

TEACH-Q: Ein valides und handhabbares Instrument zur Bewertung von Vorlesungen

Dienstleistung; Kausalanalyse; Lehrevaluation; Lehre, Qualität der; Validitätsmessung



U. Hansen



Th. Hennig-Thurau

Die im Rahmen einer umfangreichen empirischen Explorationsstudie erhobenen Erkenntnisse über Qualitätsdimensionen und -merk-

male bei der Evaluation von universitären Lehrleistungen bilden die Grundlage der Entwicklung eines Bewertungsinstrumentes für die Qualität der Lehre im Fachbereich Wirtschaftswissenschaften, das studenten- und dozentenseitige Perspektiven integriert. Dabei werden neben einem Vergleich multiattributiver Qualitätsmodelle verschiedene komplexe Reliabilitäts- und Validitätskriterien eingesetzt, um die Struktur des Konstruktes Qualität der Lehre sowie seine inhaltlichen Merkmale aus dienstleistungstheoretischer Perspektive zu ermitteln.

1. Ausgangspunkt und Vorgehensweise der Untersuchung

Die Evaluation von Lehrleistungen hat als Thema universitärer Diskussion neue Konjunktur und erfährt eine zunehmende *inner- und außeruniversitäre Aufmerksamkeit*. [1] Bereits in den 60er und 70er Jahren war diesbezüglich eine engagierte Bewegung entstanden, die aus universitären Reformabsichten heraus die *didaktischen Leistungen* der Professoren und des akademischen Mittelbaus verbessern und restrukturieren wollte. [2] Die gegenwärtig geführte Debatte weist über diesen Aspekt hinaus zusätzlich einen *bildungsökonomischen* Charakter auf: Angesichts leerer öffentlicher Kassen richtet sich das gesellschaftliche Interesse auf eine effiziente Verwendung finanzieller Res-

* Ursula Hansen ist Professorin am Lehrstuhl Marketing 1: Markt und Konsum der Universität Hannover, Königsworther Platz 1, 30167 Hannover; Dipl.-Kfm. Thorsten Hennig-Thurau ist Wissenschaftlicher Mitarbeiter am selben Lehrstuhl; Dr. Holger Wochnowski war bis September 1995 ebenfalls Mitarbeiter an diesem Lehrstuhl und ist seitdem bei der BMG Bertelsmann Music Group tätig. Die Autoren danken den studentischen Teilnehmern der Marktforschungspraktika 1994 und 1995, ohne deren aktive Unterstützung dieser Aufsatz nicht möglich gewesen wäre und insbesondere cand. rer. oec. Markus Langer, cand. rer. oec. Petra Kiepe und Dipl.-Ök. Harald Zielstorff für die Durchsicht von Teilen früherer Versionen dieses Aufsatzes. Ein besonderer Dank gebührt darüber hinaus zwei anonymen Gutachtern für ihre konstruktiven Anmerkungen.

TEACH-Q: Ein Instrument zur Bewertung von Vorlesungen

sources. Die Lehrevaluation wird in diesem Zusammenhang als Möglichkeit gesehen, die Qualität des Leistungsangebotes einzelner Einrichtungen der Universität (Fachbereiche, Institute, Lehrstühle) zu erfassen, um Maßstäbe für ein leistungsorientiertes Universitätsmanagement zu finden.

Vor dem Hintergrund derartiger Bestrebungen, die in anderen europäischen Ländern und insbesondere den USA bereits wesentlich weiter fortgeschritten sind, ist in den letzten Jahren an deutschen Universitäten eine Vielzahl von Verfahren und Instrumenten zur Bewertung der Lehrqualität entwickelt worden; eine Befragung an deutschsprachigen Universitäten durch die Autoren im Rahmen der Arbeitsgruppe »Evaluation der Lehre« im Frühjahr 1996 belegt den Einsatz von 168 verschiedenen Instrumenten (!) an wirtschaftswissenschaftlichen Lehrstühlen. Hinsichtlich dieses Methodenpluralismus im Bereich der Evaluationsforschung gilt es die wissenschaftliche Aufmerksamkeit insbesondere auf folgende Aspekte zu lenken:

a) Die vorhandenen Evaluationsansätze lassen sich grundsätzlich in zwei Kategorien einteilen: »kundenseitigen Verfahren«, d. h. solchen Ansätzen, die auf das Qualitätsurteil der Studenten als Leistungsempfänger abstellen, [3] stehen »anbieterseitige Verfahren« gegenüber. Letztere Gruppe umfaßt eine große Anzahl verschiedener Verfahren, die weit über das Instrument der Befragung hinausgehen (peer group reviews bzw. Expertengutachten, Dokumentenanalyse, Lehrberichte etc.). [4] Während Studentenbefragungen überwiegend durch einen induktiven Charakter gekennzeichnet sind, basieren anbieterseitige Evaluationsverfahren i. d. R. auf deduktiven Überlegungen, wobei die existierende Vielzahl an gesellschaftlichen Anspruchsgruppen der Hochschulausbildung einen breiten und durchaus nicht durchgängig harmonischen Zielkatalog zur Folge hat. Dieser enthält u. a. lerntheoretische, didaktische und funktionale Anforderungen, die bei einer Lehrevaluation entsprechend zu berücksichtigen sind, was durch den Einsatz einzelner Instrumente i. d. R. nicht zu leisten ist. Eine *Integration kunden- und anbieterseitiger* Evaluationsperspektiven findet gegenwärtig nur sehr beschränkt statt. Eine Ausnahme stellt hier das

»Heidelberger Inventar zur Lehrveranstaltungs-Evaluation« (HILVE) von Rindermann und Amelang dar, bei dem die Generierung von Indikatoren der Lehrqualität sowohl durch eine Befragung von Studenten als auch von Dozenten erfolgte. [5] Der Einsatz des Instrumentes schließlich bleibt jedoch auch bei HILVE auf eine Seite (hier: die der Studenten) begrenzt. [6]

b) Ein zweiter Punkt betrifft die empirische Reliabilität und Validität der vorhandenen Meßinstrumente. Während eine Vielzahl der existierenden Ansätze zur Lehrevaluation überhaupt keine bzw. nur eine fragmentarische theoretische Fundierung aufweist (so liegt etwa rund 87% der 168 insgesamt im Fachgebiet Wirtschaftswissenschaften zum Einsatz gelangenden Fragebögen *kein* theoretisches Konzept zugrunde), [7] können auch über die *empirische Validität* solcher Instrumente, die auf theoretischen Konzepten basieren, keine abschließenden Aussagen getroffen werden. Zugleich muß davon ausgegangen werden, daß sich die Eignung der verschiedenen existierenden Meßinstrumente für die Beurteilung der Qualität der Lehre signifikant voneinander unterscheidet. Das Resultat ist eine erhebliche *Intransparenz* bezüglich der Güte der verschiedenen Meßverfahren im Rahmen der Lehrevaluation. Hervorzuheben ist auch hier der HILVE-Fragebogen von Rindermann und Amelang, der zumindest eine Beurteilung der Reliabilität des Instrumentes auf der Grundlage von Verfahren der ersten Generation umfaßt. Weiterführende Reliabilitätsaussagen und kausalanalytische Validitätsprüfungen liegen jedoch auch bei HILVE nicht vor.

c) Aus der *Interdisziplinarität* des Forschungsgebietes Lehrevaluation schließlich resultiert das Problem, daß die Mehrzahl der vorhandenen Evaluationsinstrumente (so auch das HILVE-Inventar) für einen *fachbereichsübergreifenden* Einsatz konzipiert worden ist. Eine solche Vorgehensweise läßt jedoch die fachlichen Spezifika und die daraus resultierenden Besonderheiten in der Zielsetzung der Lehre einzelner Fachbereiche unberücksichtigt und erscheint uns deshalb als grundsätzlich problematisch. [8]

In unserer Untersuchung soll es nun darum

gehen, auf den hier in Kürze skizzierten Problemfeldern der Lehrevaluation aufbauend mit *TEACH-Q* ein Instrument vorzustellen, das die genannten Defizite der Lehrevaluation aufgreift und hinsichtlich dieser Bereiche einen kompetenten Lösungsansatz bietet. Dabei sollen der Untersuchung die folgenden grundlegenden Überlegungen vorangestellt werden:

Ein wichtiger Gegenstand der gegenwärtig intensiv geführten Evaluationsdiskussion sind die pädagogischen und didaktischen *Defizite des Leistungsangebotes* an wirtschaftswissenschaftlichen Fachbereichen bundesdeutscher Universitäten. Die zentrale Aufgabe der Evaluation der Lehre soll deshalb hier darin gesehen werden, bestehende Schwachpunkte aufzuzeigen und konkrete inhaltliche wie didaktische Anstöße zur Verbesserung der Qualität der Lehre zu liefern. Eine darüber hinaus gehende Verwendung, etwa für hochschulpolitische Zwecke, halten wir für grundsätzlich möglich, stellt jedoch *kein primäres Ziel* der Instrumentenentwicklung dar.

Universitätsangebote werden von uns grundsätzlich als *Dienstleistungen* verstanden.[9] Dies erscheint sinnvoll, da universitäre Lehrangebote sämtliche Merkmale enthalten, die den Dienstleistungen allgemein konstitutiv zugrundegelegt werden[10]:

- Universitäre Lehrangebote haben (mit Ausnahme von Vorlesungsskripten und ähnlichen Materialien) *immateriellen Charakter* und sind damit intangible, d.h. nicht greifbare Leistungen.
- Universitäre Lehrangebote stellen *Potentialangebote* dar, deren Qualität frühestens zum Zeitpunkt der Inanspruchnahme beurteilt werden kann (experience bzw. credence qualities).
- Universitäre Lehrangebote erfordern im Rahmen der Leistungserstellung bzw. -inanspruchnahme die Integration und teilweise die Mitwirkung der Studenten im Sinne eines *externen Faktors*, der nicht in der ausschließlichen Verfügungsgewalt des Dienstleisters steht.
- Produktion und Konsum von universitären Lehrangeboten haben *gleichzeitig* zu erfolgen.

Wenn den Lehrangeboten der Universitären ein Dienstleistungscharakter unterstellt werden kann, so scheint es naheliegend zu sein, dienstleistungstheoretische Ansätze aufzugreifen, um die Qualität solcher Veranstaltungen zu messen. Insbesondere bietet es sich an, auf die Erkenntnisse der Qualitätsforschung als wichtigem Teilbereich des Dienstleistungsmarketing[11] zu rekurrieren, da in diesem Bereich spezielle Methoden zur Messung der (Dienstleistungs-)Qualität entwickelt worden sind.[12] Die Urteile der *Studenten* als Adressaten der Leistungsangebote nehmen der dienstleistungstheoretischen Betrachtungsweise entsprechend insofern eine wichtige Rolle im Rahmen der Qualitätsmessung ein.

