsunfall zu verursachen, fällt bei Fahranfängern um ein Vielfaches höher aus als bei erfahrenen Fahrern. Allerdings halbiert sich dieses Unfallrisiko in den ersten sechs Monaten des Führerscheinbesitzes. Weiterhin werden rund 95 Prozent aller Verkehrsunfälle teilweise oder ganz durch menschliches Fehlverhalten verursacht. Dies legt die Vermutung nahe, dass durch eine optimale Fahranfängervorbereitung und insbesondere durch eine pädagogisch anspruchsvolle Fahrschulbildung die Unfallzahlen gesenkt werden könnten. Fahrschulen haben demnach eine besondere Verantwortung, junge Menschen auf eine sichere Teilnahme am Straßenverkehr vorzubereiten.

Im System des Fahrlehrerrechts wurden zwei unterschiedliche Möglichkeiten eröffnet, um eine hohe Qualität der Fahrschulbildung in Fahrschulen zu gewährleisten. Bei diesen beiden Alternativen handelt es sich zum einen um die Überwachung der Fahrlehrer und Fahrschulen gemäß § 33 („Überwachung“) des Fahrlehrergesetzes (FahrLG) und zum anderen um den Einsatz eines Qualitätssicherungssystems nach § 34 („Ausnahmen“) FahrLG. Die Überwachung der Fahrlehrer und Fahrschulen gemäß § 33 FahrLG eröffnet dabei auch die Möglichkeit, die staatliche Kontrolle nicht nur auf eine Formalüberwachung (z. B. der Ausstattungsstandards und Aufzeichnungspflichten) zu fokussieren, sondern darüber hinaus zusätzlich eine Überprüfung der fachlichen und pädagogisch-didaktischen Qualität der Fahrschulbildung vorzunehmen.

Im Land Brandenburg wird mit dem System der „Pädagogisch qualifizierten Fahrschulüberwachung“ („PQFÜ“; Sturzbecher, Hrsg., 2004) die Fahrschulüberwachung unter pädagogischen und methodischen Gesichtspunkten ausgestaltet. Bei der PQFÜ werden Elemente der Formalüberwachung mit einer umfassenden Qualitätskontrolle der theoretischen und praktischen Ausbildung in Fahrschulen verbunden. Das PQFÜ-System wurde mittels Runderlass des Ministeriums für Infrastruktur und Raumordnung des Landes Brandenburg (MIR) mit Wirkung vom 15.03.2005 in Kraft gesetzt. Das primäre Ziel der PQFÜ besteht darin, eine ordnungsgemäße Fahrschulbildung mit einer pädagogischen Mindestqualität des theoretischen Unterrichts und der praktischen Ausbildung sicherzustellen sowie formale Rechtsverstöße aufzudecken. Bei der praktischen Umsetzung der PQFÜ übernimmt die IVS Institut für Verkehrssicherheit gGmbH organisatorische, verwaltungsmäßige und abrechnungstechnische Aufgaben, insbesondere die Organisation der Beauftragung geeigneter Sachverständiger als Prüfer, die Verwaltung der Akten sowie die Kontrolle der ordnungsgemäßen Erstellung des Überwachungsberichts und der Rechnung der Sachverständigen.

Für die Umsetzung des PQFÜ-Systems wurde vom Institut für angewandte Familien-, Kindheits- und Jugendforschung an der Universität Potsdam (IFK) ein Manual erarbeitet, das erstens Checklisten für verschiedene Überwachungsinhalte aus dem Regelablauf der Formalüberwachung und ein Schema für typisierte Beanstandungen enthält. Zweitens beinhaltet das Manual Beobachtungsinventare für den theoretischen Unterricht und die praktische Ausbildung in der Fahrschule, die von pädagogischen Sachverständigen als einheitliche „Messlatte“ für die Qualitätskontrolle der Ausbildung genutzt werden.

Die Erprobungsuntersuchung zum Beobachtungsinventar für den theoretischen Unterricht erfolgte 2003 im DEKRA-Niederlassungsgebiet Potsdam mit einer Stichprobe von 100 Fahrschulen und 123 Fahrlehrern. Die in diesem Beobachtungsinventar verwendeten 13 pädagogischen Qualitätskriterien wurden im Rahmen dieser Erprobungsuntersuchung auch anhand externer Kriterien wie der Bestehensquote der Fahrschulen validiert. Es konnte gezeigt werden, dass bei fachgerecht durchgeführten Fahrschulüberwachungen ein signifikanter Zusammenhang zwischen den bei der Überwachung des theoretischen Unterrichts erzielten Ergebnissen einer Fahrschule und den von ihr erreichten Bestehensquoten besteht (Sturzbecher, Schellhas & Hermann, 2004).

Im Rahmen der 2006 und 2007 im Auftrag des Ministeriums für Infrastruktur und Raumordnung des Landes Brandenburg durchgeführten Begleituntersuchung wurde die PQFÜ anhand der Auswertungsprotokolle und einer schriftlichen Befragung der Sachverständigen und Fahrlehrer evaluiert. Das Ziel dieser Begleituntersuchung bestand in der Überprüfung der Praktikabilität und Akzeptanz der PQFÜ sowie in der Aufdeckung von Optimierungspotentialen. Die Fortsetzung der Begleituntersuchung und die weitere Begleitung der Überwachung der praktischen Fahrausbildung 2008 dienten darüber hinaus auch der Prüfung der Umsetzungspraxis von Maßnahmeempfehlungen. Ausgehend von den Ergebnissen der Begleituntersuchung sollten letztlich Empfehlungen zur Optimierung der Umsetzung des PQFÜ-Systems abgeleitet werden.

Im Rahmen der Begleituntersuchung war eine Umsetzung der letztgenannten Ziele allerdings nicht vollständig möglich: Voraussetzung hierfür ist ein systematischer Vergleich der bei der Verfahrensentwicklung eingesetzten Analysestrategien und erzielten Ergebnisse einerseits mit den Strategien und Ergebnissen der Begleituntersuchung andererseits. Darüber hinaus setzt die Weiterentwicklung des PQFÜ-Systems voraus, dass die Beobachtungsinventare, mit denen die pädagogische Qualität des theoretischen Unterrichts und der praktischen Ausbildung erfasst werden, sowohl zuverlässig als auch valide sind. Um dies zu überprüfen, müssen beispielsweise die Überwachungsergebnisse der Fahrschulen mit den von ihnen im Überwachungszeitraum erzielten Bestehensquoten verglichen werden. Beide Voraussetzungen wurden mit der vorliegenden Validierungsstudie im Jahre 2008 geschaffen, um gezielte Empfehlungen zur Weiterentwicklung des PQFÜ-Systems geben zu können.

Ausgehend von der Erprobungsuntersuchung im Jahr 2003, in der erstmalig die Zuverlässigkeit und Validität der Beobachtungsinventare zur PQFÜ untersucht wurde, und von der Begleituntersuchung in den Jahren 2006 und 2007, wurden mit der vorliegenden Validierungsstudie 2008 mehrere Ziele verfolgt: Zum einen sollte ergänzend zu den Ergebnissen der Begleituntersuchung und als Fortsetzung der Analysen der Erprobungsuntersuchung die Zuverlässigkeit und Validität der Beobachtungsinventare zur PQFÜ untersucht werden. Zum anderen sollte die Effizienz der Verwaltung des PQFÜ-Systems überprüft werden (Ist-Stand der Wirksamkeit der PQFÜ). Auf der Grundlage dieser Befunde sollte dann ein Konzept entwickelt werden, in dem Möglichkeiten zur kontinuierlichen Evaluation sowie für die Pflege und stetige Weiterentwicklung des PQFÜ-Systems (Soll-Stand der Wirksamkeit von PQFÜ) abgeleitet werden. Schließlich sollte im Rahmen dieses Konzepts ein Vorschlag zum Aussehen und zum Betrieb einer Datenbank für die kontinuierliche Evaluation der PQFÜ und eine damit verbundene Berichterstattung zu ihrer Wirksamkeit unter Berücksichtigung des Datenschutzes erarbeitet werden.

