

Ergebnisse geschlechtsspezifischer Auswertung studentischer Lehrveranstaltungsbeurteilung

Wirkungen des Geschlechts der Lehrenden auf die Lehrevaluation in ausgewählten Fächern der Geistes-, Gesellschafts- und Naturwissenschaften

2008



Inhaltsverzeichnis:

1.	Einleitung: Zielsetzung des Projekts	2
2.	Forschungsstand und Fragestellungen	2
3.	Methode und Datenbasis	5
4.	Ergebnisse	10
4.1	Ergebnisse zu Globalbewertungen	10
4.2	Ergebnisse zum Lehrfaktor „Organisation“	12
4.3	Ergebnisse zum Lehrfaktor „Interaktion“	18
5.	Diskussion der Ergebnisse und Ausblick	23
	Literaturverzeichnis	26
	Anhang	28

1. Einleitung: Zielsetzung des Projekts

Die Universität Duisburg-Essen (UDE) hat sich in ihren Ziel- und Leistungsvereinbarungen mit dem Land NRW zu einer umfassenden Gleichstellung verpflichtet, die auf allen Qualifikationsstufen weibliche Universitätsangehörige nachhaltig in ihrer weiteren Entwicklung fördert. Im Bereich der Lehre bedeutet dies, dass die UDE sich besonders um den Studienerfolg weiblicher Studierender und die Lehrqualifizierung des weiblichen wissenschaftlichen Nachwuchses bemüht. Im Rahmen der hochschulinternen Ziel- und Leistungsvereinbarungen zwischen dem Rektorat und den Fachbereichen werden diese daher nicht allein auf eine kontinuierliche Qualitätssicherung in der Lehre, sondern auch auf die Durchführung von Maßnahmen zur Verbesserung der Genderkompetenz festgelegt. In diesem Zusammenhang können auch die Ergebnisse der studentischen Lehrveranstaltungsbeurteilung eine Rolle spielen. So können geschlechtsspezifisch aufbereitete Ergebnisse eventuell Aufschluss über bestehende genderbezogene Problematiken geben.

Seit vier Semestern wird die studentische Lehrveranstaltungsbeurteilung durch das Zentrum für Hochschul- und Qualitätsentwicklung (ZfH) an der UDE zentral organisiert. Im vergangenen Sommersemester 2007 haben Lehrende aus 10 Fachbereichen in Kooperation mit dem ZfH ihre Lehrveranstaltungen durch Studierende bewerten lassen. Neben den Lehrenden erhalten auch die Dekanate pro Semester die teilweise aggregierten Lehrerevaluationsergebnisse, um auf deren Grundlage Maßnahmen zur Qualitätsentwicklung in der Lehre ableiten zu können.

Bisher können diese Daten nicht zur Abschätzung eines genderbezogenen Entwicklungsbedarfs eingesetzt werden. So wird im Rahmen der fragebogengestützten Lehrveranstaltungsbeurteilung zwar das Geschlecht der Studierenden erhoben, so dass eine geschlechtsbezogene Auswertung der studentischen Bewertungen möglich wäre, Zusammenhänge der Evaluationsergebnisse zum Geschlecht der Lehrenden innerhalb bestimmter Fachbereiche bzw. Fächergruppen können bisher jedoch nicht systematisch untersucht werden.

Zielsetzung des Projekts zur geschlechtsspezifischen Datenauswertung ist es, die Notwendigkeit einer diesbezüglichen Analyse der Lehrerevaluationsdaten abschätzen zu können. Indem beispielhaft für drei Fächergruppen die Lehrerevaluationsergebnisse zweier Semester untersucht werden, sollen Anhaltspunkte dafür gewonnen werden, ob die Einführung einer regelmäßigen und hinsichtlich der geschlechterbezogener Zusammenhänge systematischen Auswertung der Lehrerevaluationsergebnisse vor dem Hintergrund zusätzlichen Auswertungsaufwands lohnend erscheint.

2. Forschungsstand und Fragestellungen

Lehrerevaluation an der UDE verfolgt das Ziel der Lehrentwicklung, indem die studentische Lehrveranstaltungsbeurteilung Informationen zu individuellen Optimierungsmöglichkeiten einzelner Lehrender und deren Lehrmethoden liefert. Ergänzt um Beratungsleistungen und Weiterbildungsangebote durch das ZfH, soll Lehrerevaluation auf diese Weise einen Beitrag zur Qualitätsentwicklung leisten (vgl. Gruber/Pistor 2007).

Wesentliche Voraussetzung dafür ist allerdings, dass die Ergebnisse der Lehrveranstaltungsbeurteilung die Akzeptanz der Lehrenden finden. Obwohl studentische Lehrveranstaltungsbeurteilungen im Hochschulbereich mittlerweile ein verbreitetes Instrument zur

Evaluation von Lehrerfolg darstellen, haben Lehrende häufig Zweifel an deren Wirksamkeit (vgl. Berger/Schleußner 2003). So können auf die subjektive Wahrnehmung der Studierenden bzgl. der Lehrqualität einer Lehrveranstaltung neben der eigentlichen Lehrleistung auch weitere Variablen Einfluss haben, die allerdings in keinem Zusammenhang zum Lehrgeschehen stehen. Dies führt zu Einschränkungen der Validität, denn studentische Lehrveranstaltungsbewertungen sind dann als valide anzusehen, wenn sie die Meinungsbilder der Studierenden unverzerrt erfassen können (vgl. Spiel/Grössler 2000). Um die Akzeptanz der Ergebnisse sicherzustellen ist somit transparent zu machen, wie genau die Ergebnisse den Beurteilungen der Lehrqualität entsprechen bzw. inwiefern sie von verschiedenen Störgrößen verzerrt werden.

Die Evaluationsforschung im Hochschulbereich hat in den letzten Jahren verschiedene Bias-Variablen untersucht. Dies waren vor allem der Besuchsgrund (Pflichtveranstaltungen, freiwillige oder fakultative Veranstaltungen), Schwierigkeit des Lernstoffs, Interesse und Motivation der Studierenden, Hierarchieposition des Lehrenden, Rahmenbedingungen (Lärm, Raumsituation), Termin der Veranstaltung, Sympathie zum Lehrenden, voruniversitäre Bildung (vgl. Berger/Schleußner 2003; Rindermann 2001; Spiel/Grössler 2000; Shevlin et al. 2000; Husbands/Fosh 1993).

Hinsichtlich biographischer Merkmale von Studierenden und Lehrenden können die Studien in der Regel keinen systematischen Zusammenhang mit studentischen Bewertungen feststellen (Spiel/Grössler 2000; Berger/Schleußner 2003; Diehl 2001; Winocur et al. 1989; Feldman 1993). Dies gilt insbesondere auch, wenn das Geschlecht des Lehrenden als mögliche Bias-Variable berücksichtigt wird. So kommt Diehl (2001: 80) in seiner Betrachtung verschiedener Untersuchungen zu dem Ergebnis, dass „(...) die vorgenommenen Vergleiche der durchschnittlichen Beurteilungen von Dozenten und Dozentinnen keine oder nur (praktisch) unbedeutende Unterschiede“ ergeben haben.

Es stellt sich daher die Frage, ob ein Gender-Bias in der studentischen Lehrveranstaltungsbewertung überhaupt berücksichtigt werden muss, d.h. ob das Geschlecht der Lehrenden bei der Bewertung von Lehrveranstaltungen durch die Studierenden von Bedeutung ist.

An US-amerikanischen Hochschulen durchgeführte Studien – insbesondere von Basow – weisen hinsichtlich der Wirkungen des Geschlechts der Lehrenden auf komplexere Zusammenhänge hin. Wird lediglich untersucht, ob männliche und weibliche Lehrende systematisch unterschiedliche Bewertungen von den Studierenden erhalten, so zeigen sich keine bzw. zu vernachlässigende Effekte. Wird allerdings die Interaktion des Geschlechts der Lehrenden mit weiteren Variablen untersucht, so zeigen sich deutlichere Effekte (vgl. Basow/Silberg 1987; Basow 1994; Basow 1995; Basow 2000). So weist Basow darauf hin, dass Forschungsergebnisse zur Auswirkung von Gender auf soziale Wahrnehmung zeigen, dass Verzerrungen auf diesem Gebiet häufig erst in der Interaktion mit anderen Variablen erkennbar werden. Sie kommt daher zu folgendem Ergebnis: „A complex picture emerges when student evaluations are examined as a function not only of professor gender but also of student gender, gender-typing of the discipline, status of the professor, and gender-typed characteristics of the professors.“ (Basow 1995: 656)

Als Drittvariablen berücksichtigt Basow daher in ihren Untersuchungen vor allem „student gender“ und „gender-typing of the discipline“. Werden bei der Untersuchung der Unterschiede in den Bewertungen der Lehrenden durch die Studierenden neben dem Geschlecht der Lehrenden als Drittvariablen das Geschlecht der Studierenden und das Fach

kontrolliert, zeigt sich, dass „Gender matters“ (vgl. Basow 1995). Es lassen sich vor allem zwei Wirkungsweisen erkennen:

- (1) Studentinnen und Studenten unterscheiden sich in der Bewertung von weiblichen Lehrenden, nicht aber in der Bewertung von männlichen Lehrenden.
- (2) Das Geschlecht der Lehrenden hat vor allem dann eine Auswirkung, wenn das Geschlecht auffällig ist, d.h. wenn weibliche Lehrende in einem Fach nur gering vertreten sind.

Wird das *Geschlecht der Studierenden* kontrolliert, so zeigen sich die Bewertungen von männlichen Lehrenden meist unbeeinflusst, während weibliche Lehrende von Studentinnen bessere Bewertungen erhalten als von Studenten (vgl. Basow 1994). Diese Ergebnisse blieben auch dann bestehen, wenn die Hierarchiestufen und Lehrerfahrung der Dozierenden, die von den Studierenden erwarteten Noten, ihre Durchschnittsnoten und ihr Studienjahr kontrolliert wurden (vgl. Basow 1995). Einschränkend weist Basow allerdings darauf hin, dass die gefundenen Effekte – wenn auch statistisch signifikant – doch gering ausgeprägt sind. Sie macht aber auch darauf aufmerksam, dass diese Effekte gerade an US-amerikanischen Universitäten, an denen Personalentscheidungen auch an studentischer Evaluation festgemacht werden, für den einzelnen Lehrenden keine zu vernachlässigende Größe darstellen. „Instructors whose particularities (e.g., gender, gender of students, teaching style, gender-typed personality characteristics, and discipline) all are correlated negatively with student ratings may be particularly affected.“ (Basow 1995: 664)

Diese Ergebnisse lassen vermuten, dass an weibliche Lehrende andere Erwartungen gerichtet werden als an männliche Lehrende (vgl. Basow/Silberg 1987). So zeigt sich, dass weibliche Lehrende, um gute Globalbewertungen zu erhalten, nicht nur auf der instrumentellen Ebene (Organisation der Veranstaltungen, Erklärungsleistung usw.) positiv bewertet sein müssen, sondern auch auf der sozialen-emotionalen Ebene (bspw. hohe Sensibilität für und Unterstützung der Studierenden), während für männliche Lehrende nur die instrumentelle Ebene wichtig ist (vgl. Basow 1995). Diese Bewertungsmuster hängen evtl. damit zusammen, dass weibliche Lehrende im Hochschulbereich insgesamt noch immer unterrepräsentiert sind und daher eine Ausnahme darstellen. Ihr Geschlecht fällt somit auf und daher werden zusätzlich stereotype weibliche Rollenerwartungen an sie gerichtet. Weibliche Lehrende sehen sich so höheren Erwartungen ausgesetzt, da sie nicht nur ihre professionellen Rollen, sondern zugleich auch ihre Geschlechterrollen erfüllen müssen (vgl. Winocur et al. 1989). Basow (1994: 2) kommt daher zu dem Ergebnis, dass „(...) female professors bear a double burden: they must fulfill both their gender role by being nurturant and warm, as well as their professional role by being competent and knowledgeable.“

Zudem weisen Untersuchungen für weibliche Lehrende vor allem auf dem Gebiet von sozial-emotionalem Lehrverhalten höhere Qualitäten nach als für ihre männlichen Kollegen. So zeigen Beobachtungsstudien, dass Hochschullehrerinnen in ihrem Verhalten einen deutlich höheren Respekt vor den Studierenden als Person zeigen und bspw. weniger abwertende Kommentare äußern und häufiger geschlechtsneutrale Sprache verwenden. Außerdem werden weibliche Lehrende von den Studierenden als ihnen gegenüber offener wahrgenommen. Auch zeigen Studien, dass Hochschullehrerinnen die Beteiligung der Studierenden in den Lehrveranstaltungen deutlicher fördern und bspw. häufiger Fragen stellen und weniger Vortragsverhalten zeigen. Darüber hinaus gehen Dozentinnen persönlicher auf die Studierenden ein, kennen Studierende mit Namen und geben zu de-

ren Beiträgen häufiger klare Rückmeldungen (vgl. Viebhahn 2007). Da Studentinnen solche Qualitäten und Verhaltensweisen deutlich höher bewerten als ihre männlichen Kommilitonen, kann sich hier in der Evaluation eine „same-gender preference in teaching style“ äußern. Studenten bevorzugen dagegen häufig einen eher instrumentellen gegenüber einem emotionalen Lehrstil (vgl. Basow 1994; Winocur et al. 1989; Fassinger 1995).

Diese Zusammenhänge zeigen sich noch ausgeprägter, wenn es sich um „gender-atypical-fields“ handelt. Sind Frauen in einem *Fachgebiet* unterrepräsentiert und damit als Lehrende in ihrem Geschlecht auffällig, so werden deutlicher geschlechtertypische Rollenerwartungen an sie gestellt, das Geschlecht der Lehrenden beeinflusst die Bewertung der Studierenden entsprechend deutlicher und auch die „same-gender preference“ ist ausgeprägter (vgl. Basow/Silberg 1987). „In addition, research has demonstrated that women receive the most negative ratings when they are in a nontraditional field, especially from those who hold traditional attitudes toward women’s roles, primarily men.“ (Basow 1995: 664). So stellt Basow diese Effekte besonders in den Sozialwissenschaften und weniger in den Geisteswissenschaften fest. In den Naturwissenschaften konnten keine konsistenten Muster aufgezeigt werden. Diese Ergebnisse lassen sich über Gender-Erwartungen erklären: In den Geisteswissenschaften sind weibliche Lehrende meist stark vertreten und daher nicht auffällig. In den Sozialwissenschaften sind Frauen dagegen noch unterrepräsentiert. Frauen erhalten also häufig dann schlechtere Bewertungen, wenn sie in einem für ihr Geschlecht untypischen Fach lehren.