Universitätsdienstleistungen weisen jedoch gegenüber anderen Dienstleistungsangeboten verschiedene *Spezifika* auf, denen bei der Gestaltung eines Evaluationsinstrumentes auf angemessene Weise Rechnung zu tragen ist.[13] Von zentraler Bedeutung kann dabei die Tatsache angesehen werden, daß Universitäten einen *Erziehungsauftrag* zu erfüllen haben[14], was zur Folge hat, daß eine alleinige Ausrichtung der Konzeptualisierung und Messung des Qualitätskonstruktes an den Bedürfnissen der »Kunden« nur dann gerechtfertigt erscheint, wenn diese nicht im Konflikt mit den gesellschaftlich formulierten Aufgaben einer Hochschule stehen. Eine Möglichkeit zur Integration dieser Spezifika von universitären Lehrangeboten sehen wir darin, zusätzlich zu den Urteilen der Studenten auch die Anforderungen und Meinungen der *Dozenten* als Leistungsanbieter und »Erzieher« bei der *Instrumentenentwicklung und -anwendung* einzubeziehen. Das Instrument *TEACH-Q* stellt somit eine Kombination kunden- und anbieterseitiger Evaluationsansätze dar.

Weiterhin erachten wir als zentrale Kriterien im Rahmen der Instrumentenentwicklung zum einen die *Güte* des Instrumentes, d.h. seine Zuverlässigkeit und Gültigkeit im Hinblick auf das Konstrukt Qualität der Lehre, sowie zum anderen seine *Handhabbarkeit* im Vorlesungsalltag: Da ein Instrument zur Messung der Lehrqualität breite Akzeptanz unter den Studenten finden soll, zugleich aber nur ein vergleichbar geringes Zeitkontingent zur Ausfüllung zur Ver-

fügung steht, darf eine »kritische Länge« nicht überschritten werden. Im Hinblick auf die Beurteilung der Güte des Instrumentes halten wir eine Kombination der Reliabilitätskriterien der ersten Generation mit solchen der zweiten Generation und kausalanalytischen Methoden für angemessen. Güte und Handhabbarkeit stellen unserer Ansicht nach keine grundsätzlichen *Gegensätze* dar, sondern können gemeinsam realisiert werden.

Das Leistungsangebot von Universitäten umfaßt eine *Vielzahl von Leistungselementen*, die in ihrer Gesamtheit die Qualität der Lehre determinieren. Neben den verschiedenen Formen von Lehrveranstaltungen (Vorlesungen, Übungen, Seminare, Kolloquien etc.) sind hier insbesondere das Auftreten der einzelnen Lehrstühle insgesamt und die von ihnen zu erbringenden veranstaltungsbegleitenden Leistungen (Sprechstundenangebot, Bibliothek, Betreuung von Haus- und Diplomarbeiten etc.) zu nennen. Da eine Entwicklung spezifischer Meßinstrumente für eine derartige Vielzahl von Qualitätselementen aus Zeit- und Kostengründen im Rahmen einer einzelnen Untersuchung nicht zu leisten ist, beschränken wir uns im weiteren auf die Bewertung von *Vorlesungen*. Diese erscheinen uns als *zentrale Größe* für die Qualität der Lehre, da sie in Zeiten großer Studentenzahlen eine wichtige Rolle bei der Vermittlung von Lerninhalten einnehmen; eine Verbesserung der Lehrleistungen auf diesem Gebiet erreicht eine Vielzahl von Studenten zugleich. Dabei beziehen sich unsere Ausführungen ausschließlich auf Vorlesungen innerhalb des Fachbereichs *Wirtschaftswissenschaften*. Eine Übertragbarkeit unserer Ergebnisse auf andere Fachbereiche halten wir aufgrund der geschilderten Spezifika für problematisch und bedarf weiterführender Untersuchungen.

Im Rahmen unserer Untersuchung der Vorlesungsqualität unterscheiden wir somit zwischen drei Bereichen der Instrumentenentwicklung, die in den nachfolgenden Kapiteln erörtert werden:

- die Entwicklung eines Qualitätsmodells aus *Studentensicht* (Kapitel 2)
- die Ermittlung *pädagogisch relevanter* Qualitätsmerkmale (Kapitel 3), und schließlich

- die *Zusammenführung und Abstimmung* beider Sichtweisen zu einem ganzheitlichen Forschungsinstrumentarium TEACH-Q (in Kapitel 4).

2. Entwicklung eines Qualitätsmodells aus Studentensicht

2.1. Forschungsablauf

Zur Messung der Vorlesungsqualität aus Sicht der Studenten wurde ein mehrstufiges Vorgehen gewählt. Im einzelnen können folgende Schritte unterschieden werden:

a) *Itemgenerierung*. In einer umfangreichen Explorationsphase wurden neben der Auswertung bisheriger Lehrevaluationen insgesamt rund 150 qualitative Befragungen von Studenten durchgeführt. [15] Mit Hilfe der *Methode der Kritischen Ereignisse* [16] wurden dabei solche Ereignisse identifiziert, bei denen aufgrund ihrer jeweiligen Relevanz für die befragten Studenten davon ausgegangen werden kann, daß sie einen Rückschluß auf die der Qualitätsbeurteilung zugrundeliegenden Leistungsmerkmale ermöglichen. Insgesamt wurden auf diesem Wege 77 Items generiert, bei denen ein solcher Zusammenhang zur Vorlesungsqualität angenommen werden kann. [17]

b) *Dimensionenermittlung*. Auf Grundlage der ermittelten Items wurde ein vorläufiger Fragebogen erstellt, der in 7 wirtschaftswissenschaftlichen Vorlesungen verteilt wurde. Anhand von insgesamt 509 auswertbaren Antwortbögen wurde im Anschluß daran eine faktorenanalytische Reduktion der Beurteilungen der jeweils auf fünf-stufigen Skalen abgefragten Einzelmerkmale auf gemeinsame Dimensionen vorgenommen. Unter Verwendung der Hauptkomponentenanalyse konnten auf diese Weise insgesamt 14 Faktoren bzw. Dimensionen der Vorlesungsqualität identifiziert werden, wobei jedoch von den extrahierten Faktoren nur rund 60% der Itemvarianz erklärt wurde. [18] Zusätzliche Schwierigkeiten bereitete die stark unterschiedliche Verteilung der Items auf die Faktoren sowie die z. T. sehr schwierige inhaltliche Interpretation der gebildeten Faktoren.

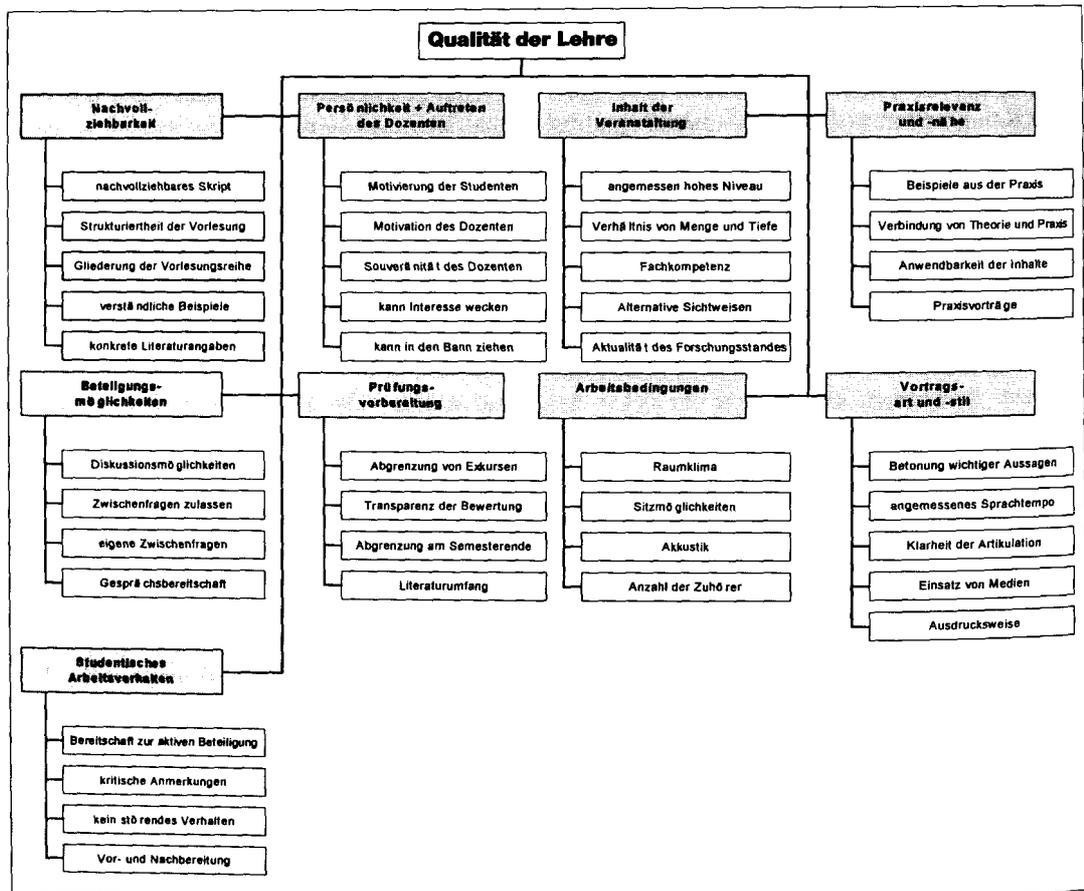


Abb. 1: Dimensionen und Items der Qualität der Lehre

Die genannten und allgemein bekannten methodenbedingten Probleme ließen eine weitere inhaltliche Überarbeitung des Instrumentes als notwendig erscheinen.

c) *Instrumentenüberarbeitung.* Die faktorenanalytisch extrahierten Qualitätsdimensionen wurden anschließend unter Einbeziehung von Studenturteilen inhaltlich restrukturiert und in einigen Fällen um weitere Items ergänzt. Das Ergebnis der qualitativen Überarbeitung des Instrumentes umfaßte schließlich 9 Dimensionen, die jeweils anhand von vier bis sechs Items operationalisiert wurden (Multi-Item-Messung). Die obige Abb. 1 zeigt die Dimensionen und die dazugehörigen Items.

d) *Validierung des Instrumentes.* Der überarbeitete Fragebogen wurde wiederum einem Feldtest unterzogen. Auf der Grundlage von 570

in 6 verschiedenen, systematisch ausgewählten Veranstaltungen (BWL/VWL; Grund-/ Hauptstudium; Wahl-/ Pflichtveranstaltung) ausgefüllten Fragebögen wurde die Relevanz der verschiedenen Fragebogenelemente für die Qualität der Lehre untersucht. Im Rahmen eines Methodenvergleiches wurden verschiedene multiattributive Qualitätsmodelle verglichen. Das vorhandene Frageninventar und die unterstellte Struktur des Konstruktes Qualität der Lehre wurde verschiedenen Methoden zur Beurteilung seiner Reliabilität und Validität unterworfen und auf diesem Wege ein Kausalmodell der Qualität von Vorlesungen aus Studentensicht an wirtschaftswissenschaftlichen Fachbereichen entwickelt.