Um diese Ziele zu erreichen, wurden bei der vorliegenden Validierungsstudie 2008 zunächst die bei der Erprobungsuntersuchung zur PQFÜ 2003 eingesetzten Analysestrategien mit den Daten aus der Begleituntersuchung 2006/2007 und den Bestehensquoten der überwachten Fahrschulen in 2006/2007 repliziert. Anhand eines systematischen Vergleichs der Befundstrukturen der Validierungsstudie und der Erprobungsuntersuchung wurden dann Rückschlüsse auf die methodische Qualität der bisher durchgeführten Überwachungen gezogen. Danach wurden weiterführende Schlussfolgerungen zu den typischen Defiziten bei der Ausbildungsqualität der überwachten Fahrschulen und zur Angemessenheit der ggf. eingeleiteten Sanktionsmaßnahmen abgeleitet. Um Aussagen zur Effizienz der Verwaltung der PQFÜ zu treffen, wurden offene Interviews mit Vertretern beteiligter Institutionen geführt; diesbezügliche Ergebnisse wurden bei der Erarbeitung von Optimierungsvorschlägen berücksichtigt. Abschließend wurden Empfehlungen für die Pflege und Weiterentwicklung des PQFÜ-Systems abgeleitet.

2 Ergebnisse der Begleituntersuchung und Schlussfolgerungen

Die vorliegende Ergebniszusammenfassung bezieht sich auf die in den Jahren 2006 und 2007 von der Arbeitsstelle für Bildungs- und Sozialisationsforschung (ABS) der Universität Potsdam durchgeführte Begleituntersuchung zur PQFÜ. Von den 295 statistisch erfassten Überwachungen (270 zum Theorieunterricht; 25 zur praktischen Fahrausbildung) im Evaluationszeitraum vom 15.03.2006 bis zum 31.12.2007 sind 280 Auswertungsbogen in die Evaluation eingegangen (255 Theorie¹; 25 Praxis). Zu den 255 in die Begleituntersuchung einbezogenen Überwachungsfällen des theoretischen Unterrichts lagen 223 Fragebogen mit Einschätzungen der Sachverständigen und 124 Fragebogen mit Einschätzungen der überwachten Fahrlehrer vor. Der Ergebnisüberblick beschränkt sich aufgrund der eingeschränkten Datenbasis zur Überwachung der praktischen Ausbildung weitgehend auf die Ergebnisse der Überwachung der theoretischen Ausbildung.

Im Rahmen der Überwachung des theoretischen Unterrichts werden 13 Qualitätsdimensionen von den Sachverständigen mittels einer vierstufigen Skala eingeschätzt, wobei die Qualitätsskala in einen zweistufigen positiven Bereich („Erfüllungsbereich“) und in einen zweistufigen negativen Bereich („Nichterfüllungsbereich“) unterteilt ist. Hinsichtlich der Durchführungspraxis und der Praktikabilität des Beobachtungsinventars für die Überwachung des theoretischen Unterrichts ziehen sowohl die Sachverständigen als auch die überwachten Fahrlehrer ein positives Fazit: Insgesamt gaben mehr als die Hälfte der überwachten Fahrlehrer an, dass der Ablauf der Überwachung als sehr praktikabel einzuschätzen ist. Die Mehrheit der überwachten Fahrlehrer sieht das PQFÜ-System zudem als ein geeignetes Verfahren an, Qualitätsunterschiede und Möglichkeiten der Qualitätsoptimierung aufzuzeigen.

Betrachtet man die Durchführungspraxis der Überwachungen im Hinblick auf die Maßnahmeempfehlungen, lässt sich zunächst feststellen, dass die Sachverständigen eher zurückhaltend agiert haben und Maßnahmeempfehlungen fast ausschließlich in Fällen erfolgten, in denen substantielle Qualitätsmängel deutlich erkennbar waren. Im Hinblick auf die zur Vereinheitlichung der Empfehlungspraxis erarbeiteten Orientierungshilfen für die Sachverständigen zeigte sich, dass sie zu einer einheitlichen Verfahrensweise bei mehrfacher Unterschreitung der Mindeststandards beitrugen: Während im ersten Jahr der PQFÜ-Durchführung in solchen Fällen vereinzelt keine Maßnahmeempfehlungen ausgesprochen wurden, erfolgte 2007 immer eine Maßnahmeempfehlung. In diesem Zusammenhang fällt jedoch auf, dass Maßnahmeempfehlungen in den meisten Fällen nur aufgrund des ersten Maßnahmenkriteriums vergeben wurden, also wenn mindestens sieben Qualitätsdimensionen im Nichterfüllungsbereich lagen. Das zweite Orientierungskriterium, die Unterschreitung von mindestens zwei Qualitätsdimensionen in einer Kriteriengruppe, zog hingegen vielfach keine Maßnahmeempfehlung nach sich. Besonders häufig war dies bei der Kriteriengruppe „Interaktion“ festzustellen, welche die vier Kriterien „Angemessenes Reagieren auf Beiträge der Fahrschüler“, „Organisation von Diskussionen“, „Organisation von Erfahrungsberichten“ und „Diskursive Unterrichtsgestaltung durch Medien“ umfasst.

Die Ursache für die festgestellte unbefriedigende Empfehlungspraxis könnte möglicherweise darin liegen, dass Akzeptanzprobleme bei der Anwendung des zweiten Maßnahme-

¹ Die im nachfolgenden Kapitel 3 zu den Ergebnissen der Validierungsstudie etwas variierenden und im Vergleich zur hier dargestellten Begleituntersuchung leicht abweichenden Fallzahlen entsprechen der tatsächlichen Auswertung der jeweils verfügbaren Datensätze und sind daher nicht als Darstellungsfehler anzusehen.

kriteriums bestehen, insbesondere bei der Kriteriengruppe „Interaktion“, die im Unterschied zu den anderen Kriteriengruppen relativ schwer zu erfüllende Einzelkriterien umfasst. Allerdings spielen diese Kriterien für eine pädagogisch anspruchsvolle Fahrschulbildung jugendlicher Fahrerlaubnisbewerber und für den Ausbildungserfolg eine besonders wichtige Rolle (Sturzbecher, Schellhas & Hermann, 2004). Ein anderer Grund für die beschriebene Empfehlungspraxis könnte auch darin bestehen, dass das zweite Maßnahmekriterium erst 2007 eingeführt wurde und sich daher erst noch etablieren muss. Wie auch immer: Das PQFÜ-System kann nur optimal wirksam sein, wenn eine unzureichende Ausbildungsqualität in einem Qualitätsbereich auch entsprechende Unterstützungsmaßnahmen nach sich zieht und die ordnungsgemäße Anwendung und Umsetzung der Maßnahmeempfehlungen verbessert werden. Dies betrifft insbesondere die Anwendung des zweiten Entscheidungskriteriums für eine Maßnahmeempfehlung.

Abschließend bleibt zur Umsetzung der Maßnahmeempfehlungen anzumerken, dass bisher kaum Rückfragen von Seiten der Erlaubnisbehörden zu erfolgten bzw. nicht erfolgten Maßnahmeempfehlungen der Sachverständigen bei der IVS gGmbH eingegangen sind. Möglicherweise ließen sich durch eine genauere Kenntnis der Sachlage einige Einsprüche bzw. Beschwerden der Fahrschulinhaber im Hinblick auf die Durchführung oder die Ergebnisse der Überwachung vermeiden. Darüber hinaus sei grundsätzlich angemerkt, dass der IVS gGmbH Informationen zur Maßnahmenanordnung durch die Erlaubnisbehörden und zur Umsetzung der angeordneten Maßnahmen in den Fahrschulen fehlen. Damit erscheint eine Kontrolle der Wirksamkeit der PQFÜ für die Erhöhung der Ausbildungsqualität in den Fahrschulen nur schwer möglich.

3 Ergebnisse der Validierungsstudie und Schlussfolgerungen

Wie aufgrund der Empfehlungen zur Durchführung der PQFÜ (Sturzbecher, 2004) zu erwarten war, wurde in den ersten zwei Jahren der PQFÜ vornehmlich die theoretische Ausbildung der Fahrschulen überwacht. Der folgenden Tabelle 1 ist zu entnehmen, dass bei insgesamt 256 Fahrschulen der theoretische Unterricht durch die PQFÜ erfasst wurde. Demgegenüber wurde bei nur 25 Fahrschulen die praktische Ausbildung nach den Kriterien der PQFÜ bewertet. Aufgrund der geringen Fallzahl der bewerteten praktischen Ausbildungseinheiten beziehen sich die folgenden Analysen ausschließlich auf die PQFÜ-Ergebnisse zum Theorieunterricht.