Aufbauend auf die Ergebnisse dieser Analysen von Basow sollen im Rahmen des Projektes folgende Fragestellungen untersucht werden:

- 1) Sind eventuell bestehende Unterschiede zwischen der Veranstaltungsbewertung von weiblichen und männlichen Lehrenden abhängig vom Geschlecht der Studierenden?
- 2) Sind eventuell bestehende Unterschiede in der Veranstaltungsbewertung von weiblichen und männlichen Lehrenden abhängig von der Fachrichtung?

3. Methode und Datenbasis

Zur Untersuchung oben angeführter Fragestellungen wurden Ergebnisse der studentischen Lehrveranstaltungsbewertung herangezogen, die seit dem WS 2004/2005 sukzessive an der UDE eingeführt wurde und an der seit dem WS 2007/2008 alle Fachbereich – mit Ausnahme der Ingenieurwissenschaften und der Medizin – teilnehmen. Da innerhalb der Evaluationsbögen neben Fragen zur Einschätzung der Lehrveranstaltungen auch das Geschlecht der Studierenden erhoben wird, eignen sich die Daten unmittelbar zu einer Analyse hinsichtlich des Geschlechts der Studierenden. Das Geschlecht der Lehrenden wird allerdings nicht erhoben, so dass vor der Analyse in einem ersten Arbeitsschritt eine entsprechende Datenaufbereitung durchgeführt wurde. Dazu wurden die vorliegenden Datensätze so aufbereitet, dass sie eine Auswertung sowohl

- a) nach dem Geschlecht der Studierenden als auch
- b) nach dem Geschlecht der Lehrenden und
- c) dem Fachgebiet

ermöglichen können.

Insofern handelt es sich bei der geschlechterspezifischen Auswertung der Lehrevaluationsergebnisse um eine Sekundäranalyse. Bei der Interpretation der Untersuchungsergebnisse ist daher zu beachten, dass die auswertbare Datenbasis nicht hinsichtlich des Status der Lehrenden und deren Lehrerfahrung, aber auch nicht nach der Art und Größe der Veranstaltung (Pflichtveranstaltungen, Veranstaltungen im Hauptstudium, Veranstaltungszeitpunkt usw.) kontrolliert werden konnte, die ebenfalls Einfluss auf das Bewertungsergebnisse durch die Studierenden haben können. Auch können sich innerhalb des Datensatzes Bewertungen unterschiedlicher Veranstaltungen des gleichen Dozierenden befinden und Studierende können – allerdings jeweils für verschiedene Veranstaltungen – mehrfach an Lehrevaluationen innerhalb eines Semesters teilgenommen haben. Zudem können Studierende nur eingeschränkt nach ihrem Studienfortschritt unterschieden werden und es wird auch nicht differenziert, ob sie Haupt- oder Nebenfachstudierende in einer Lehrveranstaltung sind.

Für die Zusammensetzung der Datenbasis wurden die Evaluationsergebnisse aus den zwei Semestern SoSe 2007 und WS 2007/2008 ausgewählt. Die Evaluationsordnung sieht vor, dass alle Lehrenden mit zwei Veranstaltungen pro Jahr an der Lehrevaluation teilnehmen. Da Fachbereiche diesbezüglich unterschiedliche Evaluationspraktiken entwickelt haben, ist durch die Berücksichtigung zweier Semester somit die Wahrscheinlichkeit gegeben, Evaluationsergebnisse jeder/jedes Lehrenden in der Auswertung berücksichtigen zu können.

Da die oben angeführten Forschungsergebnisse nahe legen, hinsichtlich geschlechtsspezifischer Auswertung der Evaluationsergebnisse insbesondere die Interaktion zwischen Studierenden und Lehrenden zu betrachten, wurden nur Evaluationsergebnisse zur Veranstaltungsform „Seminar“ herangezogen. Zudem wird eine Unterscheidung der Ergebnisse in die Fächergruppen Natur-, Geistes- und Gesellschaftswissenschaften vorgenommen. Um die Vergleichbarkeit der Ergebnisse zu gewährleisten, wurden jeweils Fächer aus diesen drei Bereichen ausgewählt, die den gleichen Fragebogen für Seminare verwenden. Daher wurden Evaluationsergebnisse der Lehrheiten Physik, Germanistik und Gesellschaftswissenschaften herangezogen.¹

¹ Die Gesellschaftswissenschaften umfassen die Fächer Politikwissenschaft und Soziologie, die im Rahmen von integrierten Diplomstudiengängen und auch im ersten Studienjahr der Bachelorprogramme gemeinsame Lehrangebote machen.

Fach	Studierende		Lehrende	
Germanistik 50 Seminare	4764 Studierende *	3556 Frauen (75%) 1208 Männer (25%)	47 Lehrende **	20 Frauen (43%) 27 Männer (57%)
	1733 Fälle	1353 Frauen (78%) 380 Männer (22%)	50 Fälle	31 Frauen (62%) 19 Männer (38%)
GesWi 103 Seminare	3661 Studierende *	1859 Frauen (51%) 1802 Männer (49%)	82 Lehrende **	38 Frauen (46%) 44 Männer (54%)
	1420 Fälle	720 Frauen (51%) 700 Männer (49%)	103 Fälle	26 Frauen (25%) 77 Männer (75%)
Physik 14 Seminare	658 Studierende *	174 Frauen (26%) 484 Männer (74%)	84 Lehrende **	11 Frauen (13%) 73 Männer (87%)
	206 Fälle	87 Frauen (42%) 119 Männer (58%)	14 Fälle	1 Frau (7%) 13 Männer (93%)

* Studierende im WS 07/08 nach Fällen, Stand: 11.03.2008

** Personal nach Köpfen und Landeshaushalt, Stand: 29.02.2008

Im Bereich der Geisteswissenschaften konnten in dem betrachteten Zeitraum 50 Seminare aus dem Bereich Germanistik in die Auswertung einbezogen werden. Dies entspricht 1733 Bewertungen durch Studierende, wobei 78% (1353) der Bewertungen durch Studentinnen und 22% (380) durch Studenten vorgenommen wurden, was die geschlechterbezogene Zusammensetzung der Studierenden in der Germanistik widerspiegelt. Die Evaluationsergebnisse beziehen sich auf 50 Lehrende, d.h. zu 62% (31) auf weibliche Lehrende und nur zu 38% (19) auf Dozenten der Germanistik. Hierin zeigt sich, dass Dozentinnen in der Germanistik im betrachteten Zeitraum häufiger ihre Lehrveranstaltungen evaluiert haben als ihre Kollegen.

In den Gesellschaftswissenschaften wurden 103 Seminare ausgewertet, wobei 1420 Fälle herangezogen werden konnten. 720 Bewertungen stammen dabei von Studentinnen (51%) und 700 von Studenten (49%), was die Geschlechterverteilung unter den Studierenden der Gesellschaftswissenschaften abbildet. Die Lehrveranstaltungsbewertungen beziehen sich auf 103 Lehrende, worunter sich 25% (26) Frauen und 75% (77) Männer befinden. Im Unterschied zur Germanistik nehmen in den Gesellschaftswissenschaften die Dozenten häufiger an Lehrevaluationen teil.

In der Physik konnten für den betrachteten Zeitraum 14 Seminare ermittelt werden. Da sich unter den evaluierten Lehrveranstaltungen nur lediglich ein Seminar befand, dass von einer Frau gehalten wurde, erscheint eine geschlechtsspezifische Auswertung der Evaluationsergebnisse in der Physik nicht sinnvoll. Der Vergleich der studentischen Beurteilungen lediglich einer weiblichen Lehrenden mit den Bewertungen ihrer männlichen Kollegen lässt keine aussagefähigen Ergebnisse zu. Im Folgenden wird daher auf die Auswertung der Ergebnisse einer Naturwissenschaft verzichtet, auch wenn vor dem Hintergrund des

oben angeführten Forschungsstandes in diesem Bereich vor allem interessante Ergebnisse zu erwarten gewesen wären.²

Die studentischen Lehrveranstaltungsbewertungen werden an der UDE durch vom ZfH entwickelte Fragebögen erhoben, die für verschiedene Veranstaltungstypen (Vorlesung, Seminar, Übungen) als Feedback-Instrumente für die Lehrenden konzipiert sind. Sie umfassen daher Items zur Konzeption und Umsetzung der Lehrveranstaltung (Lernziele, Lehrmethoden, Nachvollziehbarkeit der Erklärungsschritte usw.), zum Medieneinsatz und Arbeitsklima (Möglichkeit zur Diskussion und Fragestellung usw.), aber auch zu den Rahmenbedingungen der Veranstaltung (Betreuung, Raumsituation). Darüber hinaus werden Einschätzungen der Studierenden zu Arbeitsaufwand, Umfang der Lehrinhalte und dem Schwierigkeitsgrad der Veranstaltung erhoben. Abschließend werden die Studierenden um eine Gesamtbewertung hinsichtlich der Lerneffekte und der Lehrveranstaltung insgesamt gebeten. Mit den Ergebnissen sollen die Lehrenden Ansatzpunkte erhalten, die Qualität ihrer Lehre in der Diskussion mit den Studierenden verbessern zu können. Mit dem Fragebogen soll ausdrücklich keine Qualitätsmessung betrieben, sondern ein Beitrag zur Qualitätsentwicklung geleistet werden. Aus diesem Grund stehen „Veranstaltungs-Items“ zur Gestaltung der Lehrveranstaltung im Vordergrund, während „Lehrenden-Items“, die auf eine unmittelbare Einschätzung von Leistungen und Merkmalen des Lehrenden zielen, nur eine untergeordnete Rolle spielen. Die Evaluationsergebnisse erlauben somit nur einen indirekten Schluss auf die Lehrleistung der Dozierenden. Das Skalenniveau der verwendeten Items kann als ordinal interpretiert werden, da die Einstellungen der Studierenden mittels einer 5-Punkte-Skala erhoben werden.

Um bei aller Einschränkung aufgrund der verschiedenen Evaluationsansätze zumindest eine gewisse Vergleichbarkeit der Ergebnisse mit dem oben angeführten US-amerikanischen Forschungsstand zu ermöglichen, werden aus dem Fragebogen der UDE diejenigen Dimensionen für die geschlechtsspezifische Analyse ausgewählt, die sich auf 2 der sog. „Five Teaching Factors“ beziehen. Es handelt sich hierbei um verschiedene Dimensionen der Lehre, die von Leventhal, Perry und Abrami in „Scholarship“, „Organization/Clarity“, „Instructor-Group Interaction“, „Instructor-Individual Student Interaction“, „Dynamism/Enthusiasm“ unterschieden werden. Ergänzt werden diese lehrspezifischen Items um ein „overall teaching ability item“ (vgl. Leventhal/Perry/Abrami 1977). Auch wenn der Fragebogen der UDE als Feedback-Instrument nicht explizit auf das von Leventhal, Perry und Abrami konzipierte Erhebungsinstrument bezogen ist, so werden doch für die beiden Dimensionen „Organisation“ und „Interaktion“ vergleichbare Items erhoben (vgl. Anhang).³

Die Evaluationsergebnisse wurden daher nach den folgenden Dimensionen untersucht und die dazugehörigen Items herangezogen.

² Eine andere Naturwissenschaft konnte für die geschlechtsspezifische Auswertung aus dem Grund nicht herangezogen werden, dass in den übrigen Naturwissenschaften andere Fragebögen für die Evaluation von Lehrveranstaltungen zum Einsatz kommen und die Vergleichbarkeit der Ergebnisse so nicht gegeben ist.

³ Diese Annahme wird auch durch Faktorenanalysen bestätigt. Werden die Variablen der Faktoren „Organisation“ und „Interaktion“ entsprechenden Analysen unterzogen, so zeigen sich recht hohe Korrelationen zwischen den Variablen, so dass der Schluss auf jeweils eine gemeinsame Hintergrundvariable durchaus berechtigt ist.

	Indikatoren	Items	Fragen
1	Globalbewertungen	1.1 Lerneffekt der LV	„Den Lerneffekt dieser LV empfinde ich bisher als ... sehr klein – sehr groß“
		1.2 Gesamtbeurteilung der LV	„Insgesamt betrachtet beurteile ich diese LV als ... sehr schlecht – sehr gut“
2	Organisation	2.1 Lernziele	„Die Lernziele der LV wurden von Anfang an klar gemacht“
		2.2 Relevanz der Lehrinhalte	„Die Relevanz der Lehrinhalte für das Studienziel wurde klar verdeutlicht“
		2.3 Vorbereitung des Lehrenden	„Die/der Lehrende wirkt auf mich gut vorbereitet“
		2.4 Nachvollziehbarkeit der Erklärungsschritte	„Die Erklärungsschritte der/des Lehrenden kann ich inhaltlich gut nachvollziehen“
		2.5 Fragenbeantwortung	„Fragen werden von der/dem Lehrenden für mich ausreichend klar beantwortet“
		2.6 Transparenz von Leistungs- und Prüfungsanforderungen	„Leistungs- und Prüfungsanforderungen werden für mich transparent dargestellt“
3	Interaktion	3.1 Fragemöglichkeiten	„Es bestehen genügend Möglichkeiten, Fragen zu stellen“
		3.2 Diskussionsmöglichkeiten	„Es bestehen genügend Gelegenheiten, sich an Diskussionen zu beteiligen“
		3.3 Aufgreifen studentischer Beiträge	„Die/der Lehrende greift studentische Beiträge für mich angemessen auf“
		3.5 Ruhiger Ablauf	„Die/der Lehrende sorgt für einen ruhigen Ablauf der LV“
		3.4 Arbeitsatmosphäre	„Die Arbeitsatmosphäre in dieser LV empfinde ich als angenehm“
		3.5 Interesse am Studienerfolg	„Die/der Lehrende zeigt Interesse am Lernerfolg der Studierenden“

Da die oben angeführten Forschungsergebnisse zeigen, dass sich bei Globalbewertungen weniger Gender-Effekte zeigen lassen als bei differenzierteren Items, und daher den Einbezug möglichst vieler Bewertungsitems empfehlen, werden neben den zwei Globalbewertungen jeweils 6 Items für die Dimensionen „Organisation“ und „Interaktion“ getrennt nach den drei Fächern untersucht.⁴ Zur bivariaten Analyse werden aufgrund des nominalen Skalenniveaus der Variable Geschlecht Kreuztabellen mit einem Chi-Quadrat-Test durchgeführt.