Abb. 2 verdeutlicht die gesamte Vorgehensweise im Rahmen der Entwicklung eines kundenseitigen Qualitätsmodells zur Vorlesungsbe-

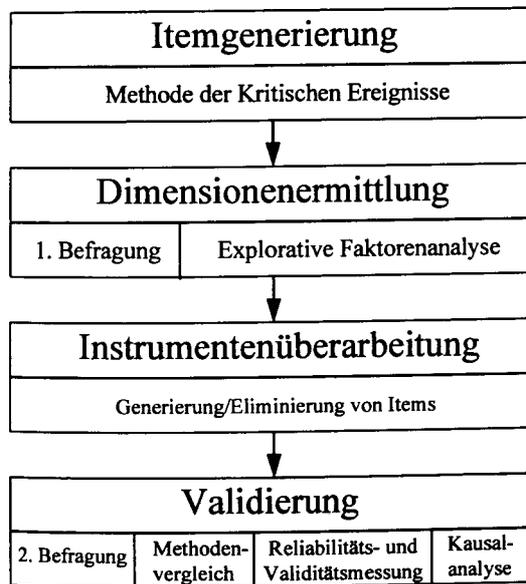


Abb. 2: Vorgehensweise bei der Entwicklung eines Instrumentes zur Messung der Qualität der Lehre aus Kundensicht

wertung. Die Abschnitte der Itemgenerierung und der Dimensionenermittlung sind bereits Gegenstand der Arbeit von Wochnowski/Devries (1994) gewesen. Im Mittelpunkt der nachfolgenden Ausführungen stehen die verschiedenen Vorgehensweisen, die die Überprüfung der Validität des überarbeiteten Qualitätsmodells zum Gegenstand haben.

2.2. Reliabilität und Validität von TEACH-Q: Prozesse und Ergebnisse

Bei der Analyse der Validität von TEACH-Q können zwei Teilbereiche unterschieden werden. Dabei wurde zum einen untersucht, welches Multiattributmodell am besten geeignet erscheint, die merkmalsbezogenen Beurteilungen der Studenten zu einem Gesamtwert der Qualität der Lehre zu verdichten. Zum anderen wurde im Rahmen eines komplexen Forschungsprozesses eine empirische Konzeptualisierung des Konstruktes Qualität der Lehre vorgenommen und deren Validität kausalanalytisch überprüft.

Methodenvergleich. Neben der eigentlichen

Beurteilung der Veranstaltung wurden die Studenten im Rahmen der zweiten Befragung gebeten, auch ihre individuellen *Erwartungen* und *Wichtigkeitsurteile* hinsichtlich der aufgeführten Leistungsmerkmale mitzuteilen. Auf Grundlage dieser Daten wurden dann von uns vier verschiedene weit verbreitete multiattributive Modelle der Qualitätsmessung hinsichtlich ihrer Eignung bei der Ermittlung ganzheitlicher Qualitätswahrnehmungen verglichen.[19] Im einzelnen waren dies

- das *SERVQUAL-Modell* von Parasuraman/Zeithaml/Berry (1988), das die Differenzen von Erwartungen und Erfahrungen über alle einbezogenen Merkmale aufsummiert. Die Gesamtqualität ist umso größer, je größer die Erfahrungen die Erwartungen überschreiten.
- das *Adequacy-Importance-Modell*[20], das die Merkmalsbeurteilung mit ihren jeweiligen Wichtigkeiten multipliziert und die daraus resultierenden Eindruckswerte aufsummiert.
- das *Idealpunktmodell* nach Trommsdorff (1975), das die Summe der merkmalsbezogenen Differenzen zwischen Idealausprägungen und tatsächlicher Leistung bildet. Im Unterschied zu SERVQUAL ist die Qualität hier umso höher, je dichter Erfahrungen und Erwartungen beieinander liegen.
- ein *Performance-Modell* der Qualitätsmessung (SERVPERF[21]), das ausschließlich die (ungewichtete) Beurteilung der attributbezogenen Leistungsbeurteilungen zu einem Gesamturteil addiert.

Bei der Abwägung der Eignung der Modelle für die Messung der Qualität der Lehre müssen verschiedene Aspekte berücksichtigt und gegeneinander abgewogen werden.[22] So weisen Multiattributmodelle gegenüber Gesamtbewertungen den wichtigen Vorteil auf, neben den Kundenurteilen auch die *Gründe* für deren Bewertung zu erfassen und somit eine Stärken-Schwächen-Analyse zu ermöglichen. Die *Gewinnung von verwertbaren Informationen* kann somit als erste wichtige Funktion der multiattributiven Qualitätsmessung gesehen werden.

Daneben ist jedoch auch stets die *Güte* der berechneten Qualitätswerte zu berücksichtigen. Von dem multiattributiv ermittelten Gesamtqualitätswert, der im Idealfall mit dem nicht

	SERVQUAL	Adequacy-Importance	Idealpunkt	Performance (SERVPERF)
Korrelation zum Gesamturteil	0,630	0,755	0,646	0,758

Tab. 1 Gütemaße verschiedener Multiattributmodelle

bekanntem »wahren« Wert identisch ist, wird ein Rückschluß auf die »wahre« Qualität der Veranstaltung vorgenommen und die Beurteilung der Vorlesung und des Dozenten daran ausgerichtet. Wir sprechen in diesem Zusammenhang von einer zweiten wichtigen Funktion der Qualitätsmessung, der sog. *Prognosefunktion*. Um einen Überblick über die prognostische Güte der verschiedenen Qualitätsmodelle zu gewinnen, haben wir die kompositionellen Qualitätswerte der verschiedenen Modelle ins Verhältnis zum Gesamturteil der Studenten gesetzt (vgl. Tab. 1). [23]

Das Ergebnis zeigt, daß das SERVPERF-Modell und der Adequacy-Importance-Ansatz mit einem r von jeweils rund 0,76 die größte Nähe zur Gesamtbewertung der Studenten aufweisen, während das SERVQUAL-Modell und das Idealpunktmodell deutlich geringere Korrelationen erzielen. Im Hinblick auf die Erfüllung der Prognosefunktion der Qualitätsmessung kann auf Grundlage der verwendeten Daten gefolgert werden, daß ein Einsatz von Differenzmodellen im Rahmen der Vorlesungsevaluation als Spezialfall der Dienstleistungsqualitätsmessung im Hinblick auf die Schätzung des wahren Qualitätswertes der Veranstaltung mit einem *signifikanten Informationsverlust* verbunden ist und insofern vermieden werden sollte [24]. [25] Die Einbeziehung von *Wichtigkeiten*, wie sie im Fall des Adequacy-Importance-Modells vorgenommen wird, führt zwar zu keinem signifikanten Verlust an Informationen, kann aber zugleich auch keine zusätzlichen Erkenntnisse beisteuern. Die den Zweikomponentenverfahren (hier: SERVQUAL, Adequacy-Importance) zugrundeliegenden mathematischen Modellannahmen erscheinen im Anwendungsfall nicht bzw. nur begrenzt in der Lage zu sein, die ko-

gnitive Algebra der Studenten bei der Qualitätsbeurteilung von Vorlesungen nachzubilden.

Da bei einer Beschränkung auf die Beurteilungen der Studenten eine Vielzahl möglicher Informationen nicht erhoben wird, diesen aber eine wichtige Stellung (insbesondere im Hinblick auf die anschließende Beseitigung bestehender Qualitätsdefizite) zukommt, schlagen wir im Umgang mit TEACH-Q eine *mehrstufige* Vorgehensweise vor. Danach finden bei der Berechnung eines Gesamtwertes der Vorlesungsqualität aus Studentensicht ausschließlich die *Beurteilungen* bzw. Erfahrungen Berücksichtigung, was dem Einkomponenten-Modell *SERVPERF* entspricht. Um eine nähere Analyse der Ursachen und Entwicklungen des Qualitätswertes zu erhalten (Aufzeigen von veranstaltungsspezifischen Studentenansprüchen, Anspruchsniveauveränderungen im Zeitablauf), werden jedoch zusätzlich auch die *Erwartungen* [26] der Studenten erfragt und auf Merkmalebene ausgewertet, wobei eine separate Erhebung am Semesterbeginn als vorteilhaft angesehen werden kann. [27] Eine Verknüpfung beider Datensätze zur Berechnung von Qualitätswerten findet hingegen aus den vorgenannten Gründen *nicht statt*.

Überprüfung und Modifikation der gewählten Konzeptualisierung von Qualität der Lehre aus Studentensicht. Anhand der Ergebnisse der zweiten Befragung von Studenten im Rahmen der Instrumentenentwicklung wurden zunächst die theoretischen Qualitätsdimensionen auf interne Konsistenz überprüft. Neben der Berechnung des Koeffizienten Cronbachs α als Reliabilitätskriterium der ersten Generation wurde für jede theoretische Dimension eine konfirmatorische Faktorenanalyse als Verfahren der zweiten Generation durchgeführt (vgl. Homburg 1995, S. 80 ff.; Homburg/Giering 1996, S. 8 ff.). [28] Weiterhin wurden die Korrelationen der einzelnen Items zum separat erhobenen Gesamturteil berechnet, die einen ersten Aufschluß über die Relevanz der einzelnen Items für die Qualitätswahrnehmung der Studenten geben. [29] Die Tab. 2 zeigt die Ergebnisse der Untersuchungen im Überblick.

Aus der Gesamtmenge der Indikatoren wurden im Anschluß wegen mangelnder Reliabilität die Indikatoren »nachvollziehbares Skript«,