Tab. 1: Beschreibung der Stichprobe

Evaluationsinstrumente	Fallzahlen zu den Evaluationsinstrumenten		
	15.03.2006 bis 31.12. 2006	01.01.2007 bis 31.12. 2007	Gesamt
Auswertungsbogen für den theoretischen Unterricht	104	152	256
Auswertungsbogen für die praktische Fahrschulbildung	13	12	25

Die deskriptiven Statistiken für die 13 Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht sind in der folgenden Tabelle 2 dargestellt. Wie an den Mittelwerten erkennbar ist, wurde der theoretische Unterricht der 256 Fahrschulen im Durchschnitt als „Gut“ (Skalenstufe „1“, Erfüllungsbereich) bzw. als „Befriedigend“ (Skalenstufe „2“, ebenfalls Erfüllungsbereich) bewertet. Dabei sind die Standardabweichungen bei allen 13 Qualitätskriterien relativ klein, das heißt die Fahrschulen unterscheiden sich hinsichtlich der pädagogisch-didaktischen Qualität des theoretischen Unterrichts insgesamt nur relativ wenig. Interessant erscheint, dass sich größere Unterschiede insbesondere bei den Kriterien zeigen, die auf interaktive Lehrformen zielen (z. B. Berücksichtigung von Erfahrungsberichten der Fahrschüler und Durchführung von Diskussionen) oder die sich auf die Beachtung der individuellen Lernvoraussetzungen der Fahrschüler (Binnendifferenzierung) und die Lernkontrolle richten. Gerade diese Kriterien beeinflussen aber stark den Ausbildungserfolg der Fahrschüler und insbesondere auch die Bestehensquoten der Fahrschulen.

Tab. 2: Deskriptive Statistiken zu den einzelnen Qualitätskriterien

Qualitätskriterium	N	Minimum	Maximum	Mittelwert	Standardabweichung
Strukturierung der Unterrichtseinheit	255	1	4	2,09	,77
Motivierung der Fahrschüler und Praxisbezug	256	1	3	1,16	,41
Fachliche Vermittlung der Unterrichtsinhalte	256	1	3	1,25	,55
Binnendifferenzierung	255	1	4	1,57	,83
Angemessenes Reagieren auf Beiträge der Fahrschüler	255	1	4	1,27	,57
Tempo der Vermittlung der Unterrichtsinhalte	256	1	4	1,27	,53
Festigung	254	1	4	1,59	,75
Visualisierung der Unterrichtsinhalte durch Medien	255	1	3	1,13	,43
Diskursive Unterrichtsgestaltung durch Medien	256	1	4	1,59	,83
Qualität der Lehrvorträge	255	1	4	1,40	,58
Organisation von Erfahrungsberichten	255	1	4	1,78	,99
Organisation von Diskussionen	256	1	4	2,00	,97
Durchführung von Lernkontrollen	253	1	4	1,67	,95

Wie bereits oben erwähnt, setzen sinnvolle Empfehlungen zur Weiterentwicklung des PQFÜ-Systems voraus, dass die Beobachtungsinventare, mit denen die pädagogische Qualität des theoretischen Unterrichts erfasst werden soll, sowohl zuverlässig als auch valide sind. Ein Verfahren gilt dann als valide, wenn es diejenigen Merkmale, die es messen soll, auch tatsächlich erfasst (Lienert & Raatz, 1998). Es stellt sich also die Frage, ob die Beobachtungsinventare wirklich die Qualität des Fahrschulunterrichts messen. Diesen Grad der inhaltlichen Gültigkeit kann man auf unterschiedlichen Wegen bestimmen (Guthke, 1990); er ist auch von unterschiedlichen Einflüssen abhängig (z. B. von der Fahrschulstichprobe oder von den dabei eingesetzten Sachverständigen). In Abhängigkeit vom Vorgehen bei der Gewinnung von Gültigkeitsaussagen unterscheidet man unter anderem zwischen Inhaltsvalidität, interner Validität und Kriteriumsvalidität.

Inhaltsvalidität liegt vor, wenn Beurteilungskriterien augenscheinlich so beschaffen sind, dass sie die zu erfassenden Merkmale abbilden. Diese Inhaltsvalidität wird einem Verfahren in der Regel durch eine Expertenbeurteilung zugebilligt. Da die inhaltliche Validität der Beobachtungsinventare bereits bei der Erprobungsuntersuchung nachgewiesen wurde und da sich weder die Beobachtungsinventare noch die vorherrschenden Expertenmeinungen seit der Erprobungsuntersuchung wesentlich verändert haben, kann auf eine erneute Untersuchung dieses Aspekts bei der Validierungsstudie verzichtet werden. Dementsprechend kann für die Beobachtungsinventare inhaltliche Validität angenommen werden.

Interne Validität ist dann gegeben, wenn die Beurteilungskriterien der Beobachtungsinventare einem inhaltlich sinnvollen Zusammenhangsmuster folgen. Die Qualität eines Unterrichts ist nämlich nicht durch einfache Zusammenhänge zwischen zwei Unterrichtsmerkmalen sinnvoll charakterisierbar; vielmehr handelt es sich um ein komplexes Zusammenspiel vieler Einflussfaktoren, die sich auf die Qualität der Unterrichtsgestaltung und schließlich auf den Unterrichtserfolg auswirken. Jedes einzelne Kriterium der beschriebenen 13 Qualitätskriterien sollte deshalb in einem positiven Zusammenhang mit den anderen Qualitätskriterien stehen, das heißt die Qualitätskriterien sollten positiv miteinander korrelieren. Darüber hinaus sollten aber inhaltlich „verwandte“ Qualitätskriterien untereinander eine größere Zusammenhangsstärke aufweisen als zu anderen, inhaltlich nicht verwandten Kriterien.

Um die interne Validität zu überprüfen, wurde in der Validierungsstudie 2008 die Korrelationsmatrix der 13 Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht ermittelt (s. Tabelle 3 auf der folgenden Seite). In dieser Korrelationsmatrix sind die bivariaten Zusammenhänge zwischen den Qualitätskriterien dargestellt; die Zusammenhangsstärken können zwischen „+1“ und „-1“ variieren. Ein positiver Wert in diesem Variationsintervall besagt, dass eine hohe Qualität bei einem der jeweils zwei betrachteten Qualitätskriterien mit einer hohen Qualität beim anderen Qualitätskriterium einhergeht; ein negativer Wert sagt aus, dass eine hohe Qualität bei einem der beiden Kriterien mit einer niedrigen Qualität beim anderen Qualitätskriterium verbunden ist. Werte um Null bedeuten indes, dass zwischen den beiden betrachteten Qualitätskriterien kein Zusammenhang besteht.

Die in der Tabelle 3 dargestellten Ergebnisse belegen erwartungsgemäß, dass die 13 Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht mit zumindest mittleren Zusammenhangsstärken signifikant positiv korrelieren und enge inhaltliche Zusammenhänge bzw. „Verwandtschaften“ zwischen zwei Kriterien sich gegebenenfalls auch plausibel in engen statistischen Zusammenhängen wieder finden lassen: Beispielsweise gelingt es jenen Fahrlehrern, die besonders erfolgreich Diskussionen im Theorieunterricht anregen und leiten (Qualitätskriterium 12) auch besonders gut, den Unterricht durch Medien interaktiv zu gestalten (Qualitätskriterium 9; $r = 0,61$), während eine diskursive Unterrichtsgestaltung weniger wahrscheinlich mit einer guten Strukturierung der Unterrichtseinheit einhergeht (Qualitätskriterium 1; $r = 0,29$).