⁴ Aufgrund der unterschiedlichen Fallzahlen in den Fächern erscheinen separate Auswertungen sinnvoll.

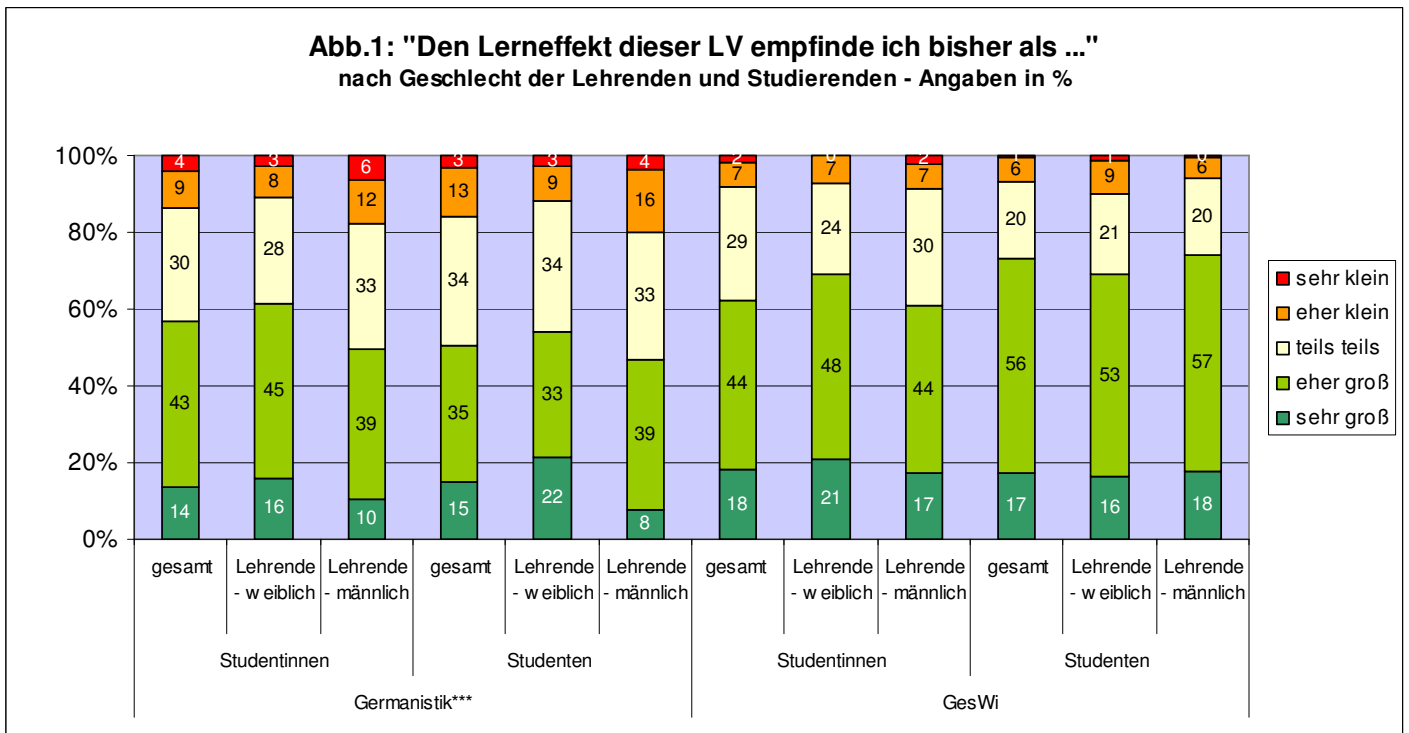
4. Ergebnisse

Die Ergebnisse der geschlechtsspezifischen Auswertung studentischer Lehrveranstaltungsbewertung werden im Folgenden nach den oben angeführten Itemgruppen dargestellt. Auf die Ergebnisse zu den Globalbewertungen für die jeweiligen Lehrveranstaltungen folgen die Ergebnisse zur Itemgruppe „Organisation“ und „Interaktion“.

4.1 Ergebnisse zu Globalbewertungen

Befragt nach ihrer **Einschätzung zu den Lerneffekten**, die sie insgesamt aus einer Lehrveranstaltung ziehen können, sprechen die Hälfte der Studenten aus der *Germanistik* von großen bis sehr großen Effekten, während die Studentinnen mit knapp 60% etwas bessere Bewertungen abgeben. Auffällig ist bei diesem Urteil, dass sowohl Studentinnen als auch Studenten die Lerneffekte in Veranstaltungen weiblicher Lehrender besser einschätzen, als die der männlichen Kollegen. Studentinnen beurteilen ihre Dozentinnen diesbezüglich allerdings wesentlich besser als ihre männlichen Kommilitonen (vgl. Abb. 1).

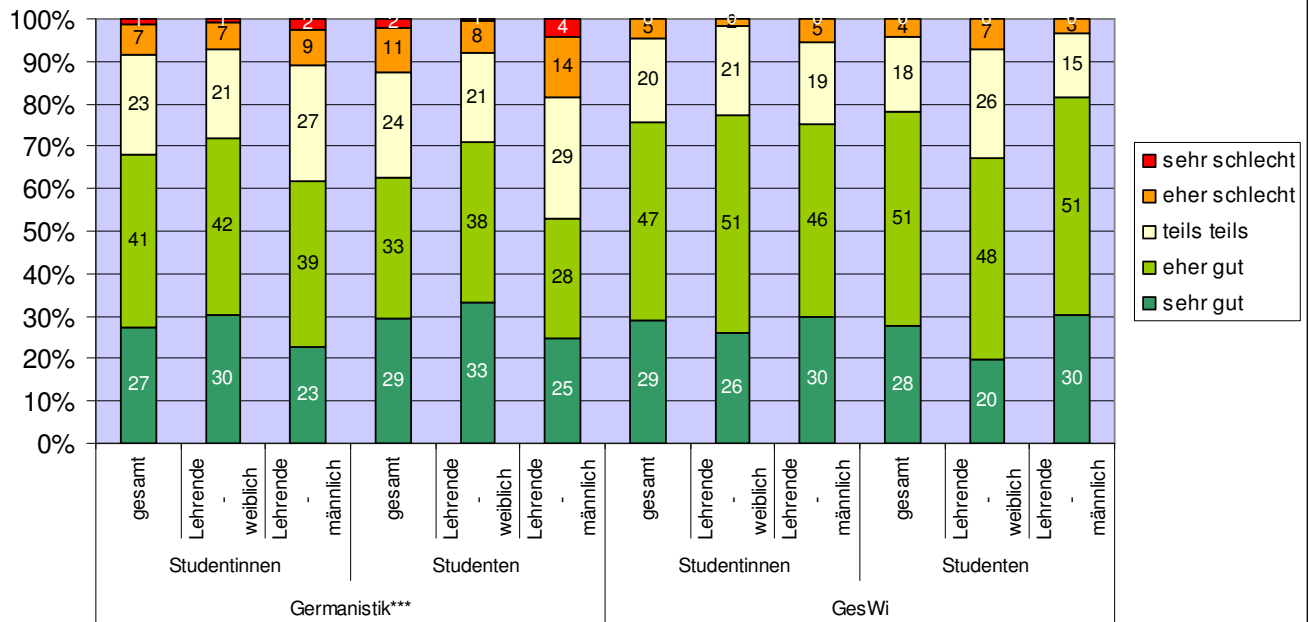
Abb.1: "Den Lerneffekt dieser LV empfinde ich bisher als ..."
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



Ein anderes Bild zeigt sich in den *Gesellschaftswissenschaften*. Hier beurteilen die Studenten (73%) den Lerneffekt von Veranstaltungen durchschnittlich besser als die Studentinnen, von denen 62% von großen bis sehr großen Lerneffekten sprechen. Während Dozentinnen nun aber von ihren Studentinnen deutlich bessere Bewertungen erhalten – 69% geben große bis sehr große Lerneffekt bei weiblichen Lehrenden an – schätzen die Studenten die Lerneffekte in Veranstaltungen männlicher Lehrender besser ein. Hier wird der Lerneffekt zu 75% als groß bis sehr groß bezeichnet (vgl. Abb.1).

Die **Gesamtbeurteilung der Lehrveranstaltungen** spiegelt in der *Germanistik* – allerdings in einer weniger deutlichen Weise – die Beurteilung der Lerneffekte wider. Auch hier geben die Studentinnen ein etwas besseres Urteil ab als die Studenten (68% bzw. 62%) und die Veranstaltungen der Dozentinnen werden sowohl von den Studentinnen als auch von den Studenten besser beurteilt als die der männlichen Lehrenden (vgl. Abb.2).

Abb.2: "Insgesamt betrachtet beurteile ich die LV als ..."
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



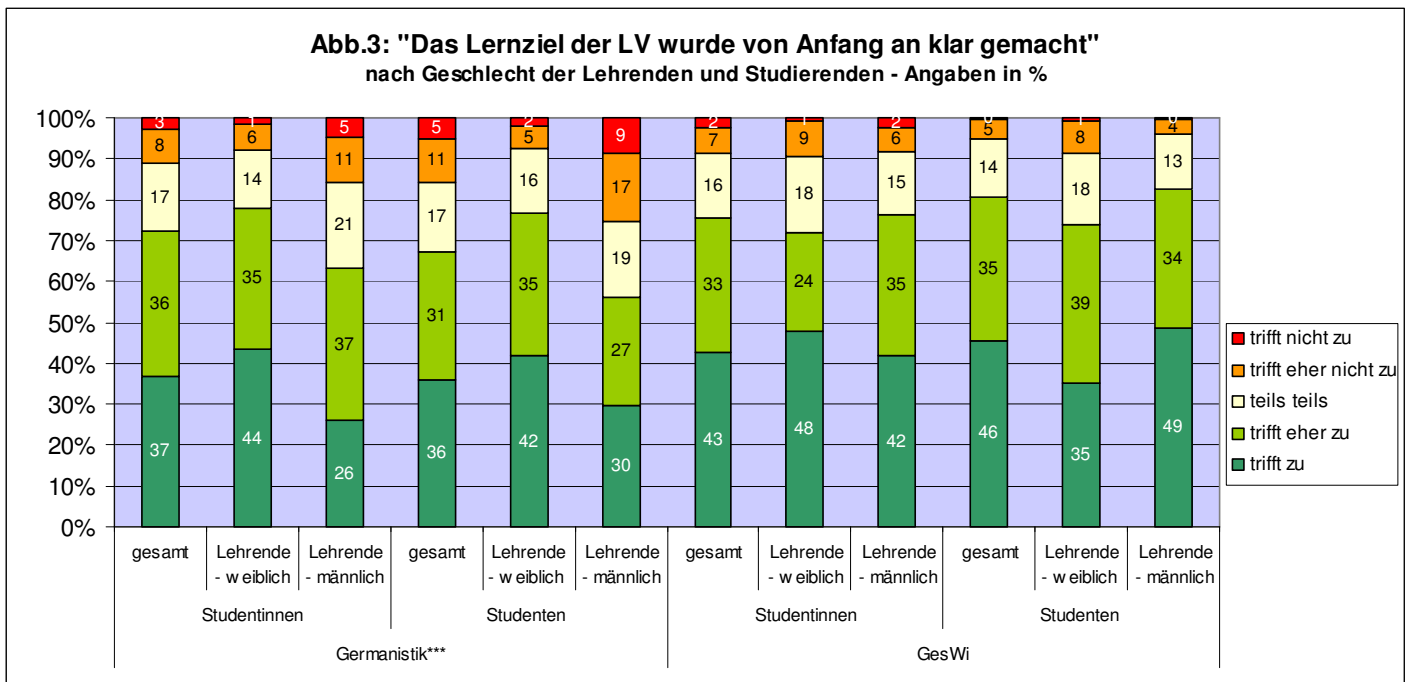
Auch in den *Gesellschaftswissenschaften* lassen sich bei der Gesamtbeurteilung den Einschätzungen zu Lerneffekten vergleichbare Bewertungen erkennen. Die Studenten geben insgesamt etwas bessere Einschätzungen als ihre Kommilitoninnen (79% gut bis sehr gut) und beurteilen die Veranstaltungen der Dozenten auch deutlich besser als die der Dozentinnen. 81% der Studenten geben das Urteil gut bis sehr gut für ihre männlichen Lehrenden, bei den Veranstaltungen weiblicher Lehrenden sind es 68% (vgl. Abb.2).

Für die Globalbewertungen zu Lehrveranstaltungen zeigen sich im Vergleich beider Fächer somit Unterschiede in der Evaluation durch die Studierenden. So liefert die geschlechtsspezifische Auswertung im Fach Germanistik hinsichtlich der Globalbewertungen signifikante Ergebnisse. Studentinnen und Studenten der Germanistik stimmen in ihrem Urteil insofern überein, als sie Lehrveranstaltungen von Dozentinnen alles in allem etwas besser bewerten. Auch wenn die Ergebnisse in den Gesellschaftswissenschaften keine signifikanten Ergebnisse liefern können, so zeigen sich doch gewisse Tendenzen. Die Studierenden der Gesellschaftswissenschaften lassen in ihrem Evaluationsverhalten tendenziell eine „Same-Gender-Preference“ erkennen. So bewerten Studentinnen die Veranstaltungen ihrer Dozentinnen im Vergleich zu den Dozenten etwas besser, während die Studenten ihre Dozenten im Vergleich zu den Dozentinnen etwas besser einschätzen.