TEACH-Q: Ein Instrument zur Bewertung von Vorlesungen

Dimension	Merkmal	Homogenität der Dimension (Cronbachs α) [41]	Homogenität der Dimension nach Entfernung des Items [42]	Korrelation zum Gesamturteil	Indikator-Reliabilität (konfirmatorische Faktorenanalyse) [43]	Verbleib des Items [44]
Nachvollziehbarkeit der Veranstaltungsinhalte	Skript Strukturierung Vorlesungsreihe Beispiele konkrete Literaturangaben	0,72	0,78	0,19	0,02/-	Nein
			0,58	0,72	0,72/0,74	Ja
			0,61	0,61	0,81/0,83	Ja
Auftreten und Wirken des Dozenten	kann motivieren ist selbst motiviert souverän Interesse wecken in den Bann ziehen	0,91	0,87	0,75	0,83/0,83	Ja
			0,89	0,65	0,69/0,69	Ja
			0,90	0,62	0,55/0,55	Ja
Vorlesungsinhalte	wissenschaftliches Niveau Verhältnis von Menge und Tiefe Fachkompetenz Alternative Sichtweisen Aktualität der Forschung	0,71	0,68	0,28	0,37/0,33	Ja
			0,47	0,47	0,24/-	Nein
			0,64	0,52	0,41/0,33	Ja
Praxisbezug der Vorlesung	aktuelle praktische Beispiele Verbindung von Theorie und Praxis Anwendbarkeit Praxisvorträge	0,81	0,64	0,32	0,48/0,39	Ja
			0,64	0,30	0,48/0,39	Ja
			0,66	0,32	0,42/0,46	Ja
Beteiligungsmöglichkeiten der Studenten	aktuelle praktische Beispiele Verbindung von Theorie und Praxis Anwendbarkeit Praxisvorträge	0,57	0,71	0,44	0,76/0,72	Ja
			0,71	0,46	0,77/0,85	Ja
			0,76	0,48	0,36/0,35	Ja
Prüfungsvorbereitung	Diskussionsmöglichkeiten Zwischenfragen zulassen Zwischenfragen stellen Gesprächsbereitschaft	0,79	0,41	0,28	0,43/0,28	Ja
			0,43	0,40	0,61/0,90	Ja
			0,64	-0,11	0,08/-	Nein
Arbeitsbedingungen	Abgrenzung Transparenz Literaturumfang Abgrenzung am Semesterende	0,63	0,48	0,40	0,29/0,26	Ja
			0,76	0,40	0,40/0,42	Ja
			0,71	0,47	0,62/0,62	Ja
Vortragsstil	Raumklima Sitzmöglichkeiten Akustik Zahl der Hörer	0,63	0,76	0,38	0,38/0,69	Nein
			0,70	0,38	0,74/0,69	Ja
			0,57	-0,13	0,33/0,39	Ja
Studentisches Verhalten während der Veranstaltung	Herausstellen wichtiger Aussagen Sprechtempo Artikulationsvermögen Medieneinsatz Ausdrucksweise	0,69	0,51	0,63	0,64/0,66	Ja
			0,66	0,49	0,57/0,50	Ja
			0,50	0,12	0,15/-	Nein
Studentisches Verhalten während der Veranstaltung	aktive Beteiligung Anmerkungen kein störendes Verhalten Nachbereitung	0,69	0,87	0,20	0,53/0,55	Ja
			0,86	0,63	0,53/0,53	Ja
			0,87	0,49	0,50/0,50	Ja
Studentisches Verhalten während der Veranstaltung	aktive Beteiligung Anmerkungen kein störendes Verhalten Nachbereitung	0,69	0,82	0,17	0,83/0,83	Ja
			0,84	-0,02	0,56/0,56	Ja
			0,83	0,71	0,77/0,77	Ja
Studentisches Verhalten während der Veranstaltung	aktive Beteiligung Anmerkungen kein störendes Verhalten Nachbereitung	0,69	0,58	0,17	0,64/0,66	Ja
			0,55	0,17	0,64/-	Ja
			0,69	-0,02	0,21/-	Nein
Studentisches Verhalten während der Veranstaltung	aktive Beteiligung Anmerkungen kein störendes Verhalten Nachbereitung	0,69	0,67	0,20	0,19/-	Nein
			0,67	0,20	0,19/-	Nein
			0,67	0,20	0,19/-	Nein

Tabelle 2: Gütebetrachtung der Faktoren und Einzelitems

»konkrete Literaturangaben« (beide Dimension Nachvollziehbarkeit), »Verhältnis von Stoffmenge und -tiefe« (Dimension Vorlesungsinhalte), »Praxisvorträge« (Dimension Praxisbezug), »Dozent stellt Zwischenfragen« (Dimension Beteiligungsmöglichkeiten), »Literaturumfang« (Dimension Prüfungsvorbereitung), »Akkustik« (Dimension Arbeitsbedingungen) sowie die Items »störendes Verhalten« und »Nachbereitung der Veranstaltung« (Dimension Studentisches Verhalten) eliminiert, so daß insgesamt 31 Indikatoren in der Analyse verbleiben.

Als zentrales Ergebnis der Aussagen der Reliabilitätsuntersuchung kann festgehalten werden, daß nach Eliminierung der genannten Items die theoretischen Faktoren (mit Ausnahme der Dimensionen »Beteiligungsmöglichkeiten« und »Arbeitsbedingungen«) ein vergleichsweise hohes Maß an empirischer Intra-Homogenität aufweisen. Die Inhomogenität der beiden genannten Dimensionen geht einher mit durchweg schwachen Korrelationen der einzelnen Items mit der ganzheitlichen Qualitätswahrnehmung, was eine geringe Relevanz der Dimensionen nahelegt. Insbesondere bei den Items der personenbezogenen Dimensionen »Dozentenpersönlichkeit« und »Vortragsstil« sowie »Nachvollziehbarkeit der Veranstaltung« bestehen hohe Korrelationen mit dem Gesamturteil.

Eine *explorative Faktorenanalyse* auf der Grundlage der Dimensionenmittelwerte zeigte, daß zwischen den Dimensionen der Qualität der Lehre keine Unabhängigkeit vorliegt. Stattdessen wurden *zwei Faktoren* identifiziert, bei denen es sich bei dem ersten Faktor um einen Zusammenschluß solcher Aspekte handelte, die durch den Dozenten unmittelbar beeinflusst werden können (Nachvollziehbarkeit, Auftreten, Vorlesungsinhalte, Praxisbezug, Beteiligungsmöglichkeiten, Prüfungsvorbereitung, Vortragsstil). Der zweite Faktor umfaßte hingegen mit den Arbeitsbedingungen und dem studentischen Verhalten solche Aspekte, die durch den Dozenten nicht bzw. nur mittelbar beeinflusst werden können.

Eine im Anschluß durchgeführte *konfirmatorische Faktorenanalyse* gab Aufschluß darüber, daß die zweite Dimension jedoch nicht über ein

ausreichendes Maß an Homogenität verfügt, sondern stattdessen Arbeitsbedingungen und Studentenverhalten eigenständige Dimensionen der Lehrqualität darstellen.[30] Eine solche *dreidimensionale Struktur* des Konstruktes Qualität der Lehre aus Studentensicht, die zwischen den Dimensionen Dozentenqualität, Studentenqualität und Umfeldqualität unterscheidet (siehe Abb. 3), weist eine hohe Diskriminanzvalidität auf und erfüllt die Gütekriterien der konfirmatorischen Faktorenanalyse.[31] Die einzelnen Dimensionen des Konstruktes Qualität der Lehre werden dabei durch 26 Indikatoren (Dimension Dozentenqualität), 2 Indikatoren (Studentenqualität) respektive 3 Indikatoren (Umfeldqualität) operationalisiert.

Kausalanalytische Überprüfung der Validität der gewählten Konzeptualisierung. Im weiteren Verlauf der Untersuchung galt es, Aussagen über die Güte der vorgenommenen Konzeptualisierung zu formulieren sowie die Relevanz der verschiedenen Dimensionen und Faktoren für das Qualitätsurteil der Studenten zu ermitteln. Dabei wurde ein *kausalanalytisches Untersuchungsdesign* gewählt, welches die Dimensionen der Qualität der Lehre als Determinanten des ganzheitlichen Qualitätsurteils interpretiert. Die Höhe der Erklärung des Gesamturteils durch die Dimensionen kann dabei als Indiz für die *Inhaltvalidität* der gewählten Konzeptualisierung und somit der Eignung der ihr zugrundeliegenden Indikatoren für die Qualitätsmessung angesehen werden.[32] Die folgende Abb. 4 verdeutlicht die untersuchte Modellstruktur und zeigt die Schätzergebnisse.[33]

Alle drei Dimensionen der Qualität der Lehre weisen erwartungsgemäß einen positiven Zusammenhang zur Gesamtbeurteilung der Studenten auf. Der mit Abstand stärkste Einfluß geht mit einem standardisierten Strukturkoeffizienten von 0,84 dabei auf die Dimension Dozentenqualität zurück, d.h. auf solche Qualitätsmerkmale, die unmittelbar durch den Dozenten variiert werden können. Hinsichtlich der beiden anderen Dimensionen kann ein weit geringerer Zusammenhang zum Gesamturteil festgestellt werden (Strukturkoeffizienten von 0,22 für die Studentenqualität bzw. 0,02 für die Umfeldqualität), wobei dieser Zusammenhang im Fall der Umfeldqualität zudem keine Signifi-

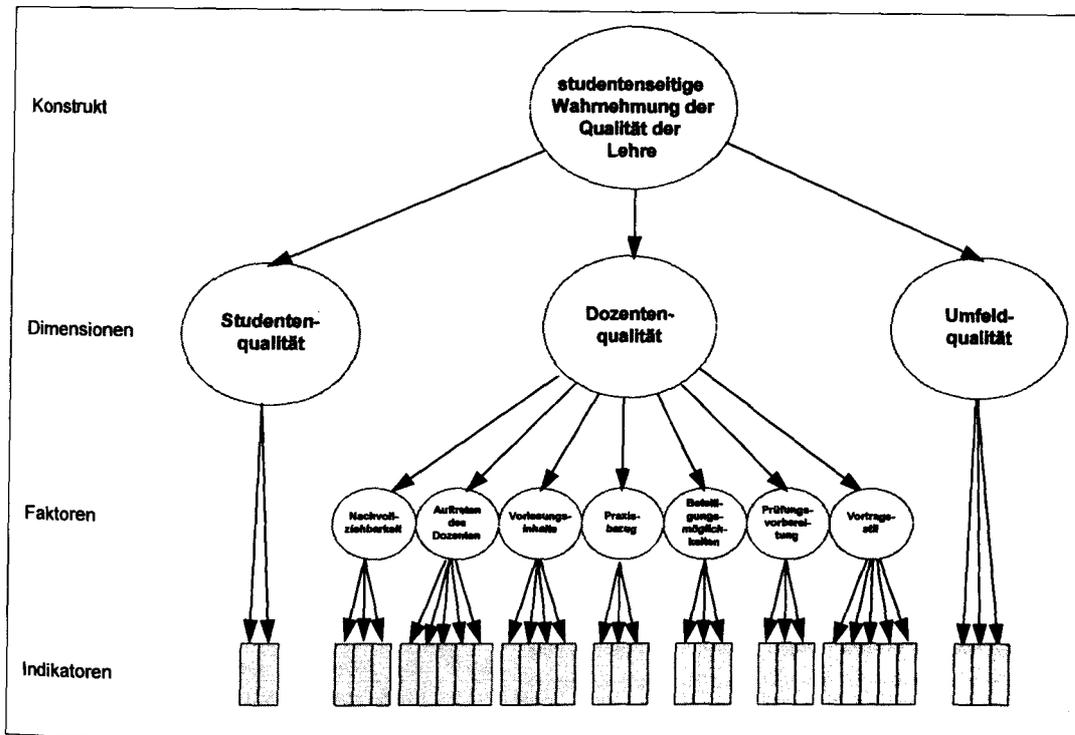


Abb. 3: Endgültige Konzeptualisierung des Konstruktes Qualität der Lehre aus Studentensicht

kanz aufweist (t-Wert des Strukturkoeffizienten kleiner 1,96 [34]). Insgesamt wird das Gesamturteil zu 76% durch die gewählte Konzeptualisierung der Qualität der Lehre erklärt, was eine *hohe Inhaltsvalidität* der einbezogenen Indikatormenge indiziert.

Die globale Anpassung des Modells kann mit Werten von GFI=0,94 und AGFI=0,92 als *gut* bezeichnet werden, [35] wobei durch die Einbeziehung auch vergleichsweise gering ladender, aber sachlich begründeter Indikatoren eine Abweichung vom methodischen Optimum bewußt in Kauf genommen wurde. Hinsichtlich der lokalen Anpassungsmaße werden die in der Literatur geforderten Richtwerte *nahezu vollständig übertroffen*. So liegen die Indikator-Reliabilitäten für alle Items mit Ausnahme der Variablen »Raumklima« und der Faktoren »Beteiligungsmöglichkeiten« und »Prüfungsvorbereitung« über 0,40. Die durchschnittlich erfaßte Varianz der Faktoren betragen 0,50 (Dimension Umfeldqualität), 0,66 (Studentenqualität) und 0,53 (Dozentenqualität).