Tab. 3²: Korrelationsmatrix der 13 Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht (Produktmomentkorrelationen nach Pearson: „Pearson's r “)

Qualitätskriterium	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Strukturierung der Unterrichtseinheit	1,00	,29	,39	,30	,27	,27	,31	,30	,24	,44	,22	,29	,33
2. Motivierung der Fahrschüler und Praxisbezug	,29	1,00	,43	,35	,43	,30	,30	,33	,34	,39	,22	,23	,22
3. Fachliche Vermittlung der Unterrichtsinhalte	,39	,43	1,00	,43	,44	,32	,34	,52	,46	,42	,29	,40	,29
4. Binnendifferenzierung	,30	,35	,43	1,00	,39	,29	,29	,29	,55	,28	,48	,52	,34
5. Angemessenes Reagieren auf Beiträge der Fahrschüler	,27	,43	,44	,39	1,00	,36	,31	,42	,40	,31	,20	,35	,23
6. Tempo der Vermittlung der Unterrichtsinhalte	,27	,30	,32	,29	,36	1,00	,28	,20	,29	,28	,26	,29	,23
7. Festigung	,31	,30	,34	,29	,31	,28	1,00	,32	,35	,44	,24	,25	,42
8. Visualisierung der Unterrichtsinhalte durch Medien	,30	,33	,52	,29	,42	,20	,32	1,00	,37	,33	,19	,30	,24
9. Diskursive Unterrichtsgestaltung durch Medien	,24	,34	,46	,55	,40	,29	,35	,37	1,00	,28	,47	,61	,34
10. Qualität der Lehrvorträge	,44	,39	,42	,28	,31	,28	,44	,33	,28	1,00	,26	,32	,32
11. Organisation von Erfahrungsberichten	,22	,22	,29	,48	,20	,26	,24	,20	,47	,26	1,00	,46	,23
12. Organisation von Diskussionen	,29	,23	,40	,52	,35	,29	,25	,30	,61	,32	,46	1,00	,37
13. Durchführung von Lernkontrollen	,33	,22	,29	,34	,23	,23	,42	,24	,34	,32	,23	,37	1,00

Um die statistischen Zusammenhänge tiefer zu analysieren und hinsichtlich ihrer inhaltlichen Plausibilität zu überprüfen, wurden bei der Erprobungsuntersuchung 2003 explorative und konfirmatorische Faktorenanalysen durchgeführt. Diese ergaben, dass sich neun der beschriebenen 13 Qualitätskriterien („Strukturierung der Unterrichtseinheit“, „Motivierung der Fahrschüler und Praxisbezug“, „Fachliche Vermittlung der Unterrichtsinhalte“, „Binnendifferenzierung“, „Angemessenes Reagieren auf Beiträge der Fahrschüler“, „Tempo der Vermittlung der Unterrichtsinhalte“, „Festigung“, „Visualisierung der Unterrichtsinhalte durch Medien“ und „Diskursive Unterrichtsgestaltung durch Medien“) zu einem statistischen Faktor bündeln ließen und die inhaltstheoretisch sinnvolle Skala „Allgemeine Lehr-

² Alle Koeffizienten sind mindestens auf dem 5%-Niveau signifikant.

kompetenz“³ bildeten (Sturzbecher, Schellhas & Hermann, 2004). Diese Skala besitzt aufgrund prinzipieller methodischer Erwägungen (s. Fußnote 3) eine größere Gültigkeit für die Beurteilung der Ausbildungsqualität einer Fahrschule als die Einzeleinschätzungen zu den 13 Qualitätskriterien. Sofern das Beobachtungsverfahren zur Beurteilung der Ausbildungsqualität im theoretischen Unterricht intern valide ist, müssten sich bei der Validierungsstudie 2008 der Faktor bzw. die Skala „Allgemeine Lehrkompetenz“ wieder finden lassen.

Das Ergebnis der diesbezüglichen Faktorenanalyse zeigt die folgende Tabelle 4: Die genannten neun Qualitätskriterien bilden auch 2008 wieder einen einzigen Faktor, eine Skalenbildung erscheint wiederum sinnvoll, und obwohl die einzelnen Qualitätskriterien inhaltlich durchaus unterschiedliche Aspekte der pädagogisch-didaktischen Professionalität eines Fahrlehrers abbilden, erreicht diese Skala mit einem Wert von Cronbach's Alpha = .85 eine hervorragende interne Konsistenz.

Tab. 4: Ergebnisse der Faktorenanalyse

Qualitätskriterium	Faktor 1	Anfänglicher Eigenwert		
		Gesamt	Prozent der Varianz	Kumulierte %
Fachliche Vermittlung der Unterrichtsinhalte	,80			
Angemessenes Reagieren auf Beiträge der Fahrschüler	,75	4,46	49,50	49,50
Visualisierung der Unterrichtsinhalte durch Medien	,73			
Diskursive Unterrichtsgestaltung durch Medien	,73	Summen von quadrierten Faktorladungen für Extraktion		
Binnendifferenzierung	,66	Gesamt	Prozent der Varianz	Kumulierte %
Motivierung der Fahrschüler und Praxisbezug	,62	3,98	44,26	44,26
Festigung	,61			
Tempo der Vermittlung der Unterrichtsinhalte	,57			
Strukturierung Unterrichtseinheit	,46			

Fassen wir zusammen: Die verschiedenen Befunde zur internen Validität zeigen, dass sich sinnvolle theoretische Annahmen über inhaltliche Zusammenhänge zwischen den Qualitätskriterien empirisch bestätigt haben und damit neben den Expertenbeurteilungen (s. „Inhaltsvalidität“) ein weiteres wichtiges Indiz für die Gültigkeit des vorliegenden Beobachtungsinventars gefunden wurde. Darüber hinaus stimmen die validitätsbezogenen Ergeb-

³ Die Skala fasst die Beobachtereinschätzungen bezüglich der neun Qualitätskriterien zu einem „Zahlenwert“ zusammen. Dies erscheint vorteilhaft, weil die „Allgemeine Lehrkompetenz“ einen sehr komplexen Sachverhalt darstellt, der zwar für die Professionalität eines Fahrlehrers aussagekräftig ist, aber nur unzureichend mit einer Einschätzung zu einem Einzelkriterium erfasst werden kann. Bei einem Einzelkriterium würde sich zudem die Fehleranfälligkeit der statistischen Ergebnisse erhöhen: Hohe Lehrkompetenz muss ein Fahrlehrer nicht schon deshalb besitzen, weil er bei einem einzigen Beurteilungskriterium die „Höchstnote“ erhielt. Wurden jedoch mehrere Qualitätskriterien mit „Gut“ bewertet, so kann man mit höherer Sicherheit auf eine entsprechende Lehrkompetenz des Fahrlehrers geschlossen werden.

nisse der Validierungsstudie 2008 nahezu völlig mit den entsprechenden Ergebnissen der Erprobungsuntersuchung im Jahre 2003 (Sturzbecher, Schellhas & Hermann, 2004) überein, d. h. man gelangt zu sehr ähnlichen Aussagen bezüglich der bivariaten Beziehungen zwischen den einzelnen Qualitätskriterien untereinander und dem gesamten Netzwerk dieser Kriterien. Dies spricht ebenfalls für die methodische Stabilität und Validität des Verfahrens.

Neben der internen Validität wurde in der Validierungsstudie auch die Kriteriumsvalidität des Beobachtungsinventars für den theoretischen Unterricht überprüft. Die Kriteriumsvalidität bezieht sich auf den Grad des Zusammenhangs zwischen den bei der Unterrichtsbeobachtung erfassten Qualitätsmerkmalen und einem unabhängig davon erhobenen Außenkriterium. Es gibt zwei Arten von kriteriumsbezogener Validität: Bei der „Übereinstimmungsvalidität“ werden das Verfahrensergebnis und das Außenkriterium zur gleichen Zeit erhoben und auf Übereinstimmungen hin untersucht, während bei der „Vorhersagevalidität“ ermittelt wird, wie sich ein in der Zukunft liegendes Außenkriterium mit dem Verfahrensergebnis vorhersagen lässt („Prognostische Validität“). Zur Überprüfung der Kriteriumsvalidität des Beobachtungsinventars für den theoretischen Unterricht wurde der Zusammenhang zwischen den PQFÜ-Ergebnissen zur Unterrichtsqualität von 2007 und den Bestehensquoten der untersuchten Fahrschulen von 2007 untersucht, das heißt es wurde die Übereinstimmungsvalidität überprüft. Die folgende Tabelle 5 zeigt die gefundenen korrelativen Zusammenhänge zwischen den beobachteten Qualitätskriterien und den Bestehensquoten der Fahrschulen.