4.2 Ergebnisse zum Lehrfaktor „Organisation“

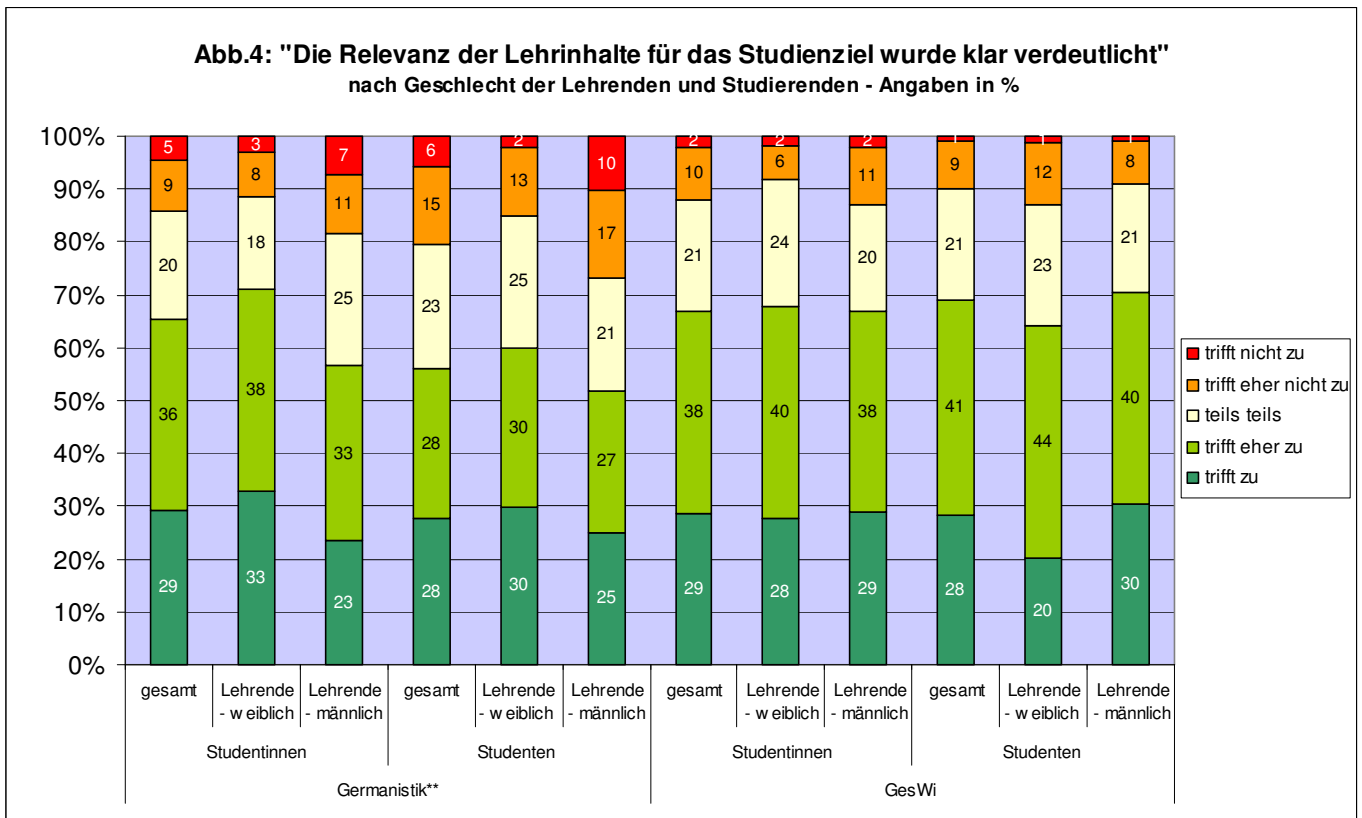
Unter dem Lehrfaktor „Organisation“ wird eine geschlechtsspezifische Analyse der Einschätzung der Studierenden bezüglich der Klarheit von Lernzielen, der Relevanz der Lehrinhalte und der Vorbereitung der Lehrenden durchgeführt. Weiterhin wird dazu die Nachvollziehbarkeit der Erklärungsschritte und die Beantwortung von Fragen durch die/den Lehrenden, aber auch die Transparenz von Leistungs- und Prüfungsanforderungen gezählt.

Hinsichtlich der Klarheit ihrer **Lernziele** unterscheiden sich die Studierenden nicht in ihrer Beurteilung der Lehrenden in der *Germanistik*. So sind sowohl den Studentinnen als auch von den Studenten in Lehrveranstaltungen weiblichen Lehrpersonals die Lernziele klarer als bei den männlichen Kollegen. Mehr als $\frac{3}{4}$ der Studentinnen und Studenten beurteilen die Lehrveranstaltungen ihrer Dozentinnen in dieser Hinsicht gut bis sehr gut. Bei den männlichen Lehrenden sind dies 63% der Studentinnen und 57% der Studenten (vgl. Abb.3).



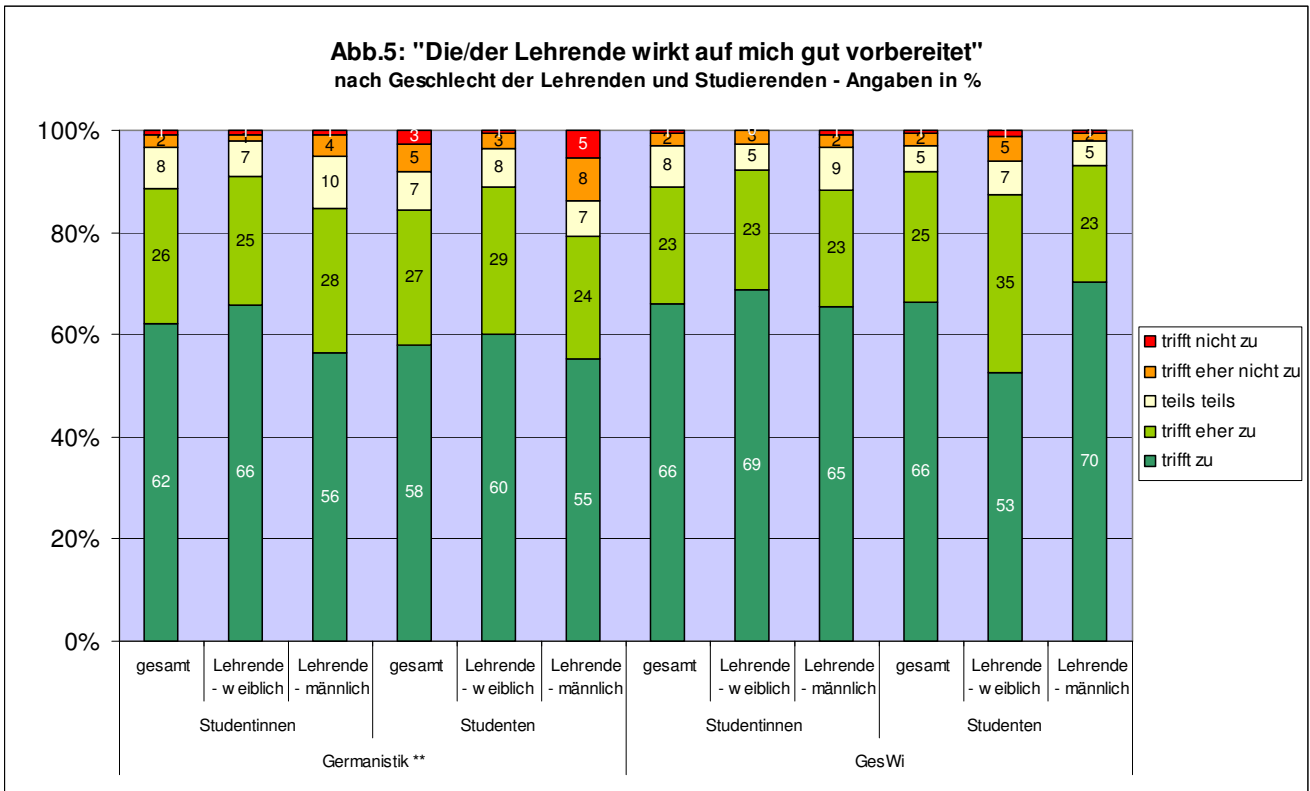
In den *Gesellschaftswissenschaften* wird dagegen das männliche Lehrpersonal hinsichtlich der Klarheit der Lernziele ihrer Lehrveranstaltungen von den Studierenden besser beurteilt. Knapp die Hälfte der Studenten, die Lehrveranstaltungen männlicher Lehrender evaluiert haben, bezeichnet die Klarheit der Lernziele als sehr hoch. Die Studentinnen beurteilen ihre Dozenten diesbezüglich insgesamt zwar auch besser – 77% der Studentinnen bezeichnen die Lernziele als klar bis sehr klar – auffällig ist aber, dass knapp die Hälfte der Studentinnen in Lehrveranstaltungen weiblicher Lehrender der Aussage, die Lernziele seien von Anfang an klar gemacht worden, voll zustimmen können (vgl. Abb.3).

Für das Fach *Germanistik* zeigen sich hinsichtlich der **Relevanz von Lehrinhalten** für das Studienziel ähnliche Ergebnisse. Auch hier werden die Veranstaltungen der Dozentinnen von allen Studierenden im Vergleich zu ihren männlichen Kollegen besser bewertet. Allerdings fällt das Urteil der Studentinnen insgesamt deutlicher aus. 71% der Teilnehmerinnen bescheinigen den Lehrveranstaltungen ihrer Dozentinnen, dass die Relevanz der Lehrinhalte für das Studienziel deutlich bis sehr deutlich wurde, bei den Studenten sind dies 60% (vgl. Abb.4).



In den *Gesellschaftswissenschaften* unterscheidet sich das Urteil der Studentinnen zur Relevanz der Lehrinhalte nicht zwischen Lehrveranstaltungen weiblicher und männlicher Lehrender. 68% bzw. 67% bezeichnen die Verdeutlichung der Relevanz in den Veranstaltungen als klar bis sehr klar. Bei den Studenten zeigen sich dagegen Unterschiede in der Beurteilung von Dozentinnen und Dozenten. Während 70% der Teilnehmer von Veranstaltungen männlicher Lehrender die Relevanz als klar bis sehr klar bezeichnen, sind es bei den weiblichen Lehrenden 64% (vgl. Abb.4).

Die Lehrenden wirken auf die Studierenden beider Fächer gleichermaßen sehr gut vorbereitet. Mehr als 80% der Studierenden sprechen von einer guten bis sehr guten **Vorbereitung**. In der *Germanistik* werden die Dozentinnen von Studentinnen und Studenten wiederum als besser vorbereitet erfahren. 91% der Teilnehmerinnen und 89% der Teilnehmer von Lehrveranstaltungen weiblicher Lehrenden schätzen die Vorbereitung ihrer Lehrenden gut bis sehr gut ein (vgl. Abb.5).

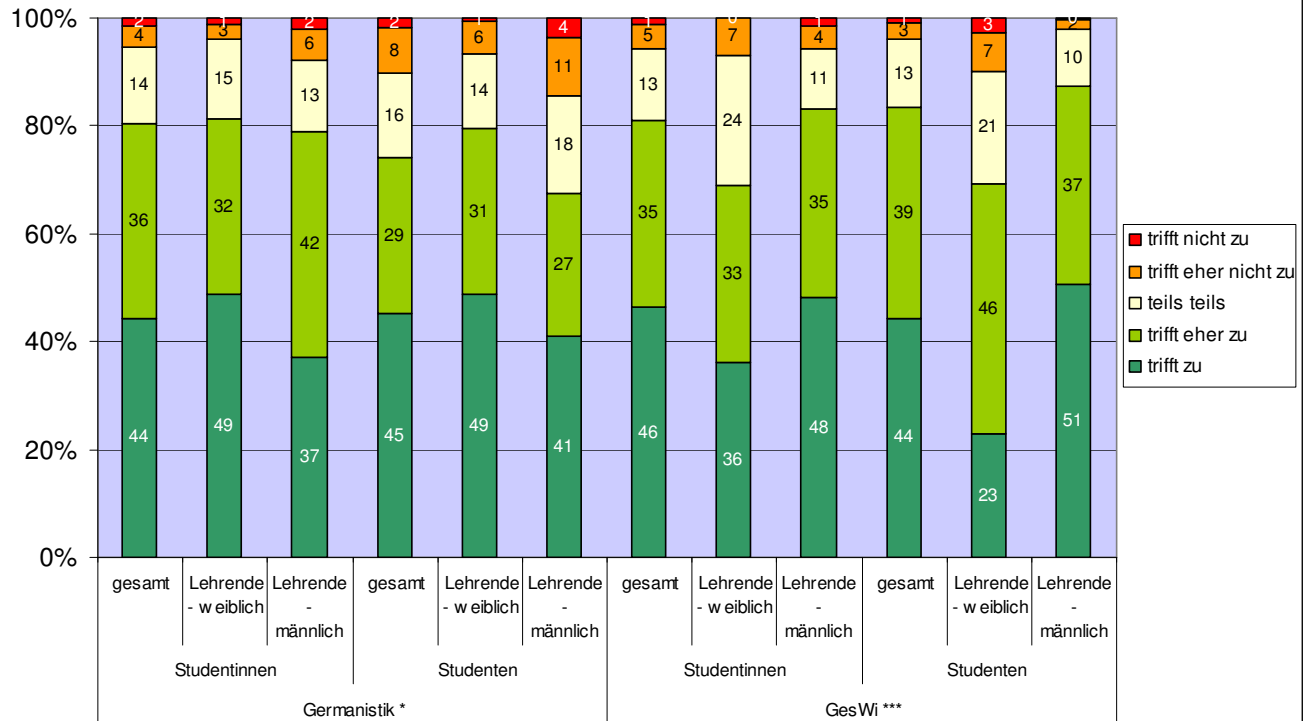


Die Studentinnen und Studenten der *Gesellschaftswissenschaften* sind sich in der insgesamt sehr positiven Beurteilung der Veranstaltungen ebenfalls einig. Die Studentinnen schätzen ihre Dozentinnen allerdings besser ein. 92% der Besucherinnen von Lehrveranstaltungen bezeichnen das weibliche Lehrpersonal als gut vorbereitet, davon sogar 69% als sehr gut vorbereitet. Die Studenten sehen dagegen ihre Dozenten als besser vorbereitet an. 93% der Teilnehmer dieser Veranstaltungen sprechen von einer guten Vorbereitung, darunter sind 70%, die ihren Dozenten sogar bescheinigen, dass sie sehr gut vorbereitet wirkten (vgl. Abb. 5).