Aufgrund der überragenden Bedeutung der Dimension Dozentenqualität stellt sich weiterhin die Frage, wie stark der Einfluß der verschiedenen Teilaspekte dieser Dimension auf die Gesamtbeurteilung der Vorlesungsqualität ist. Um die jeweilige *Wichtigkeit der Faktoren* zu ermitteln, wurden die Faktorkorrelationen mit dem Gesamturteil herangezogen. [36] Tab. 3 zeigt die Höhe der Korrelationen im einzelnen.

Die Ergebnisse zeigen zunächst, daß zwischen sämtlichen untersuchten Bereichen und der Gesamtbewertung der Vorlesung ein starker positiver Zusammenhang besteht. Darüber hinaus bestehen aber auch Unterschiede: So fließen insbesondere die Beurteilung des *Vortragstils*, die *Emotionalität und Motivation* des Dozenten sowie die Verständlichkeit seiner Ausführungen zu einem hohen Maße in die studentenseitige Qualitätsbeurteilung ein. Den Vorlesungsinhalten als eigentlichem »Kernelement« der Veranstaltung kommt hingegen ebenso wie der Prüfungsvorbereitung eine vergleichsweise geringere Bedeutung zu.

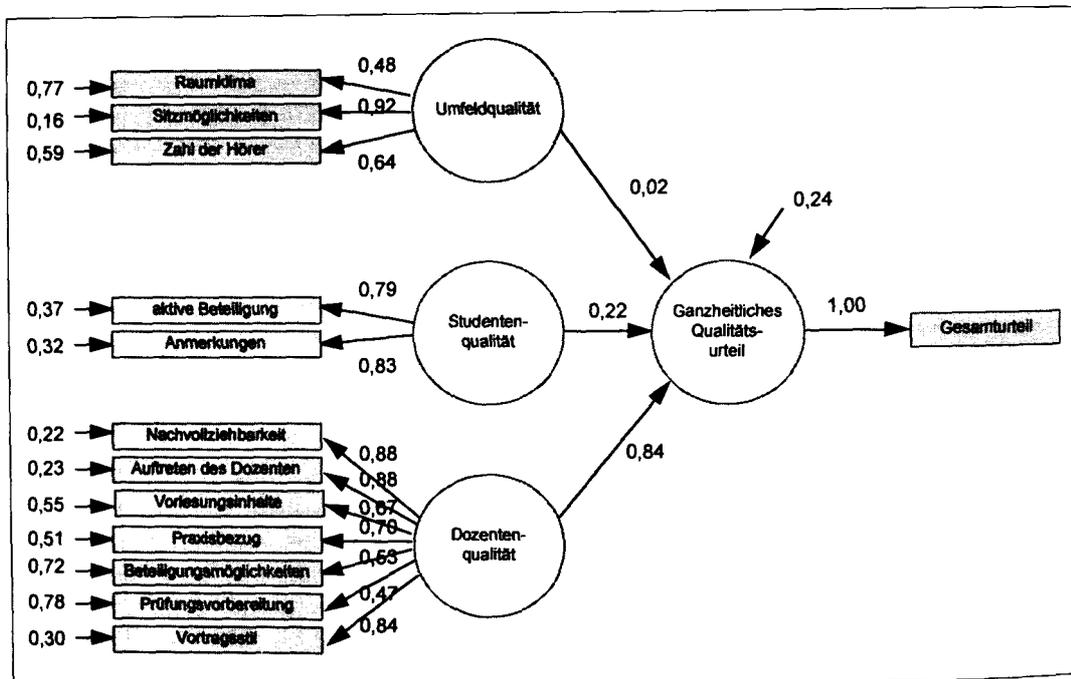


Abb. 4: Kausalmmodell der Qualität der Lehre aus Studentensicht

Zusammenfassend kann festgestellt werden, daß in erster Linie die Dozentenqualität, aber auch die Studentenqualität signifikant zur Qualitätsbeurteilung der Veranstaltung durch die Studenten beitragen und daher bei zukünftigen Befragungen von Studenten berücksichtigt werden müssen; sie stellen zentrale Elemente des

Befragungsinstrumentes TEACH-Q dar. Der in dieser Untersuchung als nicht-signifikant festgestellte Einfluß der Umfeldgestaltung auf die Vorlesungsqualität erscheint uns als äußerst dynamische Größe und soll deshalb trotz seiner zumindest zum Untersuchungszeitpunkt geringen Relevanz im Fragebogen verbleiben. [37]

Faktor	Wichtigkeit (Korrelation mit dem Gesamturteil)
Vortragstil	0,88
Auftreten und Wirken des Dozenten	0,87
Nachvollziehbarkeit der Veranstaltungsinhalte	0,84
Vorlesungsinhalte	0,64
Praxisbezug der Vorlesung	0,58
Prüfungsvorbereitung	0,57
Beteiligungsmöglichkeiten der Studenten	0,52

Tab. 3: Wichtigkeit der Faktoren der Dimension Dozentenqualität

3. Pädagogisch relevante Qualitätsaspekte

Ein wesentlicher Unterschied zwischen universitären Veranstaltungsangeboten und anderen Dienstleistungen kann, wie bereits erwähnt, darin gesehen werden, daß der kundenorientierte (subjektive) Qualitätsbegriff für Vorlesungen als Dienstleistungen aufgrund des Erziehungsauftrages der Hochschulen ergänzt werden muß um Merkmale, mit denen die eher »objektive« Perspektive der gesellschaftlichen bzw. (hochschul-)politischen Aufgabenstellung erfaßt wird. Ein Instrument zur Messung der

Qualität der Lehre muß folglich neben studentischen Qualitätsdimensionen auch *solche Merkmale enthalten*, auf die zwar von Seiten der Studenten kein bzw. nur geringer Wert gelegt wird, die aber dennoch einen Einfluß auf die Erfüllung des Bildungsauftrages der Hochschulen ausüben.

Wir sind davon ausgegangen, daß diese Qualitätsdimensionen der universitären Dienstleistung am ehesten von *Dozenten* artikuliert werden können. Um Aufschluß über solche Merkmale zu erhalten, denen bei der Bewertung von Vorlesungen eine besondere pädagogische Bedeutung zugewiesen wird, wurden von uns folglich im Rahmen der Entwicklung von TEACH-Q insgesamt 42 Dozenten des Fachbereichs Wirtschaftswissenschaften an der Universität Hannover bezüglich ihrer Erwartungen an die Qualität von Vorlesungen befragt. [38] Das Ergebnis zeigt zum einen, daß hinsichtlich der Erwartungen an einzelne Merkmale erhebliche Erwartungs-Diskrepanzen zwischen den Studenten und den Dozenten bestehen. Zum anderen wird ersichtlich, daß einzelne Merkmale existieren, die bei der Qualitätswahrnehmung aus Studentensicht keine bzw. nur eine vergleichsweise geringe Relevanz aufweisen, aber aus Sicht der »Leistungsanbieter« zentrale Dimensionen einer hohen Vorlesungsqualität darstellen. Diese Merkmale müssen folglich bei der Entwicklung eines Meßinstrumentes, das *sowohl Studenten- als auch Dozenteninteressen* berücksichtigt, beachtet werden.

Zusätzliche Bedeutung kommt aus Sicht der Dozenten vor allem solchen Aspekten zu, die das *Verhalten der Studenten* betreffen. Diese Aspekte strahlen nach Ansicht der Dozenten auf andere Qualitätsbereiche aus, die sich im primären Einflußbereich des Dozenten befinden, wie z.B. die Motivation der Dozenten und die Nachvollziehbarkeit der Vorlesungsinhalte. [39] Insbesondere können zwei Teilaspekte unterschieden werden:

- Die Störung der Veranstaltung durch studentische Gespräche während der Veranstaltung, laufendes bzw. verspätetes oder verfrühtes Kommen-und-Gehen etc.
- Die Bereitschaft der Studenten zur aktiven Beteiligung an der Veranstaltung (Bewertung

von Zwischenfragen, Verhalten bei »Cold Calls« des Dozenten etc.).

Darüber hinaus wird den Vorlesungsinhalten und hier vor allem dem »wissenschaftlichen Niveau« der Veranstaltung von Seiten der Dozenten ein deutlich höherer Stellenwert für die Qualität der Vorlesung zugewiesen als von den Studenten.

Um dem angestrebten Ziel der Entwicklung eines integrativen Evaluationsinstrumentes zu entsprechen und auf diesem Wege parteienübergreifende Akzeptanz zu erzielen, wurde das Meßinstrument TEACH-Q um solche Aspekte ergänzt, die zuvor im Rahmen der Konzeptualisierung der Vorlesungsqualität aus Studentensicht aufgrund mangelnder Reliabilität aus der Indikatorenmenge entfernt worden waren. Im einzelnen handelt es sich dabei um die Items »Störendes Verhalten von Kommilitonen« (ursprünglich Dimension Studentisches Verhalten) und »Zwischenfragen des Dozenten« (ursprünglich Dimension Beteiligungsmöglichkeiten der Studenten), die *zusätzlich* in den Fragebogen aufgenommen wurden.

4. Zusammenführung beider Sichtweisen zu einem ganzheitlichen Instrument

Die Kombination von studenten- und dozentenseitigen Qualitätsmerkmalen in TEACH-Q, wie sie in den vorangegangenen Abschnitten geschildert worden ist, wird ergänzt durch eine zeitliche und subjektbezogene Differenzierung des Instrumentes. Die folgende Abb. 5 gibt einen Überblick über die verschiedenen zeitlichen und subjektbezogenen Elemente der Messung der Vorlesungsqualität mittels TEACH-Q.

Die *Erwartungen* an die Veranstaltung werden demzufolge am Beginn eines jeden Semesters erhoben (ex-ante-Messung) (siehe Anhang 1). Am Semesterende erfolgt eine weitere Datenerhebung, bei der ausschließlich die *Bewertung der Veranstaltung* abgefragt wird (ex-post-Messung) (siehe Anhang 2). Gegenstand beider Befragungen ist jeweils die Gesamtmenge der bei der Konzeptualisierung des Konstruktes

	Dozenten	Studenten
Semesteranfang	Messung der Erwartungen	Messung der Erwartungen
Semesterende	Messung der Erfahrungen	Messung der Erfahrungen
Auswertung	Abgleich der Erfahrungen (GAP II)	Abgleich der Erwartungen (GAP I)
	Berechnung der Gesamtqualität	
	Dialogprozeß	

Abb. 5: Einsatz von TEACH-Q in der Vorlesungsevaluation

Vorlesungsqualität aus Studenten- und Dozenten-sicht generierten Indikatoren. Dem von uns zugrundegelegten kombinierten Qualitätsverständnis entsprechend werden beide Datenerhebungen sowohl bei den Studenten als auch bei dem jeweiligen Dozenten der Veranstaltung durchgeführt.