Tab. 5: Zusammenhang zwischen den PQFÜ-Ergebnissen von 2007 und den Bestehensquoten von 2007

Qualitätsmerkmale	Pearson's r
Skala „Allgemeine Lehrkompetenz“ des Fahrlehrers	,25*
Bereich „Kompetenz des Fahrlehrers zur Gestaltung von Unterrichtsformen“	
Qualität der Lehrvorträge	,22*
Organisation von Erfahrungsberichten	,28**
Organisation von Diskussionen	,23*
Durchführung von Lernkontrollen	,17

Anmerkung: „**“ bedeutet, dass der Koeffizient auf dem 1%-Niveau signifikant ist; „*“ bedeutet, dass der Koeffizient auf dem 5%-Niveau signifikant ist.

Aus der Tabelle 5 ist zu entnehmen, dass das aggregierte („Sammel“-) Qualitätskriterium „Allgemeine Lehrkompetenz“ wie auch die übrigen Qualitätskriterien (mit Ausnahme des Kriteriums „Durchführung von Lernkontrollen“) signifikant positiv mit den Bestehensquoten der Fahrschulen korrelieren. Der fehlende signifikante Einfluss des Qualitätskriteriums „Durchführung von Lernkontrollen“ auf die Bestehensquoten bedeutet nicht, dass dieses Kriterium für die Sicherung der Ausbildungsqualität nutzlos wäre; vielmehr resultiert der gefundene statistische Effekt, der unter pädagogischen Gesichtspunkten als unerwartet und wenig wünschenswert anzusehen ist, auch aus Verteilungsbesonderheiten der Beobachtungsergebnisse zu diesem Kriterium (z. B. relativ geringe Varianz). Insgesamt deuten die Befunde wie schon die entsprechenden Ergebnisse der Erprobungsuntersuchung 2003 aber auf eine akzeptable Kriteriumsvalidität der PQFÜ hin: Eine didaktisch anspruchsvolle Einbeziehung von Erfahrungsberichten der Fahrschüler und von Diskussionen in die Unterrichtsgestaltung sowie pädagogisch angemessene Lehrvorträge erwiesen sich also erwartungsgemäß neben der allgemeinen pädagogisch-didaktischen Kompetenz des Fahrlehrers als wichtige Einflussfaktoren auf den Erfolg der Fahrschüler bei der theoretischen Fahrerlaubnisprüfung.

Wie die aufgeführten Untersuchungen zur internen Validität zeigen, leisten die 13 Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht vor allem im Zu-

sammenspiel einen Beitrag zur Sicherung einer hohen Ausbildungsqualität. Um dieses Zusammenspiel bei der Gültigkeitsprüfung statistisch zu berücksichtigen, wurde mit Hilfe einer linearen Regressionsanalyse ein multipler Korrelationskoeffizient ermittelt. Die Stärke dieses gemeinsamen Zusammenhangs aller Qualitätskriterien zur Bestehensquote sollte dabei nicht wesentlich hinter die Zusammenhangsstärken der einzelnen Qualitätskriterien mit der Bestehensquote zurückfallen. Für die vorliegende Validierungsstudie wurde als multipler Korrelationskoeffizient zwischen allen beobachteten Qualitätskriterien und der Bestehensquote der Fahrschulen eine Zusammenhangsstärke von $r = 0,32$ festgestellt. Dies entspricht statistisch gesehen einer gemeinsamen Varianzaufklärung aller Beurteilungskriterien in Hinblick auf die Bestehensquote von ca. 10 Prozent: Der Erfolg bei der Theoretischen Fahrerlaubnisprüfung lässt sich also teilweise auf die von den Sachverständigen beobachtete Qualität des theoretischen Unterrichts zurückführen. Bedenkt man, dass noch andere starke Einflussfaktoren wie die Intelligenz, die Anstrengungsbereitschaft und die Prüfungsangst der Fahrerlaubnisbewerber die Prüfungsergebnisse wirkungsvoll beeinflussen (Sturzbecher, Großmann, Hermann, Schellhas, Viereck & Völkel, 2004), so erscheint die erzielte Varianzaufklärung durch die Beurteilungskriterien des Beobachtungsinventars als befriedigend. Der multiple Zusammenhang aller Qualitätskriterien zum externen Validitätskriterium fällt größer aus als alle Einzelzusammenhänge; das Verfahren erscheint also auch als kriteriumsvalide.

Es muss jedoch angemerkt werden, dass die gemeinsame Varianzaufklärung der Beurteilungskriterien bei der Erprobungsuntersuchung 2003 noch 24 Prozent betrug. Wie kann dieser Abfall der Varianzaufklärung erklärt werden? Die Bestehensquoten resultieren aus den Ergebnissen von Theoretischen Fahrerlaubnisprüfungen, die wiederum lehrzielorientierte Tests darstellen (Bönninger & Sturzbecher, 2005). Zu den methodischen Eigenarten lehrzielorientierter Tests wie der Fahrerlaubnisprüfung gehört es, dass sie oft nur eine geringe Varianz bei den Prüfungsergebnissen aufweisen, d. h. die meisten Prüfungsteilnehmer bestehen die Prüfung, und die erfolgreichen Bewerber unterscheiden sich in der Regel von den nicht erfolgreichen Bewerbern nur durch relativ geringe Leistungsunterschiede (Sturzbecher, 2008). Das letztgenannte Phänomen wird statistisch noch dadurch verschärft, dass nicht die Fehlerpunktzahlen der Bewerber, sondern nur ihre Endergebnisse im Sinne des Bestehens der Prüfung in die Bestehensquoten der Fahrschule eingehen: Die tatsächlichen Leistungsunterschiede zwischen den Bewerbern, wie sie sich in den Fehlerpunktzahlen widerspiegeln, werden durch diese Gruppierung verzerrt bzw. in reduzierter Form abgebildet. Die Bestehensquote einer Fahrschule spiegelt also das tatsächliche Kompetenzniveau der in ihr ausgebildeten Fahrschüler methodisch gesehen nur grob wider. Dies beeinträchtigt sicher die Feindifferenzierung der Fahrschulen im Mittelfeld der Ausbildungsleistungen, weniger dagegen die Erkennung von Fahrschulen mit wahrscheinlich schlechter Ausbildungsqualität.

Die dargestellten Zusammenhänge tragen im vorliegenden Fall sicher dazu bei, dass der statistische Zusammenhang zwischen den Bestehensquoten und den Überwachungsergebnissen der Fahrschulen methodisch gesehen etwas „störanfällig“ ist. Viel schwerer wiegt allerdings, dass bei der Validierungsstudie auch die Varianz bei den Überwachungsergebnissen der Fahrschulen nur gering ausfällt; d. h. sehr viele Fahrschulen bei den meisten Kriterien gute Ergebnisse aufweisen (s. o.). In der Folge lassen sich auch keine hohen bivariaten und multiplen Korrelationskoeffizienten erwarten (Amelang & Bartussek, 1997). Dies alles ist bei der Verwendung eines Messmodells, das auf kriteriums- bzw. lehrzielorientierten Messungen beruht, ein bekanntes methodisches Problem und schwer zu umgehen, wenn man nicht ganz auf die Darstellung von (statistischen) Zusammenhängen zu Außenkriterien verzichten will (Guthke, 1990). Die bei der Validierungsstudie 2008 gefundene geringe Varianzaufklärung der Bestehensquoten durch die Überwachungsergebnisse ist

also kein Indiz für eine fehlende Brauchbarkeit der Qualitätskriterien zur handwerklich soliden Einschätzung der Ausbildungsqualität an Fahrschulen, sondern lässt sich methodisch plausibel auf die Verteilungsbesonderheiten der Überwachungsergebnisse zurückführen. Zu fragen bleibt, ob die vorliegenden guten und relativ wenig differenzierten Überwachungsergebnisse im Betrachtungszeitraum durch das Überwachungsverhalten der Sachverständigen oder die Fahrschulstichprobe begünstigt wurden.