Auch in der **Nachvollziehbarkeit der Erklärungsschritte** werden die Dozentinnen in der *Germanistik* von ihren Studentinnen und Studenten besser bewertet als die männlichen Lehrenden. Knapp die Hälfte der Teilnehmerinnen und Teilnehmer der Lehrveranstaltungen weiblicher Lehrender können die Erklärungsschritte ihrer Dozentinnen gut nachvollziehen, für weitere rund 30% trifft dies noch häufig zu. Insbesondere die Studenten beurteilen die männlichen Lehrenden diesbezüglich schlechter. Mit 68% ist die Zustimmung hier unterdurchschnittlich (vgl. Abb.6).

In den *Gesellschaftswissenschaften* wird sowohl von den Studentinnen als auch von den Studenten die Nachvollziehbarkeit der Erklärungsschritte männlicher Lehrender deutlich besser bewertet. Allerdings schätzen die Studentinnen ihre Dozentinnen nicht ganz so schlecht ein, wie ihre männlichen Kommilitonen. Noch 36% der Studentinnen können die Erklärungsschritte weiblicher Lehrender sehr gut nachvollziehen, unter den Studenten sind dies nur 23% (vgl. Abb.6).

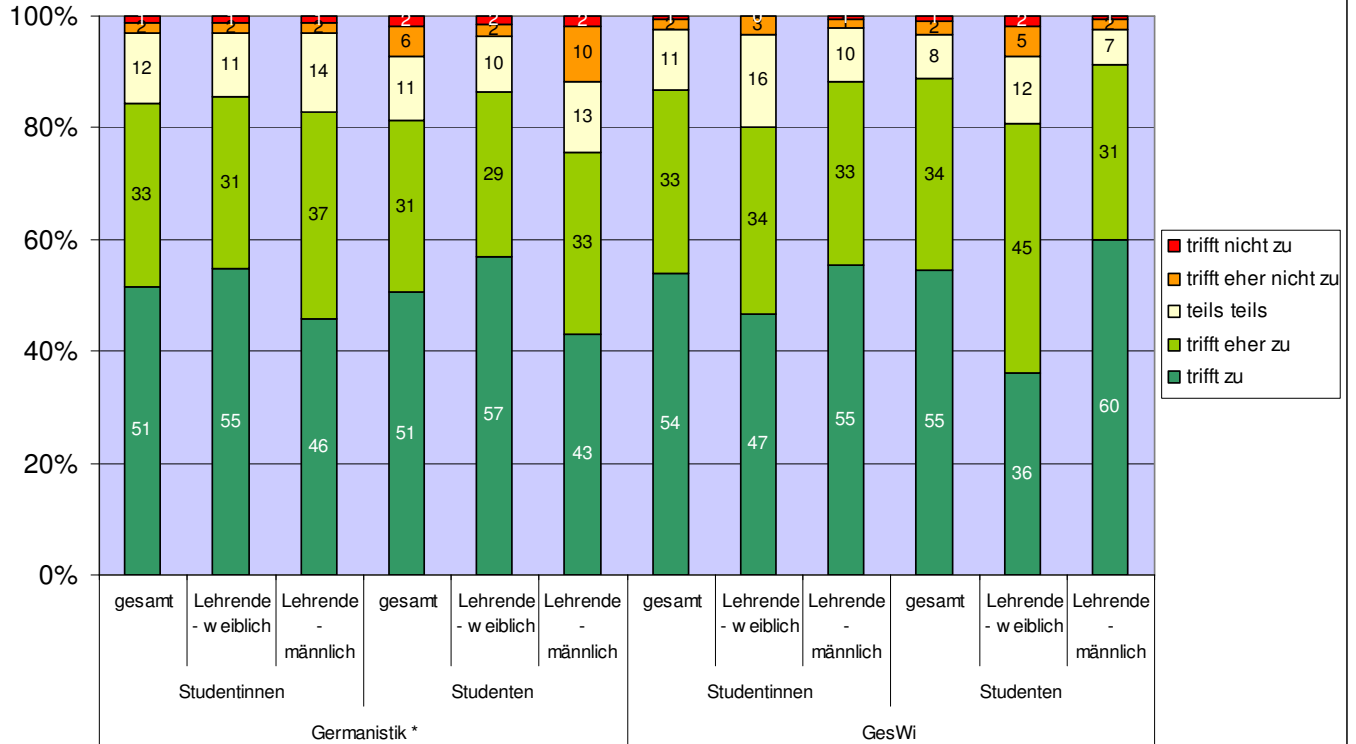
Abb. 6: "Die Erklärungsschritte der/des Lehrenden kann ich inhaltlich gut nachvollziehen" nach Geschlecht der Lehrenden und Studierenden - Angaben in %



In ihre Antwort auf die Frage, inwiefern Lehrende **Fragen ausreichend klar beantworten**, unterscheiden sich die Studentinnen und Studenten der *Germanistik* wiederum kaum. Jeweils mehr als 80% sind diesbezüglich zufrieden bis sehr zufrieden, und zwar mit den weiblichen Lehrenden etwas mehr als mit deren männlichen Kollegen. Die Studenten geben dabei ein etwas deutlicheres Urteil – insbesondere hinsichtlich der Dozenten – ab. 12% wollen nicht von einer ausreichend klaren Fragenbeantwortung durch die männlichen Lehrenden sprechen und noch 13% sind nur teilweise zufrieden (vgl. Abb.7).

Auch in den *Gesellschaftswissenschaften* stimmen die Studierenden in ihrer Bewertung überein und sind mit der Leistung ihrer Dozentinnen und Dozenten diesbezüglich sehr zufrieden. Allerdings werden in diesem Fach die männlichen Lehrenden sowohl von den Studentinnen als auch von den Studenten besser beurteilt als die Weiblichen. Auffallend ist, dass die Studentinnen zwar ebenfalls ihre Dozenten besser beurteilen, allerdings sind noch knapp die Hälfte der Studentinnen (47%), die Veranstaltung von Dozentinnen besucht haben, mit deren Fragenbeantwortung sehr zufrieden. Unter den Studenten sind es nur 36% (vgl. Abb. 7).

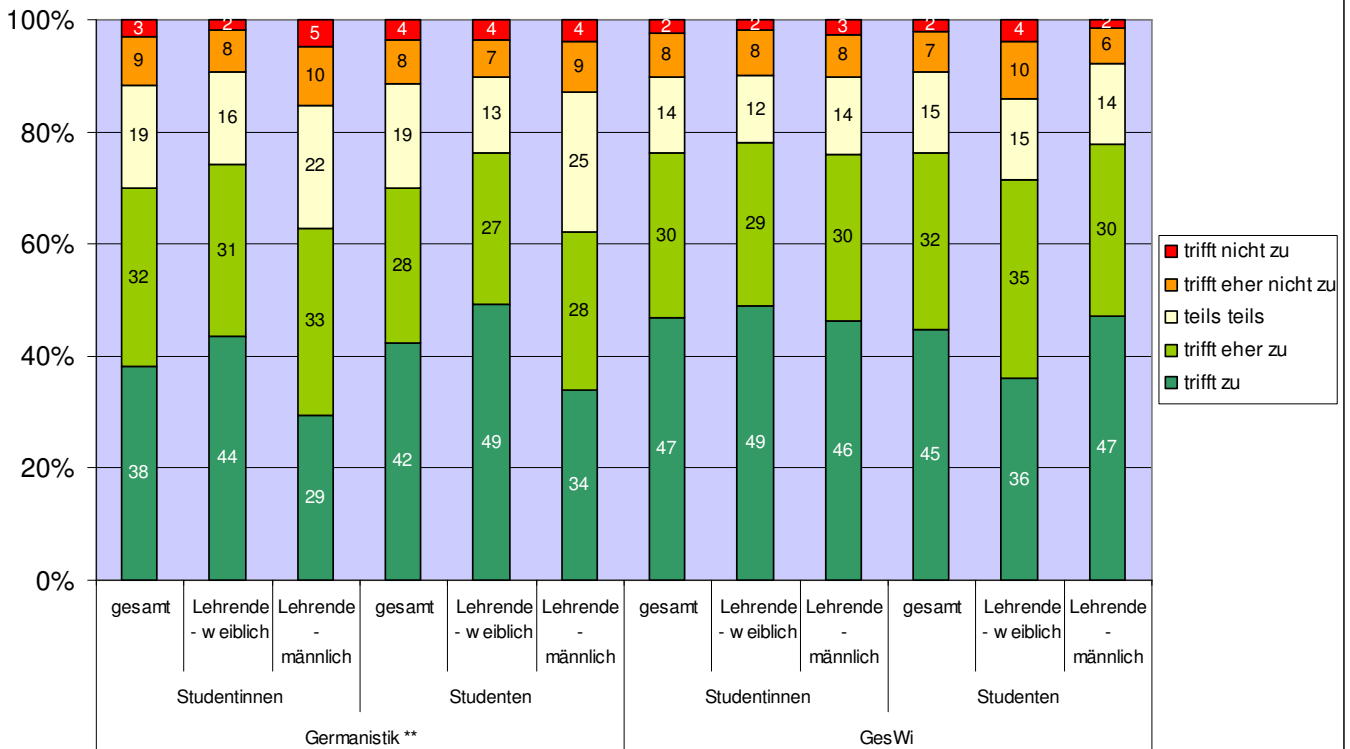
Abb. 7: "Fragen werden von der/dem Lehrenden für mich ausreichend klar beantwortet" nach Geschlecht der Lehrenden und Studierenden - Angaben in %



Die **Transparenz der Leistungs- und Prüfungsanforderungen** beurteilen die Studierenden der *Germanistik* wieder entsprechend des Beurteilungsmusters der Items im Bereich „Organisation“. So wird auch hier von Studentinnen und Studenten eine höhere Zufriedenheit in den Veranstaltungen weiblicher Lehrender angegeben. Insbesondere die Studenten beurteilen ihre Dozentinnen im Vergleich zu den Studentinnen diesbezüglich positiv. 49% sprechen von stets transparenter Darstellung in den Lehrveranstaltungen. Dagegen sind es nur 29% der Studentinnen, die Veranstaltungen von Dozenten besucht haben und eine hohe Transparenz angeben. Unter den Studenten sind es immerhin 34% (vgl. Abb. 8).

Die Ergebnisse der Lehrevaluation in den *Gesellschaftswissenschaften* zeigen hinsichtlich der Transparenz der Leistungs- und Prüfungsanforderungen weniger Unterschiede. Allerdings lassen sich Tendenzen dafür erkennen, dass die Studentinnen die weiblichen Lehrenden diesbezüglich etwas besser beurteilen – für knapp die Hälfte waren die Anforderungen stets transparent (49%) – während nur 36% der Studenten, die Veranstaltungen von Dozentinnen besucht haben, diese Urteil abgeben (vgl. Abb.8)

Abb. 8: "Leistungs- und Prüfungsanforderungen werden für mich transparent dargestellt"
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



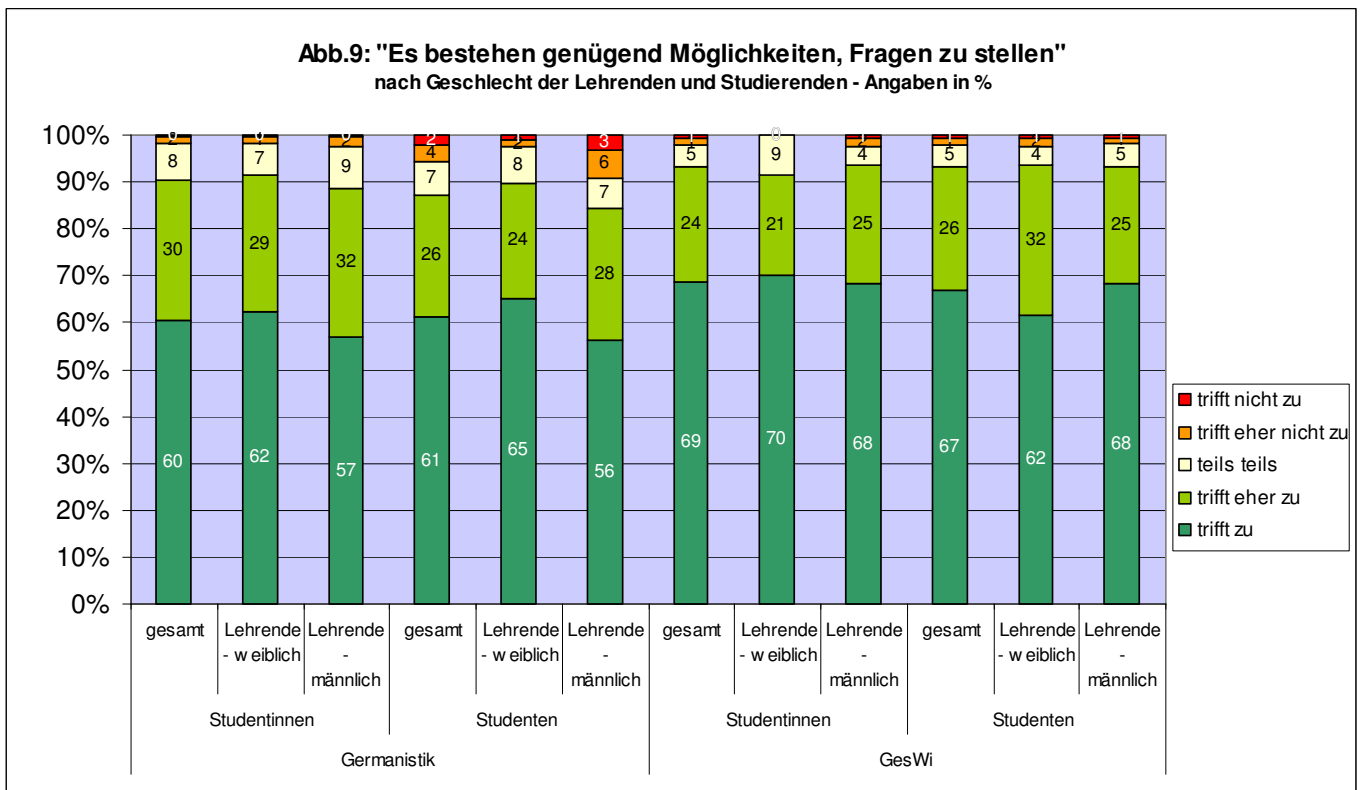
Aus der geschlechtsspezifischen Betrachtung der Evaluationsergebnisse im Bereich „Organisation“ lassen sich zwischen den Fächern relativ deutliche Bewertungsmuster unterscheiden.