Bei der Auswertung der Daten im Rahmen von TEACH-Q können drei Auswertungsebenen unterschieden werden. Neben der multiattributiven Berechnung der *Gesamtqualität* aus Sicht der Studenten erfolgt in Form einer Gap-Analyse ein Abgleich der Erwartungen von Dozenten und Studenten (*GAP I*) ebenso wie eine Gegenüberstellung der Bewertungen beider Parteien (*GAP II*). Auf Grundlage der so gewonnenen Vergleichsdaten gilt es, im Rahmen eines veranstaltungsindividuellen *Dialogvorgangs* die Gründe für bestehende Abweichungen zwischen den Ansichten der Leistungsanbieter (Dozenten) und -nachfrager (Studenten) zu ermitteln und Maßnahmen zur gegenseitigen Annäherung einzuleiten (z. B. Erklärung des eigenen Standpunktes, Veränderung des Lehrangebotes), um ein höheres Qualitätsniveau (im Sinne des von uns verwendeten kombinierten Qualitätsverständnisses) zu erzielen.

Diesem Aspekt der Aussprache zwischen Dozenten und Studenten kommt eine zentrale Bedeutung für unser Vorhaben zu: Wir verstehen die Studenturteile insofern nicht als absolute Handlungsanweisungen, die es unreflek-

tiert umzusetzen gilt, sondern als Ausgangspunkt eines *intensiven Dialogprozesses*, der eine *gemeinsame* Optimierung des Lehrangebotes von Dozenten und Studenten zur Aufgabe hat.

5. Grenzen und Ansatzpunkte der Weiterentwicklung von TEACH-Q

Ziel der vorliegenden Untersuchung war es, mit TEACH-Q ein Instrument zur Messung der Qualität von universitären Vorlesungen zu entwickeln, das die Kriterien der Validität und der Handhabbarkeit gleichermaßen erfüllt und zugleich eine Integration studenten- und dozenten-seitiger Qualitätsvorstellungen leistet. Anhand verschiedener mathematisch-statistischer Methoden sowie von Plausibilitätsüberlegungen konnte gezeigt werden, daß diese Kriterien für die vorliegenden Daten als *weitgehend erfüllt* angesehen werden können.

Hinsichtlich der Frage, ob eine Übertragbarkeit des Instrumentes auf andere wirtschaftswissenschaftliche Fachbereiche in der Bundesrepublik möglich ist, müssen jedoch zusätzliche Überlegungen angestellt werden. Als problematisch könnte sich neben kulturellen und standortbezogenen Besonderheiten bei einem solchen räumlichen Instrumententransfer insbesondere die jeweilige wahrgenommene *Engpaß-situation* der untersuchten Universität als situativer Kontext auswirken. Trotz des von uns gemessenen geringen Einflusses solcher Kontextfaktoren haben wir deshalb verschiedene situative Aspekte (wie die Anzahl der Hörer, das Raumklima) bei der Fragebogengestaltung berücksichtigt.

Eine zentrale Aufgabe weiterer Forschung in diesem Bereich kann folglich in der Überprüfung der *Generalisierbarkeit* des Instrumentes für andere Universitäten und in diesem Zusammenhang die Unterscheidung von allgemein gültigen Basisdimensionen der Qualität der Lehre und anderen, situationspezifischen Dimensionen gesehen werden. [40] Darüber hinaus besteht zumindest in einer langfristigen Perspektive die Notwendigkeit, die Bewertung

TEACH-Q: Ein Instrument zur Bewertung von Vorlesungen

von Vorlesungen durch TEACH-Q auf andere Bereiche der Lehrqualität zu erweitern. So ist insbesondere an die Messung der (Dienstleistungs-)Qualität *anderer Veranstaltungsformen* (wie Seminare, Übungen etc.) und des Lehrstuhlangebotes insgesamt zu denken.

Anhang 1:

TEACH-Q - VORLESUNGSBEWERTUNG
TEIL 1: ERWARTUNGEN

- 1. Seite -

Ich erwarte von dieser Veranstaltung, daß...	trifft voll- kommen zu	trifft überhaupt nicht zu
die Veranstaltung strukturiert ist	1 2 3 4 5	
sie sich sinnvoll mit anderen Vorlesungen ergänzt	1 2 3 4 5	
daß das Verständnis durch Beispiele gefördert wird	1 2 3 4 5	
der Dozent in der Lage ist, die Studenten zu motivieren	1 2 3 4 5	
der Dozent selber motiviert ist	1 2 3 4 5	
der Dozent souverän auftritt	1 2 3 4 5	
der Dozent das Interesse der Studenten weckt	1 2 3 4 5	
der Dozent die Studenten in seinen Bann zieht	1 2 3 4 5	
der Dozent fachkompetent ist und sich das in der Veranstaltung widerspiegelt	1 2 3 4 5	
der Dozent aktuelle Beispiele aus der Praxis einbringt	1 2 3 4 5	
Theorie und Praxis miteinander verbunden werden	1 2 3 4 5	
die vermittelten Inhalte in der Praxis anwendbar sind	1 2 3 4 5	
über den aktuellen theoretischen Forschungsstand informiert wird	1 2 3 4 5	
Möglichkeiten zur Diskussion geboten werden	1 2 3 4 5	
die Studenten fähig sind, konstruktive Anmerkungen einzubringen	1 2 3 4 5	
das Raumklima (Luft, Temperatur) angenehm ist	1 2 3 4 5	
ausreichend Sitzmöglichkeiten vorhanden sind	1 2 3 4 5	

TEACH-Q - VORLESUNGSBEWERTUNG

TEIL 1: ERWARTUNGEN

- 2. Seite -

Ich erwarte, daß...	trifft voll- kommen zu	trifft überhaupt nicht zu			
der Dozent examensrelevante Inhalte von Exkursen abgrenzt	1	2	3	4	5
der Dozent seine Kriterien und Anforderungen bei der Prüfungsbewertung offenlegt	1	2	3	4	5
der Dozent am Semesterende eine Stoffabgrenzung vornimmt	1	2	3	4	5
wichtige Aussagen herausgestellt werden	1	2	3	4	5
der Dozent sich klar und deutlich artikuliert	1	2	3	4	5
der Dozent weder zu schnell noch zu langsam spricht	1	2	3	4	5
Medien (Tafel, Folien) eingesetzt werden	1	2	3	4	5
der Dozent eine verständliche Ausdrucksweise verwendet	1	2	3	4	5
der Dozent Zwischenfragen stellt	1	2	3	4	5
Kommilitonen die Veranstaltung nicht durch Gespräche untereinander stören	1	2	3	4	5
die Studenten sich aktiv beteiligen	1	2	3	4	5
der Dozent auch außerhalb der Veranstaltung ansprechbar ist	1	2	3	4	5
die Veranstaltung ein hohes wissenschaftliches Niveau aufweist	1	2	3	4	5
der Dozent studentische Zwischenfragen zuläßt	1	2	3	4	5
alternative Sichtweisen eines Themas berücksichtigt werden	1	2	3	4	5
die Vorlesung einen "persönlichen" Charakter besitzt	1	2	3	4	5

insgesamt eine sehr gute Vorlesung	1	2	3	4	5

TEACH-Q: Ein Instrument zur Bewertung von Vorlesungen

Anhang 2:

TEACH-Q - VORLESUNGSBEWERTUNG
TEIL 2: ERFahrungen

- 1. Seite -

Ich bin der Meinung, daß...	trifft voll- kommen zu	trifft überhaupt nicht zu
der Dozent examensrelevante Inhalte von Exkursen abgegrenzt hat	1 2 3 4 5	
der Dozent seine Kriterien und Anforderungen bei der Prüfungsbewertung offengelegt hat	1 2 3 4 5	
der Dozent am Semesterende eine Stoffabgrenzung vorgenommen hat	1 2 3 4 5	
wichtige Aussagen herausgestellt worden sind	1 2 3 4 5	
der Dozent sich klar und deutlich artikuliert hat	1 2 3 4 5	
der Dozent weder zu schnell noch zu langsam gesprochen hat	1 2 3 4 5	
Medien (Tafel, Folien) sinnvoll eingesetzt worden sind	1 2 3 4 5	
der Dozent eine verständliche Ausdrucksweise verwendet hat	1 2 3 4 5	
der Dozent gute Zwischenfragen gestellt hat	1 2 3 4 5	
Kommilitonen die Veranstaltung nicht durch Gespräche untereinander gestört haben	1 2 3 4 5	
die Studenten sich aktiv beteiligt haben	1 2 3 4 5	
der Dozent auch außerhalb der Veranstaltung ansprechbar war	1 2 3 4 5	
die Veranstaltung ein hohes wissenschaftliches Niveau aufgewiesen hat	1 2 3 4 5	
der Dozent studentische Zwischenfragen zugelassen hat	1 2 3 4 5	
alternative Sichtweisen eines Themas berücksichtigt worden sind	1 2 3 4 5	
die Vorlesung einen "persönlichen" Charakter besaß	1 2 3 4 5	

Insgesamt war die Vorlesung <i>sehr gut</i>	1 2 3 4 5	

TEACH-Q - VORLESUNGSBEWERTUNG
TEIL 2: ERFAHRUNGEN

- 2. Seite -

Ich bin der Meinung, daß...	trifft voll- kommen zu	trifft überhaupt nicht zu
der Dozent examensrelevante Inhalte von Exkursen abgegrenzt hat	1 2 3 4 5	
der Dozent seine Kriterien und Anforderungen bei der Prüfungsbewertung offengelegt hat	1 2 3 4 5	
der Dozent am Semesterende eine Stoffabgrenzung vorgenommen hat	1 2 3 4 5	
wichtige Aussagen herausgestellt worden sind	1 2 3 4 5	
der Dozent sich klar und deutlich artikuliert hat	1 2 3 4 5	
der Dozent weder zu schnell noch zu langsam gesprochen hat	1 2 3 4 5	
Medien (Tafel, Folien) sinnvoll eingesetzt worden sind	1 2 3 4 5	
der Dozent eine verständliche Ausdrucksweise verwendet hat	1 2 3 4 5	
der Dozent gute Zwischenfragen gestellt hat	1 2 3 4 5	
Kommilitonen die Veranstaltung nicht durch Gespräche untereinander gestört haben	1 2 3 4 5	
die Studenten sich aktiv beteiligt haben	1 2 3 4 5	
der Dozent auch außerhalb der Veranstaltung ansprechbar war	1 2 3 4 5	
die Veranstaltung ein hohes wissenschaftliches Niveau aufgewiesen hat	1 2 3 4 5	
der Dozent studentische Zwischenfragen zugelassen hat	1 2 3 4 5	
alternative Sichtweisen eines Themas berücksichtigt worden sind	1 2 3 4 5	
die Vorlesung einen "persönlichen" Charakter besaß	1 2 3 4 5	