Würde man die Gründe für die vorrangig guten und wenig differenzierten Überwachungsergebnisse bei den Sachverständigen suchen, so ließe sich vermuten, dass sie die Ausbildungsqualität in den überwachten Fahrschulen systematisch überschätzen. Dies ist methodisch gesehen nicht völlig auszuschließen, da bei der Erprobungsuntersuchung 2003 die Varianz der Einschätzungen zu den Qualitätskriterien deutlich höher ausfiel, obwohl es sich bei der damaligen Fahrschulstichprobe um eine Positivauswahl handelte (Sturzbecher, Schellhas & Hermann, 2004). Bei den Beurteilern handelte es sich 2003 – wie bei den Sachverständigen heute – um sorgfältig ausgebildete und geprüfte Pädagogen, die allerdings nicht den Fahrlehrerberuf ausübten. Fokussiert man die Suche nach den Gründen für die vorrangig guten und wenig differenzierten Überwachungsergebnisse hingegen nicht auf die Sachverständigen, sondern auf die überwachten Fahrschulen, so lässt sich genauso gut vermuten, dass gegenwärtig vorrangig Fahrschulen mit relativ guter Ausbildungsqualität überwacht werden. Sofern ein solcher unbefriedigender Zustand tatsächlich vorliegt, könnte man ihm durch sinnvolle einheitliche Kriterien für die Anordnung der PQFÜ und anlassbezogene Überwachungen beispielsweise bei auffällig schlechten Bestehensquoten abhelfen; derartige anlassbezogene Überwachungen stellen bislang die Ausnahme dar.

Um zu überprüfen, ob eine anlassbezogene Anordnung der PQFÜ auch unter methodischen Gesichtspunkten sinnvoll erscheint, wurde in der Validierungsstudie untersucht, ob Fahrschulen mit einer Bestehensquote unter 50 Prozent ein Jahr später bei der PQFÜ weniger gute Ergebnisse erzielen als Fahrschulen mit einer Bestehensquote über 50 Prozent. Wäre dies der Fall, könnte man eine anhaltend schlechte Ausbildungsqualität in diesen Fahrschulen vermuten: einerseits wegen der schlechten Bestehensquoten, andererseits wegen der bei der Überwachung deutlich später – also nachlaufend – festgestellten schlechten Ausbildungsqualität. Um solche Fahrschulen mit vermutlich andauernder schlechter Ausbildungsqualität gegebenenfalls zu unterstützen, wäre eine frühzeitige PQFÜ wünschenswert, bei der die pädagogisch-didaktischen Probleme der Unterrichtsgestaltung präzise festgestellt und durch entsprechende Maßnahmenanordnungen bearbeitet werden könnten. Diese frühzeitige Unterstützung könnte dadurch ausgelöst werden, dass man ein Absinken der Bestehensquote einer Fahrschule beispielsweise unter 50 Prozent zum Anlass nimmt, die Anordnung einer PQFÜ zu erwägen. Bei diesen Erwägungen muss aber auch geprüft werden, ob eine unbefriedigende Bestehensquote einer Fahrschule nicht im Einzelfall durch andere Gründe als eine schlechte Ausbildungsqualität verursacht sein könnte.

Wie wurde die aufgeworfene Frage nun in der Validierungsstudie untersucht? Zunächst wurden zwei Gruppen von Fahrschulen gebildet: Die erste Gruppe umfasst alle Fahrschulen, bei denen 2006 nicht mehr als 50 Prozent der Fahrschüler die Theoretische Fahrerlaubnisprüfung bestanden hatten ($N = 7$); die zweite Gruppe beinhaltet alle Fahrschulen, die 2006 eine Bestehensquote über 50 Prozent aufwiesen ($N = 87$). Wie der Tabelle 6 entnommen werden kann, schnitten Fahrschulen mit schlechten Bestehensquoten in 2006 ein Jahr später bei der PQFÜ hinsichtlich der gefundenen Lehrkompetenz der Fahrlehrer wie auch hinsichtlich der Lehrqualität bei einigen didaktischen Ausbildungselementen (z. B.

Lehrvorträge, Lernkontrollen) deutlich schlechter ab als Fahrschulen, die 2006 eine Bestehensquote über 50 Prozent aufwiesen⁴.

Tab. 6: Mittelwerte bei Qualitätskriterien des Beobachtungsinventars für den theoretischen Unterricht, aufgeschlüsselt nach der Bestehensquote der Fahrschulen im Jahr 2006 (N = 94)

Qualitätsmerkmale	Bestehensquote 2006	Mittelwert	N
Skala „Allgemeine Lehrkompetenz“	Gruppe 1: Quote nicht größer als 50%	1,62	7
	Gruppe 2: Quote größer als 50%	1,34	87
Qualität der Lehrvorträge	Gruppe 1: Quote nicht größer als 50%	2,00	7
	Gruppe 2: Quote größer als 50%	1,36	87
Organisation von Erfahrungsberichten	Gruppe 1: Quote nicht größer als 50%	1,57	7
	Gruppe 2: Quote größer als 50%	1,45	87
Organisation von Diskussionen	Gruppe 1: Quote nicht größer als 50%	2,00	7
	Gruppe 2: Quote größer als 50%	1,86	87
Durchführung von Lernkontrollen	Gruppe 1: Quote nicht größer als 50%	1,86	7
	Gruppe 2: Quote größer als 50%	1,52	87

Die dargestellten Ergebnisse des Gruppenvergleichs stellen einerseits ein Indiz für die guten Differenzierungsmöglichkeiten dar, welche die PQFÜ für die Erkennung von Fahrschulen mit wahrscheinlich schlechter Ausbildungsqualität bietet; andererseits zeigen sie die prognostische Validität sehr geringer Bestehensquoten für eine schlechte Ausbildungsqualität: Geringe Bestehensquoten, die 50 Prozent nicht überschreiten, sind – wenn man nicht von größeren kurzfristigen Qualitätssprüngen bei der Ausbildungsqualität in einer Fahrschule ausgeht – mit deutlich erkennbaren Qualitätsdefiziten bei der PQFÜ verbunden. Deshalb erscheinen uns extrem schlechte Bestehensquoten im Regelfall als Anlasskriterium für die Anordnung einer PQFÜ als geeignet.

Insgesamt deuten die Ergebnisse der vorliegenden Validierungsstudie von 2008 also genauso wie die Ergebnisse der Erprobungsuntersuchung zum PQFÜ-System in 2003 darauf hin, dass die PQFÜ gut geeignet ist, um die pädagogisch-didaktische Qualität des theoretischen Fahrschulunterrichts festzustellen. Das Beobachtungsinventar für den theoretischen Unterricht erweist sich nicht nur als hoch zuverlässig (hohe interne Konsistenz), es weist auch eine zufriedenstellende Inhaltsvalidität, interne Validität und kriteriumsbezogene Validität auf. Schließlich erlauben es die Ergebnisse der Validierungsstudie, eine anlassbezogene Anordnung der PQFÜ aufgrund extrem schlechter Bestehensquoten empirisch zu rechtfertigen.

Neben der Überprüfung der Validität des Beobachtungsinventars für den theoretischen Unterricht wurde mit der Validierungsstudie 2008 auch noch das Ziel verfolgt, einen ersten Einblick in die Regelmäßigkeit, Einheitlichkeit und Effizienz der Umsetzung und Verwaltung der PQFÜ zu gewinnen. Zu diesem Zweck wurden Diskussionen und offene Interviews mit ausgewählten Vertretern der beteiligten Institutionen geführt (z. B. Sachverständige, IVS gGmbH, Fahrschulen). Dabei zeigten sich folgende Ergebnisse:

- (1) Der Überwachungsturnus wird in den Landkreisen und kreisfreien Städten nicht einheitlich gehandhabt. Insbesondere wird von der Option der Verlängerung des Überwachungsturnus bei untadeliger Ausbildungsqualität in unterschiedlichem Ausmaß Gebrauch gemacht.

⁴ Bei der Interpretation der dargestellten Mittelwerte ist zu beachten, dass höhere Werte beim Gruppenvergleich auf eine schlechtere Ausbildungsqualität hindeuten.

- (2) Das Ziel der Einführung des PQFÜ-Systems bestand darin, die unter verkehrspädagogischen Gesichtspunkten weitgehend nutzlose Formalüberwachung der Fahrschulstrukturen durch eine zweiteilige Kombination von Überwachungselementen zu ersetzen, die aus
- einer verkürzten stichprobenartigen Formalüberwachung („Kurzbericht“) mit einer adaptiven Prüfstrategie⁵ einerseits und
 - einer pädagogisch-didaktisch qualifizierten Einschätzung der Ausbildungsqualität des Theorieunterrichts und der praktischen Fahrausbildung andererseits besteht.