In der *Germanistik*, für die die geschlechtsspezifische Auswertung stets signifikante Ergebnisse erbracht hat, geben die Studierenden beiderlei Geschlechts bzgl. jedes Items eine bessere Bewertung der Dozentinnen ab. Ob zur Klarheit der Lernziele, Verdeutlichung der Relevanz der Lehrinhalte oder auch Transparenz der Anforderungen, stets werden die Veranstaltungen der weiblichen Lehrenden besser evaluiert. Die Dozentinnen werden aber auch als besser vorbereitet, in ihren Erklärungsschritten besser nachvollziehbar und in ihrer Fragenbeantwortung als klarer bezeichnet.

Die Evaluationsergebnisse der *Gesellschaftswissenschaften* zeigen Tendenzen zu einer „Same-Gender-Preference“ insbesondere für die Einschätzungen zur Vorbereitung der Lehrenden und zur Transparenz der Leistungs- und Prüfungsanforderungen auf. Hier beurteilen die Studentinnen die Veranstaltungen der Dozentinnen besser, während die Studenten den männlichen Lehrenden gegenüber den Dozentinnen bessere Bewertungen geben. Hinsichtlich der übrigen Items (Klarheit der Lernziele, Relevanz der Lehrinhalte, Nachvollziehbarkeit der Erklärungsschritte und Fragenbeantwortung) erhalten zwar insgesamt die Dozenten bessere Evaluationsergebnisse, auffallend ist dabei jedoch, dass die Studentinnen ihre Dozentinnen im Vergleich zu den Studenten positiver einschätzen.

4.3 Ergebnisse zum Lehrfaktor „Interaktion“

Der Lehrfaktor „Interaktion“ umfasst die Items „Fragemöglichkeiten“ und „Diskussionsmöglichkeiten“, „Aufgreifen studentischer Beiträge“, „Sorge für einen ruhigen Ablauf“, „Arbeitsatmosphäre“ und „Interesse des Lehrenden am Studienerfolg“. Für diese wird im Folgenden eine geschlechtsspezifische Analyse der Einschätzung der Studierenden vorgenommen.



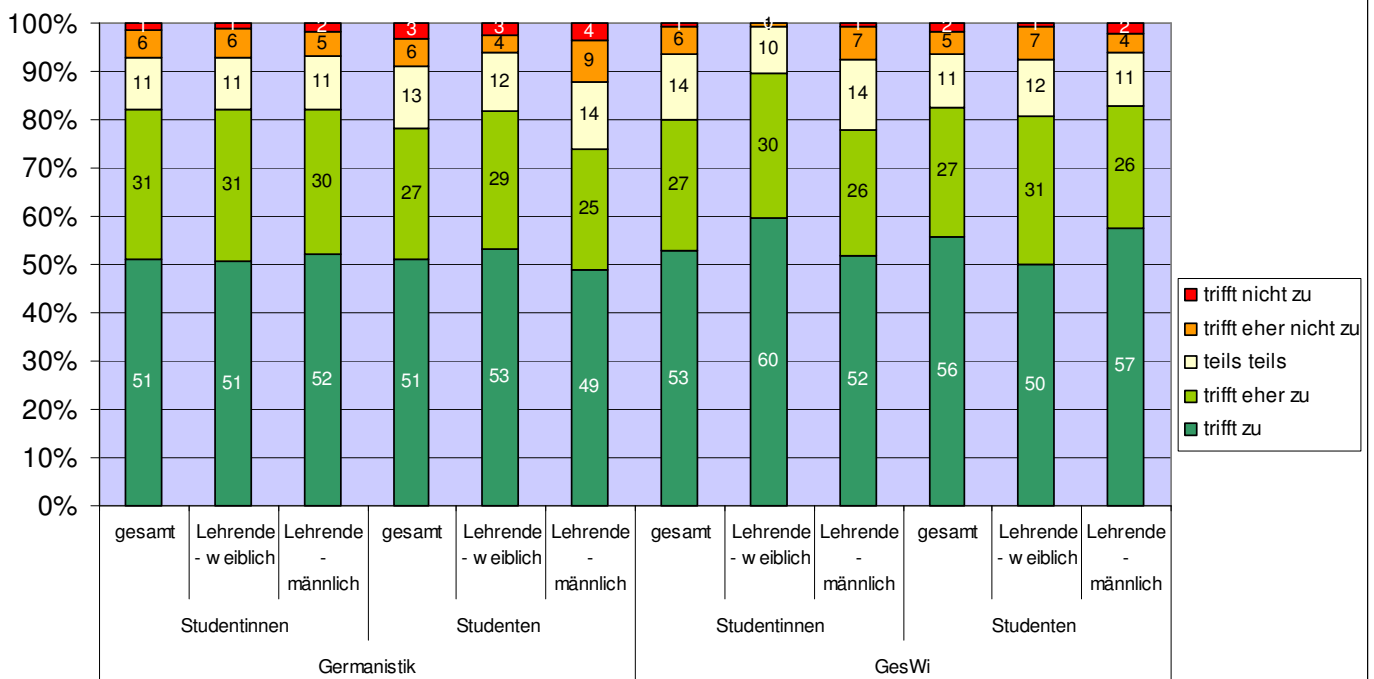
Mit der Möglichkeit, **Fragen in den Veranstaltungen** stellen zu können, sind die Studierenden beider Fächer sehr zufrieden. Die Studentinnen und Studenten der *Germanistik* sind sich in ihren Bewertungen auch auf diesem Gebiet relativ einig: Veranstaltungen von Dozentinnen werden in der Tendenz etwas besser beurteilt als die der Dozenten (vgl. Abb. 9).

Die Studierenden der *Gesellschaftswissenschaften* evaluieren die Veranstaltungen diesbezüglich noch etwas besser. Während sich die Studentinnen und Studenten in der Bewertung der Veranstaltungen von Dozenten einig sind, geben die Studentinnen in Veranstaltungen weiblicher Lehrender ein etwas besseres Urteil ab als ihre männlichen Kommilitonen. 70% der Studentinnen im Vergleich zu 62% der Studenten sprechen von stets genügend Fragemöglichkeiten in Lehrveranstaltungen von Dozentinnen (vgl. Abb.9).

Gelegenheiten zur Diskussionsbeteiligung werden von den Studierenden insgesamt im Vergleich zu den Fragemöglichkeiten etwas weniger gut bewertet. Während sich in der *Germanistik* zwischen Studentinnen und Studenten minimale und daher zu vernachlässigende Unterschiede bei der Einschätzung hinsichtlich ihrer Dozentinnen und Dozenten

erkennen lassen, sind im Antwortverhalten der Studierenden der *Gesellschaftswissenschaften* diesbezüglich Tendenzen erkennbar. In ihrem Gesamturteil sind sich Studentinnen und Studenten zwar relativ einig, das Urteil der Studentinnen in Veranstaltungen weiblicher Lehrender fällt jedoch besser aus, als das der Studenten. 60% der Studentinnen im Vergleich zu 50% der Studenten, die Lehrveranstaltungen von Dozentinnen besucht haben, bezeichnen die Diskussionsmöglichkeiten stets als genügend (vgl. Abb. 10).

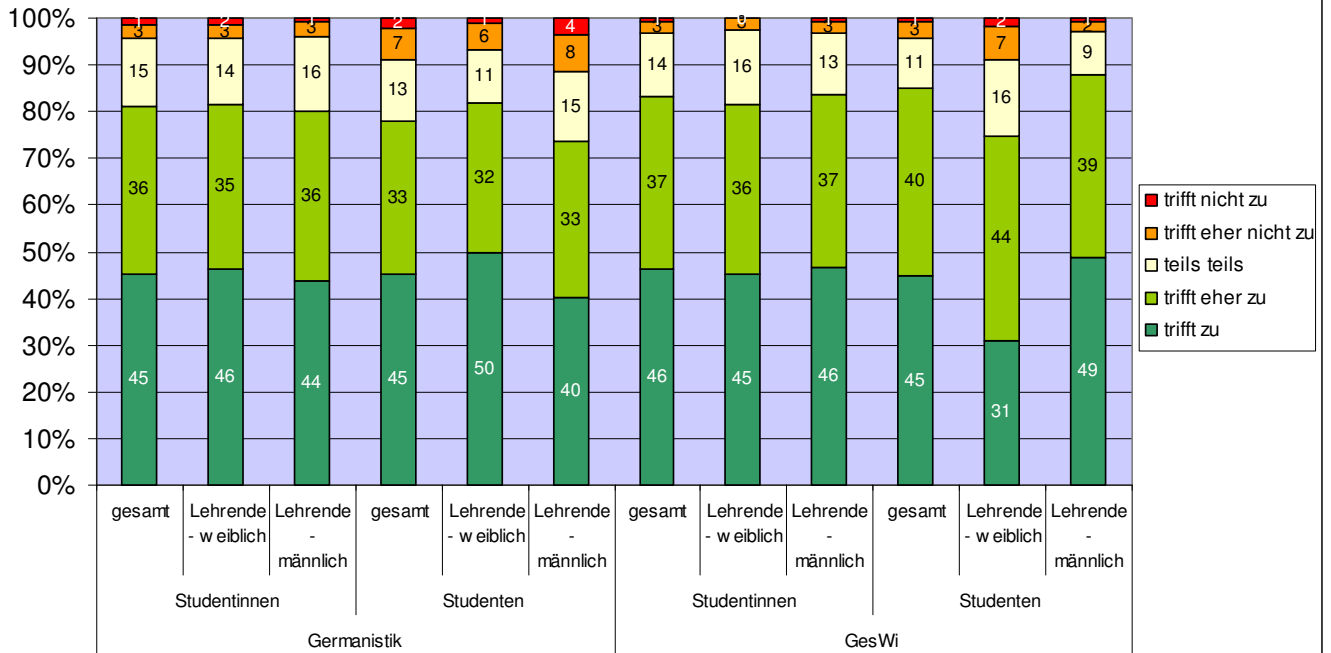
Abb. 10: "Es bestehen genügend Gelegenheiten, sich an Diskussionen zu beteiligen"
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



Das **Aufgreifen studentischer Beiträge** wird von den Studierenden der *Germanistik* tendenziell in den Veranstaltungen der Dozentinnen etwas besser beurteilt. Insbesondere sind die Studenten mit ihren Dozentinnen diesbezüglich zufriedener (50%) als mit ihren Dozenten (40%) (vgl. Abb. 11).

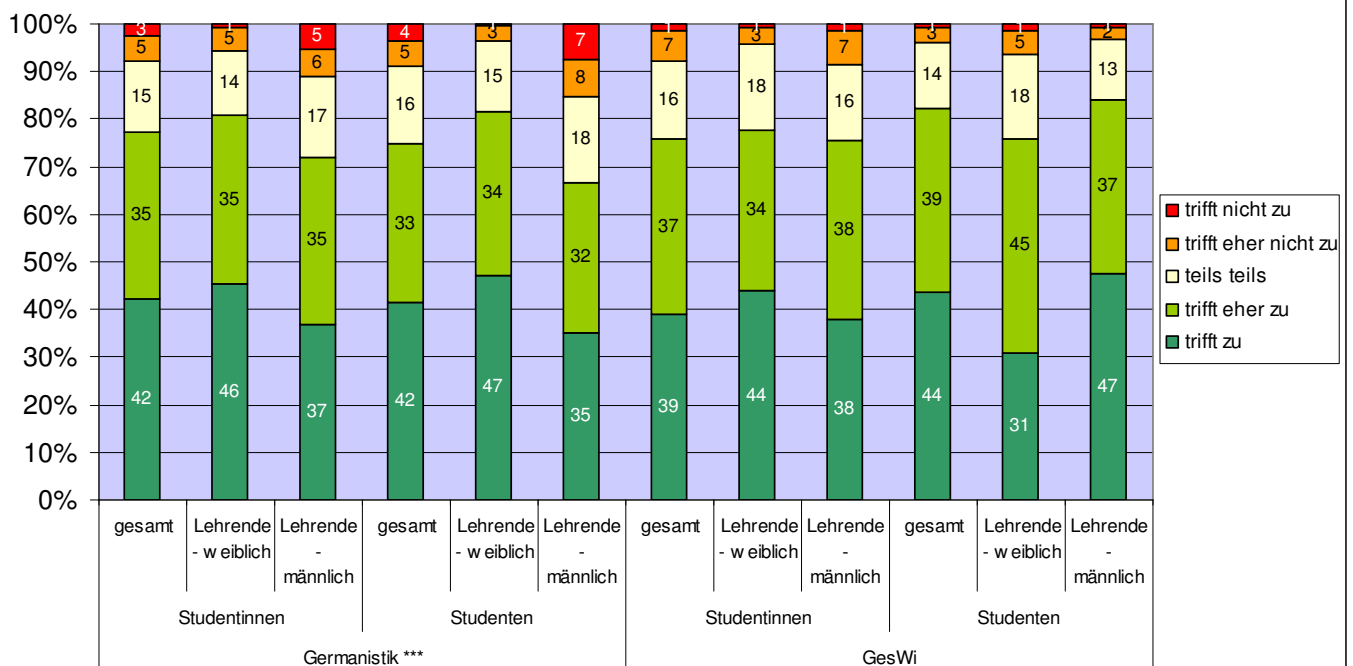
Die Bewertungen der Studentinnen in den *Gesellschaftswissenschaften* unterscheiden sich in den Veranstaltungen ihrer Dozentinnen und Dozenten diesbezüglich nicht. Die Studenten beurteilen ihre Lehrenden hinsichtlich des Geschlechts tendenziell aber unterschiedlich. So geben nur 31% der Studenten in Veranstaltungen weiblicher Lehrender an, studentische Beiträge würden ihrer Meinung nach stets angemessen aufgegriffen, während davon in Veranstaltungen von Dozenten knapp die Hälfte der Studenten (49%) sprechen (vgl. Abb. 11).