Insgesamt war die Vorlesung <i>sehr gut</i>	1 2 3 4 5	

Anmerkungen

- [1] Vgl. Schmela (1994); Webler (1992); Freter (1994).
- [2] Vgl. z. B. Spindler (Hrsg.) (1968); Flechsig / Ritter (1975).
- [3] Vgl. z. B. die Untersuchungen von Freter (1994), Daniel (1994, 1995), Kromrey (1993), Preissler (1991) und Klatzer (1992).
- [4] Vgl. den Überblick bei Holtmann / Schnitzer (1992), S. 163 ff. und Ederleh (1994), S. 56 ff.
- [5] Vgl. Rindermann / Amelang (1994).
- [6] Rindermann und Amelang haben zudem einen Fragebogen für die Dozenten entwickelt; dieser enthält jedoch *keine differenzierte Beurteilung* der eigenen Veranstaltung. Vgl. Rindermann / Amelang (1994), Anhang.
- [7] Ebenso Rindermann / Amelang (1994), S. 7
- [8] Ebenso Daniel (1994, 1995) und Ederleh (1994), S. 56, der nahelegt, »zukünftig mehr auf die Initiativen von Fakultätentagen oder der fachlichen Vereinigungen zu setzen, die fachspezifische Varianten vergleichbar entwickeln und für den bundesweiten Einsatz empfehlen«.
- [9] Ebenso z. B. Gelfert (1993).
- [10] Vgl. Meyer (1986), S. 9; Stauss (1989), S. 48; Hilke (1989), S. 6.
- [11] Vgl. Parasuraman (1995), S. 144.
- [12] Vgl. die Übersichten bei Haller (1993) und Benkenstein (1993).
- [13] Vgl. Holtmann (1992), S. 167.
- [14] Vgl. Meffert (1995), S. 54.
- [15] Dieser Teil der Untersuchung war Gegenstand eines Marktforschungspraktikum an der Universität Hannover im Wintersemester 1994/95. Vgl. Wochnowski / Devries (1994).
- [16] Vgl. Stauss / Hentschel (1990).
- [17] Vgl. Wochnowski / Devries (1994), S. 61. Im Rahmen der Itemgenerierung wurde der dienstleistungstheoretischen Ausrichtung dieses Teils der Untersuchung entsprechend eine umfassende Interpretation des Konstruktes Qualität der Lehre zugrundegelegt. Folglich wurde bewußt auf eine Beschränkung des Objektbereichs etwa auf solche Merkmale, die dem unmittelbaren Einflußbereich des Dozenten unterliegen, verzichtet.
- [18] Als Extraktionskriterium wurde die Eigenwert-Regel nach Kaiser (vgl. Stewart 1981, S. 58) herangezogen.
- [19] Bei der Auswahl der Modelle wurden Modelle der Qualitäts-, Einstellungs- und Zufriedenheitsmessung gleichermaßen herangezogen, da das Verhältnis von Qualität und Zufriedenheit bzw. Einstellungen in der Marketingwissenschaft Gegenstand anhaltender Diskussionen ist (vgl. Taylor / Baker 1994; Cronin / Taylor 1992); Holbrook beschreibt in diesem Zusammenhang den Stand der (Dienstleistungs-)Qualitätsforschung als »wallow in a sea of confusion« (Holbrook 1994, S. 23).
- [20] Vgl. Freter (1979).
- [21] Vgl. Cronin / Taylor (1992).
- [22] Vgl. Zeithaml / Berry / Parasuraman (1996), S. 39f.
- [23] Bezüglich der Verwendung des Gesamturteils als Validitätsindikator muß unterschieden werden, ob das Gesamturteil den Rang eines »wahren« Wertes annimmt oder nur einen alternativen Ansatz der Messung darstellt; im ersten Fall handelt es sich um einen Spezialfall der Kriteriumsvalidität, im zweiten um einen Ansatz zur Bestimmung der Konstruktvalidität. Wir halten aufgrund der mit der Single-Item-Messung eines derartigen komplexen Konstruktes unvermeidlich verbundenen Meßfehler letztere Interpretation für sinnvoll: »How comfortable would we feel having our intelligence assessed on the basis of our response to a single question?« (Jacoby 1978, S. 93).
- [24] Vgl. ebenso Cronin / Taylor (1992), (1994); Babakus. / Boller (1992).
- [25] Für Differenzverfahren haben verschiedene Autoren gezeigt, daß diese generell über eine mangelnde theoretische Reliabilität verfügen (vgl. Peter / Churchill / Brown 1993; Brown / Churchill / Peter 1993; Spreng 1994), was seinen Ausdruck in niedrigen empirischen Validitätswerten auch in anderen Untersuchungsbereichen findet. So stellen selbst Parasuraman et al. als Proponenten des Differenzmodells Servqual fest: »However, in terms of predictive power alone, the perceptions-only measure is superior to the disconfirmation measures and, within the latter, the direct measure is superior to the difference-score measure« (Parasuraman / Zeithaml / Berry 1994, S. 214).
- [26] Den Erwartungen wird hier der Vorzug vor der Erhebung von Wichtigkeiten gegeben, da das meßtechnische Problem der Anspruchsinflation bei der Abfrage von Wichtigkeiten u. E. nach stärker anfällt als bei der Messung von Erwartungen. Alternative Formen der Wichtigkeiten-Messung, die trade-off-Prozesse der Konsumenten einbeziehen (z. B. Konstanzsummenmethoden, Conjoint-Measurement) sind bei der vorliegenden Vielzahl an Merkmalen nicht einsetzbar.
- [27] Eine ex-post-Messung von Erwartungen wird aus Gründen mangelnder Validität von den Autoren abgelehnt. Das mit der ex-ante-Messung von Erwartungen verbundene Problem dynamischer Einflüsse auf die Menge qualitätsrelevanter Dimensionen wird im vorliegenden Fall der Vorlesungsevaluation aufgrund der i. d. R. umfangreichen Erfahrungen der Studenten mit dem »Produkt« als weniger relevant angesehen.
- [28] Als Schätzverfahren wurde dabei (wie auch bei allen folgenden kausalanalytischen Berechnungen) das ULS-Verfahren (Unweighted Least-Squares) eingesetzt. Bei der Anwendung anderer Schätzverfahren (Maximum Likelihood, Generalised Least Squares) ergaben sich nur geringe Abweichungen der geschätzten Parameter, so daß von einer hohen Stabilität der Ergebnisse ausgegangen werden kann.

- [29] Die Messung des Gesamturteils, das der Ermittlung von Korrelationen hier zugrundliegt, erfolgte über die Frage: »Was ist Dein Gesamteindruck von dieser Vorlesung?« im Anschluß an die Abfrage der Einzelitems.
- [30] Ein Gutachter hat zurecht darauf hingewiesen, daß zwischen dem Verhalten der Studenten und der Dozentenqualität ein sachlogischer Zusammenhang der Gestalt besteht, daß eine geringe Dozentenqualität eine Verschlechterung der Studentenqualität bewirkt. Allerdings ist auch ein umgekehrter Zusammenhang denkbar: Eine geringe Studentenqualität vermindert die Motivation des Dozenten und senkt in der Folge seine Dozentenqualität. In der vorliegenden Untersuchung zeigen sich diese Zusammenhänge in einer Korrelation beider Dimensionen von 0,36.
- [31] Die globalen Anpassungsmaße betragen $GFI=0,98$ und $AGFI=0,97$. Die Indikator-Reliabilität betrug bei allen Indikatoren 0,30 oder größer; Ausnahmen stellten nur die Indikatoren »Raumklima« und »Prüfungsvorbereitung« dar (Indikator-Reliabilitäten von 0,24 bzw. 0,2). Die durchschnittlich erklärte Varianz der einzelnen Dimensionen liegt bei 0,52 (Situative Qualität), 0,67 (Studentenqualität) und 0,52 (Dozentenqualität). Vgl. für einen Überblick über die Gütemaße konfirmatorischer Faktorenanalysen Homburg / Giering 1996, S. 10 ff.
- [32] Vgl. zu dieser Interpretation auch Homburg (1995), S. 116.
- [33] Als Software wurde das Programm LISREL von Jöreskog und Sörbom in der Version 8.12 eingesetzt (vgl. Jöreskog. / Sörbom 1993), wobei wiederum die Methode der ungewichteten kleinsten Quadrate (ULS) als Schätzverfahren verwendet wurde. Den theoretischen Überlegungen entsprechend wurde die Fehlervarianz der Indikatorvariable »Gesamturteil« auf Null fixiert. Vgl. zum Ansatz der Kausalanalyse z.B. Backhaus et al. (1994).
- [34] Vgl. Backhaus et al. (1994), S. 405 ff.
- [35] Für einen Überblick über die verschiedenen Gütemaße kausalanalytischer Untersuchungen vgl. Fritz (1992) und Korte (1995), sowie die Hinweise von Homburg / Baumgartner (1995).
- [36] Im Output von LISREL finden sich diese Angaben unter der Bezeichnung »Correlations of Eta and Ksi«.
- [37] In der ersten Untersuchung im Jahre 1994, die noch unter erheblich schlechteren räumlichen Bedingungen in Hannover stattfand, ist entsprechend eine größere Bedeutung des Faktors Arbeitsbedingungen feststellbar.
- [38] Vgl. Wochnowski / Devries (1994), S. 18 f.
- [39] Vgl. Wochnowski / Devries (1994), S. 36 ff.
- [40] Alle Leser, die an einem Einsatz des Instrumentes interessiert sind, seien daher aufgefordert, sich mit den Autoren des Beitrages in Verbindung zu setzen.
- [41] Vgl. z. B. Churchill (1991), S. 498 f. Für die Beurteilung der Homogenität anhand des Koeffizienten Cronbachs α existieren keine absoluten Vorgabewerte. Eine häufig zitierte Konvention stellt die Annahme einer Dimension bei einem α -Wert von größer 0,7 dar.
- [42] Ebenfalls Cronbachs α -Werte nach Entfernung des jeweiligen Items.
- [43] Der erste Wert gibt die Reliabilität jedes Indikators im Rahmen einer konfirmatorischen Faktorenanalyse mit sämtlichen Indikatoren einer Dimension an. Der zweite Wert entspricht der Indikatorreliabilität nach Entfernung inhomogener Items. Der Gesamtfit sämtlicher konfirmatorischer Faktorenanalysen mit mehr als drei Indikatoren war mit GFI-Werten zwischen 0,99 und 1 und AGFI-Werten zwischen 0,96 und 1 jeweils sehr gut. Für Modelle mit drei Indikatoren verfügt die Analyse über keinerlei Freiheitsgrade und kann folglich auch keine globalen Anpassungsmaße berechnen.
- [44] Bei der Entscheidung über die Eliminierung und den verbleib von Indikatoren wurde nicht ein einzelner Wert betrachtet, sondern die verschiedenen Reliabilitätsmaße gemeinsam betrachtet. Zugleich wurde betrachtet, wie gut eine Dimension durch die Gesamtheit der Indikatoren gemessen wurde. Lagen etwa ausreichend andere Merkmale mit höheren Gütewerten zur Messung einer Dimension vor, erfolgte die Eliminierung eines Items bereits bei einer vergleichsweise hohen Indikator-Reliabilität, während Items mit ähnlich hohen Reliabilitätswerten bei weniger homogenen Dimensionen in der Analyse behalten wurden.
- [45] Bei nur zwei Indikatoren sind die Freiheitsgrade einer konfirmatorischen Faktorenanalyse negativ; Indikator-Reliabilitäten können folglich für diesen Fall nicht berechnet werden.