Mit der empfohlenen Verkürzung der Formalüberwachung auf Sachverhalte, die im jeweiligen „Kurzbericht zur Formalüberwachung“ (für die Überwachung des Theorieunterrichts s. Sturzbecher, Schellhas & Hermann, 2004) klar beschrieben werden, sowie mit der Fokussierung der Fahrschulüberwachung auf die pädagogische Prozessqualität der Ausbildung sollten nicht nur seit langem bekannte und anerkannte Grundsätze der Bildungsevaluation zum Tragen gebracht werden; zugleich wurde beabsichtigt, einen substantiellen Anstieg der finanziellen Belastungen der Fahrschulen durch die Einführung der PQFÜ zu vermeiden. Die festgestellte Anordnung von vollständigen Formalüberwachungen ohne speziellen Anlass neben der PQFÜ konterkariert diese Erwartungen und verkehrspädagogischen Ansprüche: Sie treibt die Überwachungskosten von Fahrschulen mit untadeliger Ausbildungsqualität in die Höhe und schmälert die Akzeptanz gegenüber der PQFÜ.

- (3) Wird eine vollständige Formalüberwachung neben der PQFÜ angeordnet, so findet sie in der Regel vor der Überwachung des Theorieunterrichts statt. Dafür gibt es zwei Gründe: Die zeitnahe Durchführung der Formalüberwachung im Zusammenhang mit der Unterrichtsüberwachung erscheint zum einen geboten, um Reisekosten zu vermeiden, wie sie ein eigenständiger Termin für eine Formalüberwachung mit sich bringen würde; zum anderen erscheint eine Durchführung der vollständigen Formalüberwachung nach dem Theorieunterricht, der meist abends stattfindet, schon aus Zeitgründen schwer möglich. Eine unmittelbar vor dem Theorieunterricht stattfindende vollständige Formalüberwachung erschwert dem Fahrlehrer aber eine angemessene Unterrichtsvorbereitung und belastet insbesondere im Falle festgestellter Unregelmäßigkeiten die Unterrichtsatmosphäre.
- (4) Die Sachverständigen wurden unterschiedlich oft eingesetzt, und einige von ihnen weisen nur eine relativ geringe Einsatzhäufigkeit auf. Die Gründe für die unterschiedlichen Einsatzhäufigkeiten liegen einerseits in der suboptimalen regionalen Verteilung der Geschäftsstellen bzw. Fahrschulen der Sachverständigen im Land Brandenburg und andererseits in den im Runderlass des Ministeriums für Infrastruktur und Raumordnung festgelegten Vorgaben zum örtlichen Einsatz der Sachverständigen. Diese Vorgaben erscheinen aus methodischer Sicht zwar als sinnvoll und nützlich, trotzdem sollten sie noch einmal hinsichtlich ihrer Zweckmäßigkeit geprüft werden, denn es gibt aus dem Fahrerlaubniswesen stichhaltige Befunde, dass mit der Einsatzhäufigkeit von Fahrerlaubnisprüfern bei der praktischen Fahrerlaubnisprüfung die Strenge der Bewertungen und die methodische Sicherheit des Prüferurteils zunahm (Hampel, 1976). Aufgrund der methodischen Ähnlichkeit der Beurteilungsprozesse, welche Fahrerlaubnisprüfer und Überwachungssachverständige ausführen, ist zu vermuten, dass sich mit dem häufigeren Einsatz der Sachverständigen auch das o. g. Problem der geringen Beurteilungsvarianz lösen würde, denn Beurteilungsunsicher-

⁵ Unter einer „adaptiven Prüfstrategie“ ist zu verstehen, dass der Sachverständige seine stichprobenartigen Kontrollen gemäß Kurzbericht ausdehnt oder die Durchführung einer vollständigen Formalüberwachung empfiehlt, sobald er bei seinen Stichproben auf Unregelmäßigkeiten stößt.

heit führt in der Regel aufgrund von Einspruchsbefürchtungen zu einer Mildetendenz bzw. einer Extremwertescheu bei den Beurteilern.

- (5) Bisher fokussiert die PQFÜ auf die Fahrschulinhaber; hinzukommt, dass in jeder Fahrschule bislang nur ein Fahrlehrer durch PQFÜ überwacht wurde. Eine nachhaltige Optimierung der Fahrausbildung in den Fahrschulen durch die PQFÜ setzt indes voraus, dass diejenigen Fahrlehrer einer Fahrschule überwacht werden, die den Theorieunterricht tatsächlich durchführen.
- (6) Will man sich einen Einblick in die Wirksamkeit und Nachhaltigkeit der PQFÜ im Land Brandenburg verschaffen, so erscheint es unverzichtbar, Daten zu den Überwachungsergebnissen, zu den gegebenenfalls angeordneten Maßnahmen (insbesondere zur Einhaltung der Anordnungskriterien und zur inhaltlichen Adäquatheit der Maßnahmen zu den gefundenen Beanstandungen), zur Erfüllung der angeordneten Maßnahmen und zu den Bestehensquoten fahrschulbezogen zu erfassen und zusammenzuführen sowie anhand einer periodischen Evaluation und Berichterstattung auszuwerten. Derzeit existieren weder eine solche einheitliche und datenschutzkompatible Datenbasis noch eine Berichterstattung, die systematische Validitätsanalysen oder Wirksamkeitsbetrachtungen zur PQFÜ zuließen. Damit wird auch eine Einschätzung erschwert, inwieweit die PQFÜ landesweit einheitlich in den Landkreisen und kreisfreien Städten praktiziert wird. Augenscheinlich ist dies nicht der Fall, was zu Wettbewerbsverzerrungen und zur Unzufriedenheit unter den Fahrschulen beiträgt.

4 Fazit und Empfehlungen für die Weiterentwicklung der PQFÜ

Die Ergebnisse der Validierungsstudie 2008 belegen, dass das System der Pädagogisch qualifizierten Fahrschulüberwachung methodisch gesehen funktioniert und den üblichen psychometrischen Anforderungen gerecht wird: Die PQFÜ stellt ein wirksames Verfahren zur Abbildung der Ausbildungsqualität einer Fahrschule dar. In weiteren Untersuchungen wird jedoch zu prüfen sein, ob die PQFÜ auch schon wesentlich zur Förderung der Ausbildungsqualität in den Fahrschulen beiträgt. Die nach wie vor im Ländervergleich relativ hohen Nichtbestehensquoten bei den Fahrerlaubnisprüfungen in Brandenburg und die diesbezüglichen großen Unterschiede zwischen den einzelnen Fahrschulen verweisen darauf, dass weitere Anstrengungen zur Optimierung der Qualität der Fahrschulbildung unternommen werden müssen. Zu diesen Anstrengungen sollte es gehören, unzweifelhaft vorhandene Reserven bei der fachgerechten und einheitlichen Umsetzung der PQFÜ durch die Sachverständigen und Behörden auszuschöpfen.

Nachfolgend werden auf der Grundlage der Begleituntersuchung, mit der die Einführung des PQFÜ-Systems in den Jahren 2006 und 2007 flankiert wurde, und der vorliegenden Validierungsstudie einige Empfehlungen für die künftige Optimierung und Weiterentwicklung des PQFÜ-Systems zusammenfassend dargestellt. Diese Empfehlungen richten sich einerseits auf die methodische und organisatorische Umsetzung des bestehenden PQFÜ-Systems und andererseits auf die Einführung einer periodischen Berichterstattung zur Wirksamkeit und Effizienz dieses Systems:

- (1) Es wird empfohlen, die Anordnung vollständiger Formalüberwachungen zu Gunsten der Überwachung der pädagogisch-didaktischen Ausbildungsqualität im Theorieunterricht und in der praktischen Fahrausbildung zu reduzieren und die Formalüberwachung auf die Checklisten der Kurzberichte zu beschränken, sofern es keinen besonderen Anlass (z. B. Beanstandungen, Beschwerden) für eine vollständige Formalüberwachung gibt. Dabei empfiehlt sich eine adaptive Überwachungs- bzw. Prüfstrategie, d. h. die Sachverständigen sollen gezielt und stichprobenartig nach den für die Anfertigung der Kurzberichte bereitgestellten Checklisten die Verhältnisse in den Fahrschulen kontrollieren und bei Auffälligkeiten genauer nachprüfen oder den Fahrerlaubnisbehörden eine vollständige Formalüberwachung empfehlen. Eine solche adaptive Prüfstrategie erfordert also ein flexibles Vorgehen der Sachverständigen; der Aufwand der Kontrollen muss an die jeweilige Situation angepasst werden.
- (2) Es wird empfohlen, im Rahmen von Beratungen mit den Erlaubnisbehörden eine einheitliche Umsetzung der PQFÜ im Land Brandenburg zu forcieren. Dabei ist insbesondere eine einheitliche Anwendung der Verlängerungsoption für den Überwachungsturnus bei Fahrschulen mit untadeliger Ausbildungsqualität und eine einheitliche Anordnungspraxis für Fördermaßnahmen bei Beanstandungen im Rahmen der Überwachung anzustreben; darüber hinaus ist die Kontrolle der Maßnahmeumsetzung zu vereinheitlichen.
- (3) Es wird empfohlen, die fachgerechte und einheitliche Umsetzung der PQFÜ auch durch einen verstärkten Erfahrungsaustausch der Sachverständigen und vertiefende Schulungen zur Anwendung der Orientierungshilfen für Maßnahmeempfehlungen (hinsichtlich des Theorieunterrichts und der praktischen Fahrausbildung) zu fördern. Für derartige Schulungen sind entsprechende Angebote zu entwickeln, die sich strukturell und inhaltlich auf die verschiedenen Gruppen von Qualitätskriterien und ihre didaktische Umsetzung richten müssen sowie diesbezügliche Umsetzungsprobleme (z. B. bei den Kriterien zur interaktiven Unterrichtsgestaltung) aufgreifen sollten. Darüber hinaus empfiehlt es sich, den Sachverständigen Supervisionen anzubieten.

- (4) Es wird empfohlen, neben begründet erscheinenden Beschwerden vor allem extrem geringe Bestehensquoten (z. B. Bestehensquoten unter 50%) als Auslöser für anlassbezogene PQFÜ heranzuziehen, um die gezielte Kontrolle und Unterstützung von Fahrschulen mit schlechter Ausbildungsqualität zu gewährleisten.
- (5) Es wird empfohlen, alle Fahrlehrer einer Fahrschule, die substantielle Ausbildungsanteile leisten, in die PQFÜ einzubeziehen.
- (6) Es wird empfohlen, die Einsatzregeln für die Sachverständigen („50-Kilometer-Regel“, kein zweimaliger Einsatz hintereinander in einer Fahrschule) hinsichtlich ihrer Praktikabilität sowie ihrer methodischen und organisatorischen Implikationen zu überdenken.
- (7) Es wird empfohlen, die Qualifikationsvoraussetzungen der Sachverständigen zu diskutieren. Dabei erscheint neben der Berufserfahrung als Fahrlehrer der Abschluss einer grundlegenden schulpädagogischen Ausbildung als wünschenswert.
- (8) Es wird empfohlen, die Anforderung einer ausführlichen Begründung der PQFÜ-Ergebnisse durch einen Sachverständigen als Bestellung eines Gutachtens zu werten, damit diese Leistung für den Besteller kostenpflichtig wird und den Sachverständigen vergütet werden kann. Sofern diese Empfehlung umgesetzt wird, ist für derartige Begründungen ein Anforderungsschema zu erarbeiten, das die zusätzlich bereitzustellenden Informationen und ihre formale Darstellung beschreibt.
- (9) Es wird empfohlen, die Überwachungsergebnisse sowie die gegebenenfalls von den Sachverständigen empfohlenen bzw. von den Fahrerlaubnisbehörden angeordneten Maßnahmen und ihre Umsetzung einheitlich zu erfassen sowie die Anordnungs- und Umsetzungspraxis fortlaufend hinsichtlich ihrer Regelmäßigkeit und inhaltlichen Sachgerechtigkeit zu evaluieren und über die Evaluationsergebnisse periodisch (z. B. jährlich) Bericht zu erstatten. Für diese periodische Berichterstattung zur Wirksamkeit und Effizienz des PQFÜ-Systems ist es notwendig:
 - einen einheitlichen Identifikationscode für die Fahrschulen einzuführen, der die Zusammenarbeit der einzelnen Institutionen (Fahrerlaubnisbehörden, Sachverständige, IVS gGmbH, Evaluationsinstitut) erleichtert und die Zusammenführung von Daten aus den verschiedenen Institutionen ermöglicht,
 - bei der IVS gGmbH einen einheitlichen Stammdatensatz für alle Fahrschulen anzulegen, der als Informationen enthalten muss
 1. den Identifikationscode der einzelnen Fahrschule,
 2. ihre Überwachungsergebnisse,
 3. ggf. die Maßnahmenempfehlungen der Sachverständigen und die Maßnahmenanordnungen der Fahrerlaubnisbehörden unter Angabe der Maßnahmemethode und der spezifischen inhaltlichen Ausrichtung der Maßnahme entsprechend der Qualitätskriterien,
 4. Angaben zur Maßnahmeumsetzung und
 5. Angaben zur Wirksamkeitskontrolle der Maßnahme.
- (10) Es wird empfohlen, alle zur Pflege, Evaluation und Weiterentwicklung des PQFÜ-Systems notwendigen Institutionen, Prozesse und Ergebnisse zusammenfassend in einem Handbuch zu beschreiben, das allen Beteiligten – oberste Landesbehörden, Fahrerlaubnisbehörden, Sachverständige, Fahrschulen, IVS gGmbH – als aktuelle Orientierungs- und Handlungsgrundlage dient.

Literatur

- Amelang, M. & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung*. Stuttgart: Kohlhammer.
- Bönninger, J. & Sturzbecher, D. (2005). *Qualität der Fahrerlaubnisprüfung. Ein Reformvorschlag für die theoretische Fahrerlaubnisprüfung*. Bergisch-Gladbach: Wirtschaftsverlag NW. Verlag für neue Wissenschaft GmbH.
- Guthke, J. (1990). Grundzüge der Testkonstruktion und Testauswertung. In J. Guthke, H.R. Böttcher & L. Sprung (Hrsg.). *Psychodiagnostik* (S. 201-324). Berlin: Deutscher Verlag der Wissenschaften.
- Hampel, B. (1976). Die Führerscheinprüfung als Maßnahme der Selektion. *Informationen & Mitteilungen der Sektion Verkehrspsychologie im BDP*, 8, 100-111.
- Lienert, G.A. & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- Sturzbecher, D. (Hrsg., 2004). *Manual für die pädagogisch qualifizierte Fahrschulüberwachung*. Potsdam: Universität.
- Sturzbecher, D. (2004). Grundlagen und Ausgestaltungsmöglichkeiten der „Pädagogisch qualifizierten Fahrschulüberwachung“. In D. Sturzbecher (Hrsg.), *Manual für die pädagogisch qualifizierte Fahrschulüberwachung*. Potsdam: Universität.
- Sturzbecher, D. & Labitzke, P. (2004). Überwachungsberichte zur Formalüberwachung. In D. Sturzbecher (Hrsg.), *Manual für die pädagogisch qualifizierte Fahrschulüberwachung*. Potsdam: Universität.
- Sturzbecher, D., Schellhas, B. & Hermann, U. (2004). Beobachtungsinventar und Checkliste zur Erfassung der theoretischen Fahrschulausbildung. In D. Sturzbecher (Hrsg.), *Manual für die pädagogisch qualifizierte Fahrschulüberwachung* (Modul 2). Potsdam: Universität.
- Sturzbecher, D., Großmann, H., Hermann, U., Schellhas, B., Viereck, K. & Völkel, P. (2004). Einflussfaktoren auf den Erfolg bei der theoretischen Fahrerlaubnisprüfung. In D. Sturzbecher (Hrsg.), *Einflussfaktoren auf den Erfolg bei der theoretischen Fahrerlaubnisprüfung*. Jugendliche und Risikoverhalten im Straßenverkehr. Hannover: Degener.
- Sturzbecher, D. (2008). Methodische Grundlagen der Praktischen Fahrerlaubnisprüfung. In D. Sturzbecher, J. Bönninger & M. Rüdell (Hrsg.), *Optimierung der Praktischen Fahrerlaubnisprüfung – Methodische Grundlagen und Möglichkeiten der Weiterentwicklung*. Dresden: TÜV | DEKRA arge tp 21.