Abb. 11: "Die/der Lehrende greift studentische Beiträge für mich angemessen auf"
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



Die **Sorge für einen ruhigen Ablauf der Lehrveranstaltung** bescheinigen in der *Germanistik* die Studierenden überwiegend ihren weiblichen Lehrenden. 46% der Studentinnen und 47% der Studenten geben in Veranstaltungen von Dozentinnen einen ruhigen Veranstaltungsablauf an, während davon bei Dozenten 37% der Studentinnen und 35% der Studenten sprechen (vgl. Abb. 12).

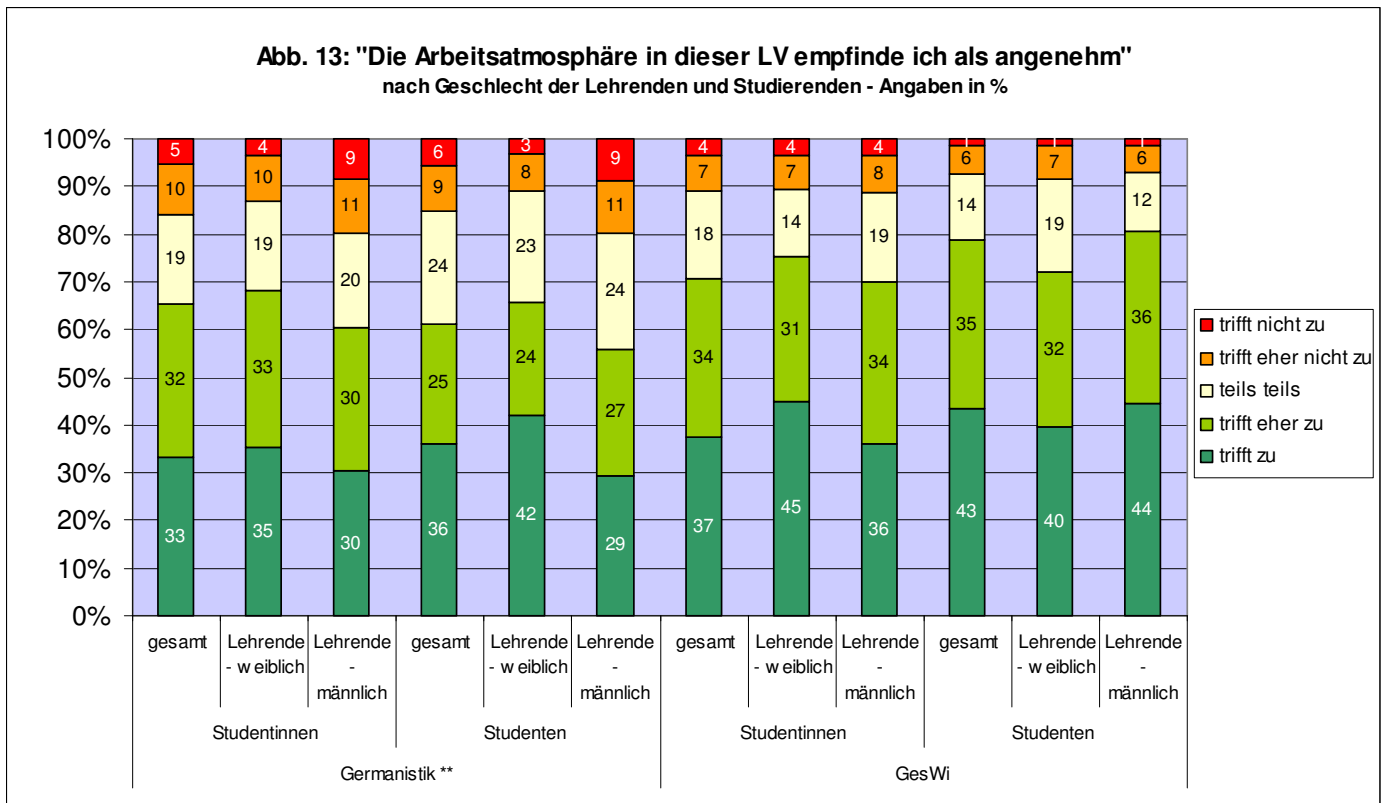
Abb. 12: "Die/der Lehrende sorgt für einen ruhigen Ablauf der LV"
nach Geschlecht der Lehrenden und Studierenden - Angaben in %



Auch bei diesem Item stimmen Studentinnen und Studenten der *Gesellschaftswissenschaften* in ihrem Urteil nicht überein. Während die Studentinnen ihre Dozentinnen im Vergleich zu den Dozenten tendenziell etwas besser beurteilen (44% bzw. 38%), sind es bei den Studenten vor allem die Dozenten, denen ein ruhiger Veranstaltungsablauf bescheinigt wird (47% bzw. 31%) (vgl. Abb. 12).

Eine ähnliche Bewertung geben die Studierenden beider Fächer auch hinsichtlich der **Arbeitsatmosphäre in den Lehrveranstaltungen** ab. Bei den Studierenden der *Germanistik* – insbesondere den Studenten (42%) – fallen die Evaluationsergebnisse in Veranstaltungen der Dozentinnen im Vergleich zu denen der Dozenten etwas besser aus (vgl. Abb. 13).

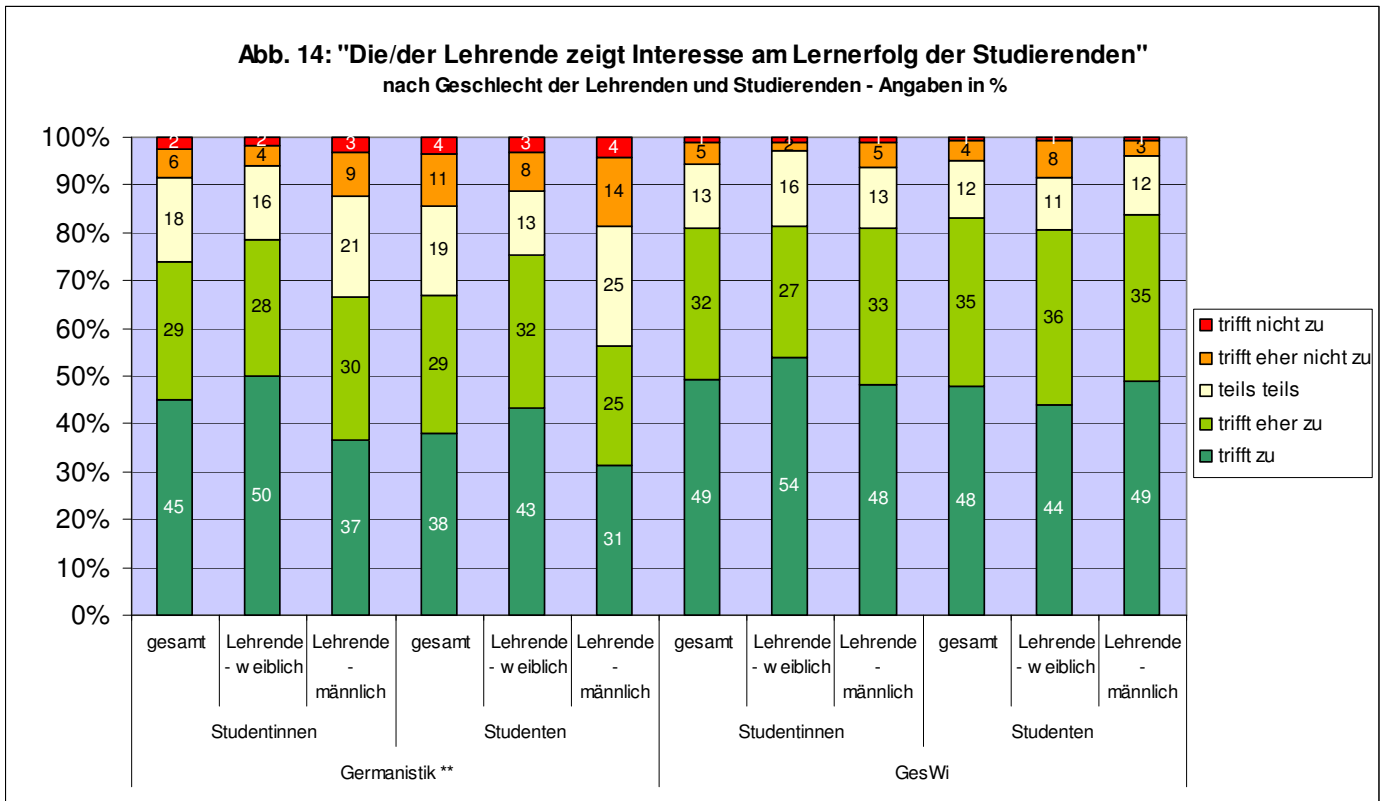
Die Studentinnen der *Gesellschaftswissenschaften*, die Veranstaltungen von Dozentinnen besucht haben (45%), beurteilen diese hinsichtlich der Arbeitsatmosphäre im Vergleich zu den Veranstaltungen der Dozenten als angenehmer. Die Studenten zeigen in ihrer Bewertung dagegen eher eine umgekehrte Tendenz (vgl. Abb. 13).



Auch bei ihrer Einschätzung des **Interesses am Studienerfolg** sind sich die Studentinnen und Studenten der *Germanistik* einig. Ein Interesse am Studienerfolg wird in diesem Fach insbesondere den Dozentinnen bescheinigt. Hierbei sind es vor allem die Studentinnen, die ein großes Interesse feststellen. 50% der Teilnehmerinnen im Vergleich zu 43% der Teilnehmer von Veranstaltungen weiblicher Lehrender geben dieses Urteil ab (vgl. Abb. 14).

Das Interesse am Studienerfolg ihrer männlichen und weiblichen Lehrenden beurteilen die Studentinnen und Studenten der *Gesellschaftswissenschaften* dagegen unterschiedlich.

Während das Interesse der Dozenten von den Studierenden gleich beurteilt wird – 48% der Studentinnen und 49% der Studenten sprechen von diesbezüglichem Interesse –, unterscheiden sie sich tendenziell in der Bewertung der Dozentinnen. Mehr als die Hälfte der Studentinnen (54%), aber 44% der Studenten bescheinigen den weiblichen Lehrenden Interesse am Studienerfolg ihrer Studierenden (vgl. Abb. 14).



Auch auf dem Gebiet „Interaktion“ zeigt die geschlechterspezifische Auswertung der Evaluationsergebnisse der Studierenden der *Germanistik* dasselbe Bewertungsmuster auf. Auch wenn in diesem Fall nicht alle Ergebnisse signifikant sind, so werden Veranstaltungen von Dozentinnen hinsichtlich interaktionsbezogener Aspekte in der Tendenz sowohl von den Studentinnen als auch von den Studenten besser bewertet. Eine Ausnahme bilden die Evaluationsergebnisse zu Diskussionsmöglichkeiten, die keine Unterschiede in der Bewertung der Dozentinnen und Dozenten erkennen lassen.

Für die Studentinnen und Studenten der *Gesellschaftswissenschaften* zeigt die Analyse in diesem Bereich für den überwiegenden Teil der Items die oben angeführte Präferenzstruktur. Hinsichtlich der Möglichkeit, in Veranstaltungen Fragen stellen zu können, sich an Diskussionen zu beteiligen, aber auch bezogen auf den ruhigen Ablauf und die Arbeitsatmosphäre der Veranstaltung sowie auf das Interesse am Studienerfolg evaluieren die Studentinnen die weiblichen Lehrenden vergleichsweise besser als die Studenten, die ihre Dozenten diesbezüglich positiv bewerten. Eine Ausnahme bildet das Aufgreifen studentischer Beiträge, das die Studierenden unabhängig vom Geschlecht eher den Dozenten bescheinigen. Dabei fällt aber wiederum auf, dass die Studentinnen weibliche Lehrende im Vergleich zu den Studenten in dieser Hinsicht deutlich besser beurteilen.

5. Diskussion der Ergebnisse und Ausblick

Aus oben angeführten Ergebnissen der geschlechterspezifischen Analyse der Evaluationsdaten lassen sich Tendenzen dafür erkennen, dass Wirkungen des Geschlechts der Lehrenden keineswegs als zu vernachlässigende Aspekte studentischer Veranstaltungsbewertung bezeichnet werden können. Es ist allerdings einschränkend hervorzuheben, dass die Zusammenhänge nicht sonderlich stark ausgeprägt sind und nicht in allen Fällen signifikante Ergebnisse umfassen.

Auf die oben angeführten Fragestellungen, ob sich Unterschiede zwischen der Veranstaltungsbewertung von weiblichen und männlichen Lehrenden als abhängig vom Geschlecht der Studierenden und der Fachrichtung nachweisen lassen, erlaubt die Datenlage keine klaren Antworten. Hier kommen die eingangs schon erwähnten Probleme der Datenbasis zum Tragen. Aus einer Sekundäranalyse von Lehrevaluationsdaten, die mit Hilfe eines allein auf ein Feedback an den Lehrenden ausgelegten Fragebogens – also ohne Items zur direkten Einschätzung der Leistung des Lehrenden – gewonnen werden, lassen sich keine eindeutigen Zusammenhänge zwischen der Beurteilung der Studierenden und dem Geschlecht der Lehrenden erkennen. Hinzu kommt, dass die Beteiligung der Lehrenden an dem Evaluationsverfahren in den meisten Fällen auf Freiwilligkeit beruht. Das Geschlecht der Lehrenden kann auf dieser Grundlage nicht eindeutig als eine, die Beurteilung der Studierenden verzerrende Bias-Variable nachgewiesen werden. Hierbei ist vielmehr zu vermuten, dass weitere Einflussfaktoren – insbesondere der Status der Lehrenden und deren Lehrerfahrung – einwirken, die im Rahmen einer Sekundäranalyse allerdings nicht kontrolliert werden konnten.