Verzeichnis der zitierten Literatur

- Babakus, E.; Boller, G. W. (1992): An Empirical Assessment of the SERVQUAL Scale, in: Journal of Business Research, Vol. 24, S. 253-268.
- Backhaus, K.; Erichson, B.; Plinke, W.; Weiber, R. (1994): Multivariate Analysemethoden - Eine anwendungsorientierte Einführung, Berlin u. a. 1994.
- Benkenstein, M. (1993): Dienstleistungsqualität - Ansätze zur Messung und Implikationen für die Steuerung, in: Zeitschrift für Betriebswirtschaft, 63. Jg. (1993), Heft 11, S. 1095-1116.
- Brown, T. J.; Churchill, G. A.; Peter, J. P. (1993): Improving the Measurement of Service Quality, in: Journal of Reatiling, Vol. 69 (1993), Spring, S. 127-139.
- Churchill, G. A. (1979): A Paradigm for Developing Better Measures of Marketing Constructs, in: Journal of Marketing Research, Vol. 16 (1979), February, S. 64-73.

TEACH-Q: Ein Instrument zur Bewertung von Vorlesungen

- Churchill, G. A. (1991): *Marketing Research: Methodological Foundations*, 5. A., Fort Worth 1991.
- Cronin, J. J.; Taylor, S. A. (1992): *Measuring Service Quality: A Reexamination and Extension*, in: *Journal of Marketing*, Vol. 56 (1992), July, S. 55-68.
- Cronin, J. J.; Taylor, S. A. (1994): *SERVPERF Versus SERVQUAL: Reconciling Performance-Based and Perception-Minus-Expectations Measurements of Service Quality*, in: *Journal of Marketing*, Vol. 58 (1994), January, S. 125-131.
- Daniel, H.-D. (1994): *Hörerbefragung an der Universität Mannheim: Konzeption, Erhebung, Auswertung*, in: *Empirische Pädagogik, Themenheft »Evaluation der Lehre«*, 8. Jg. (Heft 2), S. 109-129.
- Daniel, H.-D. (1995): *Bewertung der Lehre aus Sicht der Studierenden und Absolventen*, in: Müller-Böling, D. (Hrsg.): *Qualitätssicherung in Hochschulen*, Gütersloh 1995, S. 160-185.
- Ederleh, J. (1994): *Evaluation in Deutschland - Aktuelle Entwicklungen*, in: Seidel, H. (Hrsg.): *Evaluation der Lehre - Europäische Erfahrungen, deutsche Perspektiven*, Hannover 1994, S. 51-63.
- Flechsigg, K.-H.; Ritter, U. P. (1975): *Einführung zur Kursentwicklung in den Wirtschaftswissenschaften - Konstanzer Werkstattseminar, 2. A.*, Göttingen 1975.
- Freter, H. (1979): *Interpretation und Aussagewert mehrdimensionaler Einstellungsmodelle im Marketing*, in: Meffert, H.; Steffenhagen, H.; Freter, H. (Hrsg.): *Konsumentenverhalten und Information*, Wiesbaden 1979, S. 163-184.
- Freter, H. (1994): *Die studentische Evaluation von Lehrveranstaltungen, Arbeitspapier des Lehrstuhls Marketing der Universität Siegen*, Siegen 1994.
- Fritz, W. (1992): *Marktorientierte Unternehmensführung und Unternehmenserfolg*, Stuttgart 1992.
- Gelfert, H. D. (1993): *Universität der dritten Art. Die Universität muß ein modernes Dienstleistungsunternehmen werden*, in: *Deutsche Universitäts-Zeitung*, 50. Jg., Heft 17, S. 26-27.
- Haller, S. (1993): *Methoden zur Beurteilung von Dienstleistungsqualität*, in: *Zeitschrift für betriebswirtschaftliche Forschung*, 45. Jg. (1993), Heft 1, S. 19-40.
- Hilke, W. (1989): *Grundprobleme und Entwicklungstendenzen des Dienstleistungsmarketing*, in: Hilke, W. (Hrsg.): *Dienstleistungs-Marketing*, Wiesbaden 1989, S. 5-44.
- Holbrook, M. B. (1994): *The Nature of Customer Value*, in: Rust, R. T.; Oliver, R. L. (Hrsg.): *Service Quality: New Directions in Theory and Practice*, London 1994, S. 21-71.
- Holtmann, R. (1992): *Bericht aus der Arbeitsgruppe*, in: Ermert, K. (Hrsg.): *Lehre an Hochschulen: Über Kriterien und Instrumente zu ihrer Evaluation und Förderung*, Loccum Protokolle 11/92, Rehburg-Loccum 1992, S. 165-169.
- Holtmann, R.; Schnitzer, K. (1992): *Problemaufriß*, in: Ermert, K. (Hrsg.): *Lehre an Hochschulen: Über Kriterien und Instrumente zu ihrer Evaluation und Förderung*, Loccum Protokolle 11/92, Rehburg-Loccum 1992, S. 157-164.
- Homburg, C. (1995): *Kundennähe von Industriegüterunternehmen: Konzeption-Erfolgsaussichten-Determinanten*, Wiesbaden 1995.
- Homburg, C.; Baumgartner, H. (1995): *Beurteilung von Kausalmodellen. Bestandaufnahme und Anwendungsempfehlungen*, in: *Marketing ZFP*, 17. Jg., Heft 3, S. 162-176.
- Homburg, C.; Giering, A. (1996): *Konzeptualisierung und Operationalisierung komplexer Konstrukte*, in: *Marketing ZFP*, 18. Jg. (1996), Heft 1, S. 5-25.
- Jacoby, J. (1978): *Consumer Research: A State of the Art Review*, in: *Journal of Marketing*, Vol. 42 (April), S. 87-96.
- Jöreskog, K. G.; Sörbom, D. (1993): *LISREL 8: User's Reference Guide*, Chicago 1993.
- Klatzer, E. (1992): *Evaluation der Lehre durch das Zielpublikum, die StudentInnen*, in: Altrichter, H.; Schratz, M. (Hrsg.): *Qualität von Universitäten*, Innsbruck 1992, S. 197-211.
- Korte, C. (1995): *Customer Satisfaction Measurement: Kundenzufriedenheitsmessung als Informationsgrundlage des Hersteller- und Handelsmarketing am Beispiel der Automobilwirtschaft*, Frankfurt 1995.
- Kromrey, H. (1993): *Studentische Vorlesungskritik - Empirische Daten und Konsequenzen für die Lehre*, in: *Soziologie*, Heft 1, S. 39-56.
- Meffert, H. (1995): *Wir haben noch zahlreiche Defizite*, in: *Horizont*, 12. Jg., Nr. 49, S. 54.
- Meyer, A. (1983): *Dienstleistungs-Marketing, Erkenntnisse und praktische Beispiele*, 2. A., Augsburg 1986.
- Parasuraman, A. (1995): *Measuring and Monitoring Service Quality*, in: Glynn, W. J.; Barnes, J. G. (Hrsg.): *Understanding Services Management*, Chichester et al. 1995, S. 143-177.
- Parasuraman, A.; Zeithaml, V. A.; Berry, L. L. (1988): *SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality*, in: *Journal of Retailing*, Vol. 64 (1988), Spring, S. 12-40.
- Parasuraman, A.; Zeithaml, V. A.; Berry, L. L. (1994): *Alternative Scales for Measuring Service Quality: A Comparative Assessment Based on Psychometric and Diagnostic Criteria*, in: *Journal of Retailing*, Vol. 70 (1994), No. 3, S. 201-230.
- Peter, J. P.; Churchill, G. A.; Brown, T. J. (1993): *Caution in the Use of Difference Scores in Consumer Research*, in: *Journal of Consumer Research*, Vol. 19 (1993), March, S. 655-662.
- Preissler, R. (1991): *Die Qualität der Lehre: Ergebnisse einer Befragung an der Technischen Universität Berlin*, in: *Beiträge zur Hochschulforschung*, Heft 4, S. 401-429.
- Rindermann, H.; Amelang, M. (1994): *Das Heidelberger Inventar zur Lehrveranstaltungs-Evaluation (HILVE)*, Heidelberg 1994.
- Schmela, M. (1994): *Lehren muß gelernt sein*, in: *Rheinischer Merkur* v. 9. 12. 1994, S. 9.

Die Erfindung der doppelten Buchführung



Luca Pacioli

Abhandlung über die Buchhaltung von 1494
1997. 2., unveränderter Nachdruck der Ausgabe von 1933. 176 Seiten. Gebunden, DM 68,-

ISBN 3-7910-1188-X

Im Jahre 1494 erschien die erste gedruckte Darstellung der Buchhaltung in dem Sammelwerk („Summa“) des bekanntesten italienischen Mathematikers des 15. Jahrhunderts, des Franziskanerpaters Luca Pacioli. Es stellt die heute noch gebräuchliche Form der doppelten Buchführung dar, die als italienische Buchhaltung bekannt war. Die Übersetzung ist mit einer Einleitung über die italienische Buchhaltung sowie über das Leben und Werk von Luca Pacioli versehen. Der Reprint dieser Ausgabe ist ein kulturhistorisches Dokument.

Schäffer-Poeschel Verlag
Postfach 10 32 41 · 70028 Stuttgart
<http://www.schaeffer-poeschel.de>

**SCHÄFFER
POESCHEL**

- Spindler, D. (Hrsg.) (1968): Hochschuldidaktik: 25 Dokumente zur Hochschul- und Studienreform, Bonn 1968.
- Spreng, R.A. (1994): A Reassessment of the Reliability of Difference Scores in the Measurement of Disconfirmation of Expectations, in: Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, Vol. 7 (1994), S. 114-118.
- Stauss, B. (1989): Beschwerdepolitik als Instrument des Dienstleistungsmarketing, in: Jahrbuch der Absatz- und Verbrauchsforschung, 35. Jg. (1989), Heft 1, S. 41-62.
- Stauss, B.; Hentschel, B. (1990): Verfahren der Problementdeckung und -analyse im Qualitätsmanagement von Dienstleistungsunternehmen, in: Jahrbuch der Absatz- und Verbrauchsforschung, 36. Jg. (1990), Heft 3, S. 232-259.
- Stewart, D.W. (1981): Application and Misapplication of Factor Analysis in Marketing Research, in: Journal of Marketing Research, Vol. 18 (1981), No. 1, S. 51-62.
- Taylor, S.A.; Baker, T.L. (1994): An Assessment of the Relationship Between Service Quality and Customer Satisfaction in the Formation of Customer's Purchase Intentions, in: Journal of Retailing, Vol. 70 (1994), No. 2, S. 163-178.
- Thompson, K.N.; Getty, J.M. (1994): Structural Model of Relations Among Quality, Satisfaction, and Recommending Behavior in Lodging Decisions, in: Structural Equation Modeling, Vol. 1 (1994), Heft 2, S. 146-160.
- Trommsdorff, V. (1975): Die Messung von Produktimages für das Marketing, Köln u. a. 1975
- Webler, W.-D. (1992): Qualität der Lehre - Zwischenbilanz einer unübersichtlichen Entwicklung, in: Das Hochschulwesen, Heft 2, S. 153-173.
- Wohnowski, H.; Devries, J. (1994): Messung der Dienstleistungsqualität universitärer Dienstleistungsangebote - Entwicklung von TEACHQUAL, Lehr- und Forschungsbericht Nr. 27 des Lehrstuhls Markt und Konsum der Universität Hannover, Hannover 1994.
- Zeithaml, V.A.; Berry, L.L.; Parasuraman, A. (1996): The Behavioral Consequences of service Quality, in: Journal of Marketing, Vol. 60 (1996), April, S. 31-46.