Auch wenn sich auf Basis der Analyseergebnisse keine systematischen Zusammenhänge nachweisen lassen, so deuten sie doch auf einen Forschungsbedarf auf diesem Gebiet hin. So weisen die Ergebnisse auf Tendenzen hin, die eine genauere Forschung sinnvoll erscheinen lassen.

Die Evaluationsergebnisse in der Germanistik zeigen in nahezu allen Bereichen eine bessere Beurteilung der Veranstaltungen von Dozentinnen im Vergleich zu Dozenten auf, die nicht allein dem Zufall geschuldet sein können. So handelt es sich um ein Fach, in das deutlich mehr Studentinnen eingeschrieben sind als Studenten. Unterscheidet man die Evaluationsergebnisse der männlichen und weiblichen Lehrenden nach dem Geschlecht der Studierenden, so werden die Dozentinnen von den Studentinnen und den Studenten gleichermaßen besser bewertet. Dies nicht allein hinsichtlich der Interaktion mit den Studierenden, für die die Forschung höhere Kompetenz von Frauen auf dem Gebiet sozial-emotionaler Aspekte des Lehrverhaltens nachweist, sondern auch im Bereich der Organisation der Lehrveranstaltung, die sich dann auch in den Globalbewertungen niederschlägt. In diesem Fach können diese Ergebnisse damit zusammenhängen, dass Frauen weder unter den Lehrenden noch unter den Studierenden eine Ausnahme darstellen. So ist der Anteil von Frauen und Männern relativ ausgeglichen⁵ und die Studentinnen machen $\frac{3}{4}$ der Studierenden aus (vgl. Tabelle S.7). Ist das Geschlecht der Lehrenden daher ein weniger auffälliges Merkmal, spielt für die Studierenden bei ihrer Evaluation die eigentliche Lehrleistung ohne geschlechtsspezifische Erwartungen die zentrale Rolle. Das für die Dozentinnen durchweg positive Urteil könnte daher auf deren hohes Engagement in der Lehre zurückzuführen sein. Dafür spricht auch, dass sich überwiegend weibliche Lehrende an den beiden Lehrevaluationsphasen beteiligt haben. So weisen die

⁵ Allerdings ist hierbei keine Unterscheidung zwischen Statusgruppen berücksichtigt.

den beiden Lehrevaluationsphasen beteiligt haben. So weisen die Forschungsergebnisse in der Lehrevaluation darauf hin, dass insbesondere diejenigen Lehrenden ihre Veranstaltungen durch die Studierenden evaluieren lassen, die sich in der Lehre besonders engagieren.⁶ Allerdings können für dieses positive Urteil auch weitere Faktoren wie bspw. Art der Veranstaltung oder auch die Veranstaltungszeiten eine Rolle spielen, die im Rahmen dieser Untersuchung allerdings nicht kontrolliert werden konnten.⁷

In den Gesellschaftswissenschaften zeigen sich dagegen Tendenzen, dass die Evaluationsergebnisse für die Veranstaltungen von Dozentinnen und Dozenten je nach Geschlecht der Studierenden anders ausfallen. Gerade im Bereich der Items zur „Interaktion“ – mit Ausnahme des Aufgreifens studentische Beiträge – bewerten die Studentinnen ihre Dozentinnen besser als ihre männlichen Kommilitonen, die dafür ihren Dozenten höhere Bewertungen geben. Dies könnte als ein Hinweis auf die eingangs erwähnten „Same-Gender-Preference in teaching-style“ interpretiert werden. Im Bereich „Organisation“ werden die Dozenten dann unabhängig vom Geschlecht der Studierenden besser bewertet. Allerdings zeigt sich hierbei, dass die Studentinnen weibliche Lehrende stets deutlich besser bewerten als die Studenten, die Veranstaltungen von Dozentinnen besucht haben. Einschätzungen zur Vorbereitung der Lehrenden und zur Transparenz der Leistungs- und Prüfungsanforderungen bilden hierbei insofern eine Ausnahme, als die Studentinnen wiederum ihre Dozentinnen diesbezüglich höher einstufen, während die Studenten ihre Dozenten besser bewerten.

Dass im Zusammenhang mit der studentischen Lehrevaluation „Gender matters“, ist somit zwar nicht eindeutig zu belegen, aber auch nicht von der Hand zu weisen. Aus dem damit verbundenen Forschungsbedarf sollte allerdings nicht der Schluss gezogen werden, die Verfahren der regelmäßig stattfindenden Lehrevaluation entsprechenden Anpassungen zu unterziehen. Aufgrund der methodischen Probleme hinsichtlich der Entwicklung eindimensionaler Qualitätskriterien und entsprechender Untersuchungsdesign zur Qualitätsmessung sind Lehrevaluationen derzeit vor allem in Form von Feedback-Verfahren sinnvoll einsetzbar. Vor dem Hintergrund der damit verbundenen Zielsetzung, zur Qualitätsentwicklung der Lehre beitragen und daher dem Lehrenden konkrete Hinweise zur Verbesserung ihrer/seiner Lehre liefern zu wollen, erscheint dieser Ansatz auch gerechtfertigt. Die damit geschaffene Datenbasis liefert allerdings für eine geschlechterspezifische Analyse nur eine sehr eingeschränkte Grundlage. Eine systematische geschlechterbezogene Datenaufbereitung in das bisherige Verfahren studentischer Lehrveranstaltungsbeurteilung zu implementieren, erscheint aus diesem Grund nicht empfehlenswert. Der so erreichte Informationsgewinn rechtfertigt den damit verbundenen zusätzlichen Auswertungsaufwand nicht.

Die im Rahmen der Auswertung gefunden Hinweise diesbezüglichen Forschungsbedarfs sollten aber auch nicht ignoriert werden. Lassen sich tatsächlich geschlechterbezogene Präferenzen im Lehrstil nachweisen, so würde sich daraus unter Umständen für die Lehrenden ein hochschuldidaktischer Weiterbildungsbedarf ergeben, um den Studentinnen und Studenten hinsichtlich ihrer unterschiedlichen Bedürfnisse gerecht werden zu können. Wären zudem in Fächern, in denen Frauen als Lehrende noch unterrepräsentiert sind,

⁶ Dies bestätigen die hier untersuchten Evaluationsergebnisse insofern auch, als bei jedem betrachteten Item mehr als die Hälfte der Studierenden gute bis sehr gute Bewertungen abgeben.

⁷ So weisen Forschungsergebnisse bspw. darauf hin, dass Pflichtveranstaltungen durchweg schlechter beurteilt werden und auch die Veranstaltungszeiten Auswirkung auf die Evaluation durch die Studierenden haben.

neben den fachspezifischen auch geschlechtsspezifische Erwartungshaltungen der Studierenden nachweisbar, so wäre dieser Zusammenhang nicht nur bei der Dateninterpretation durch die Leitungsebenen zu beachten, sondern entsprechende Beratungsangebote für Frauen zum Umgang mit solchen Erwartungshaltungen zu entwickeln.

Voraussetzung dazu wäre allerdings die Schaffung einer geeigneten Datenbasis unabhängig von der regelmäßig stattfindenden studentischen Lehrveranstaltungsbeurteilung. Es erscheint somit sinnvoll, entsprechende Auswertungen als eigenständige Projekte bspw. im Zusammenhang mit institutionellen Evaluationsverfahren in den Fachbereichen anzubieten. Entscheidend wäre in diesem Zusammenhang vor allem, angemessene Erhebungsinstrumente zu entwickeln, die sich weniger auf die Lehrveranstaltung allgemein und deutlicher auf die Lehrleistung des jeweiligen Dozierenden beziehen und angemessene Untersuchungsdesigns zur Kontrolle weiterer Störgrößen einzusetzen. Sollen die Beurteilungen der Studierenden zunehmend Konsequenzen für das Lehrangebot der Universität haben, so ist der Einsatz diesbezüglich angemessener Erhebungsinstrumente und die Sensibilisierung aller, die mit diesen Evaluationsdaten arbeiten, wesentliche Voraussetzung für eine datengestützte Qualitätsentwicklung der Lehre.

Literaturverzeichnis

- BASOW, SUSAN A. und NANCY T. SILBERG (1987): Student evaluations of college professors: Are female and male professors rated differently? In: *Journal of Educational Psychology*, 79(3) 1987: 308–314.
- BASOW, SUSAN A. (1994): Student Ratings of Professors are not Gender Blind. In: *AWM Newsletter*, Vol. 24, No. 5, Sept-Oct. 1994.
- BASOW, SUSAN A. (1995): Student evaluations of college professors: When gender matters. In: *Journal of Educational Psychology*, 87(4) 1994: 656-665.
- BASOW, SUSAN A. (2000): Best and Worst Professors: Gender Patterns in Students' Choices. In: *Sex Roles*, Vol. 43, No. 5/6 2000: 407-417.
- BERGER, UWE und CHRISTINE SCHLEUBNER (2003): Hängen Ergebnisse einer Lehrveranstaltungs-Evaluation von der Häufigkeit des Veranstaltungsbesuchs ab? In: *Zeitschrift für Pädagogische Psychologie*. 17(2), 2003: 125-131.
- DIEHL, JOERG M. (2001): Studentische Lehrevaluation in den Sozialwissenschaften: Fragebögen, Normen, Probleme. In: Keiner, E. (Hrsg.): *Evaluation (in) der Erziehungswissenschaft*. Weinheim: 63-90.
- FASSINGER, POLLY A.(1995): Understanding classroom interaction: students' and professors' contributions to students' silence. In: *Journal of Higher Education*. 66(1) 1995: 82-96.
- FELDMAN, KENNETH A. (1993): College students' views of male and female college teachers: Part II—Evidence from students' evaluations of their classroom teachers. In: *Research in Higher Education*, Volume 34, Number 2 / April 1993: 151- 211
- GRUBER, SIMONE und PETRA PISTOR (2007): Lehrevaluation - Entscheidend ist, was man daraus macht! Qualitätsentwicklung durch studentische Lehrveranstaltungsbeurteilung am Beispiel des Lehrevaluationsverfahrens an der Universität Duisburg-Essen. In: *Personal- und Organisationsentwicklung in Einrichtungen der Forschung und Lehre (2/2007)*: 39-44.
- HUSBANDS, CHRISTOPHER T. und PATRICIA FOSH (1993): Students' Evaluation of Teaching in Higher Education: experiences from four European countries and some implications of the practice. In: *Assessment & Evaluation in Higher Education*. Vol. 18, No. 2 1993: 95-114.
- LEVENTHAL, LES; PERRY, RAYMOND und PHILIP C. ABRAMI (1977): Effects of Lecturer Quality and Student Perception of Lecturer's Experience on Teacher Ratings and Students Achievement. In: *Journal of Educational Psychology*. Vol. 69, No. 4 1977: 360-374.
- RINDERMANN, HEINER (2001). Die studentische Beurteilung von Lehrveranstaltungen – Forschungsstand und Implikationen. In: Spiel, Christiane (Hrsg.). *Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck*. Münster: 61-88.
- SHEVLIN, MARK ET AL. (2000): The Validity of Student Evaluation of Teaching in Higher Education: love me, love my lectures? In: *Assessment & Evaluation in Higher Education*. Vol. 25, No. 4 2000: 397-405

- SPIEL, CHRISTIANE und P. MARTIN GÖSSLER (2000): Zum Einfluss von Biasvariablen auf die Bewertung universitärer Lehre durch Studierende. In: Zeitschrift für Pädagogische Psychologie. 14(1) 2000: 38-47.
- VIEBAHN, PETER (2007): Hochschullehrer/innen in der Interaktion mit Lernenden: Die Kategorie "Geschlecht". In: Kamphans, Marion und Nicole Auferkorte-Michaelis (Hrsg.): Gender Mainstreaming – Konsequenzen für Forschung, Studium und Lehre. Studien Netzwerk Frauenforschung NRW Nr. 8, Dortmund: 17-23
- WINOCUR, SHARON; SCHOEN, LINDA G., und SIROWATKA, ASTRID H. (1989). Perceptions of male and female academics within a teaching context. In: Research in Higher Education, 30(3) 1989: 317–329.

Anhang

Five Teaching Factors nach Leventhal, Perry und Abrami (1977)

1	Scholarship	1.1	"Discusses points of view other than his own"
		1.2	"Contrasts implications of various theories"
		1.3	"Discusses recent developments in the field"
		1.4	"Presents origins of ideas and concepts"
		1.5	"Gives references for more interesting and involved points"
2	Organization/Clarity	2.1	"Explains clearly"
		2.2	"Is well prepared"
		2.3	"Gives lectures that are easy to outline"
		2.4	"Is careful and precise in answering questions"
		2.5	"Summarizes major points"
3	Instructor-Group Interaction	3.1	"Encourages class discussion"
		3.2	"Invites students to share their knowledge and experiences"
		3.3	"Clarifies thinking by identifying reasons for questions"
		3.4	"Invites criticism of his own ideas"
		3.5	"Knows if class is understanding him or not"
4	Instructor-Individual Student Interaction	4.1	"Has a genuine interest in students"
		4.2	"Is friendly toward students"
		4.3	"Relates to students as individuals"
		4.4	"Recognizes and greets students out of class"
		4.5	"Is accessible to students out of class"
5	Dyna- mism/Enthusiasm	5.1	"Is a dynamic and energetic person"
		5.2	"Has an interesting style of presentation"
		5.3	"Seems to enjoy teaching"
		5.4	"Is enthusiastic about his subject"
		5.5	"Seems to have self-confidence"
	Overall teaching ability		"How would you rate the instructor in general (all-around) teaching ability? (1) an outstanding and stimulating instructor - (5) a poor and inadequate instructor"