



MÜNSTERANER FRAGEBOGEN ZUR EVALUATION VON SEMINAREN - REVIDIERT (MFE-SR)

INSTRUMENT

INSTRUKTION

Die hier dokumentierte Itembatterie zur Bewertung von Seminaren ist ein Basismodul des „Münsteraner Fragebogen zur Evaluation“ und wird ohne eigene Instruktion präsentiert. Allgemeine Informationen zur Befragung erfolgen beim erstmaligen Zugriff auf das System (Verwendung der Daten, technische Voraussetzungen, u. ä. siehe Appendix A). Dabei werden auch demografische Merkmale (siehe Appendix A) der Teilnehmer erfragt. Dieses Basismodul trägt das Kürzel „MFE-Sr“.

ITEMS

Dozent & Didaktik:

8. Ich finde, das Seminar gab einen guten Überblick über das Themengebiet.
9. Der/Die Lehrende benutzte oft Beispiele, die zum Verständnis der Lehrinhalte beitragen.
10. Ich finde, der/die Lehrende ging auf Fragen und Anregungen der Studierenden angemessen ein.
11. Der/Die Lehrende hat das Thema interessant aufgearbeitet.
12. Ich konnte im Verlauf des Seminars die Gliederung immer nachvollziehen.
13. Ich finde, der/die Lehrende teilte die zur Verfügung stehende Zeit gut ein.

Überforderung:

14. Die Inhalte des Seminars waren zu schwierig für mich.
15. Das Tempo der Stoffvermittlung war für mich zu hoch.
16. Der mit dem Seminar verbundene Zeitaufwand hat mich überfordert.

Teilnehmer:

17. Die meisten Teilnehmer waren gut auf die einzelnen Termine vorbereitet.
18. Die meisten Teilnehmer brachten sich aktiv ein.
19. Die meisten Teilnehmer verfolgten das Seminar aufmerksam und mit Interesse.

Materialien:

20. Die in der Vorlesung verwendeten Medien (Folien, Filme, Skizzen, etc.) trugen zum Verständnis der Inhalte bei.
21. Die Qualität der in der Vorlesung verwendeten Medien (Folien, Filme, Skizzen, etc.) war gut.
22. Die Qualität der zusätzlichen Materialien war gut.

ANTWORTVORGABEN

Antwortvorgaben: 7-stufiges Antwortformat mit den Benennungen: 1 = „stimme gar nicht zu“, 2 = „stimme nicht zu“, 3 = „stimme eher nicht zu“, 4 = „neutral“, 5 = „stimme eher zu“, 6 = „stimme zu“ und 7 = „stimme vollkommen zu“ verankert, zusätzlich steht die Antwortoption „nicht sinnvoll beantwortbar“ zur Verfügung.

Ergänzend vorgegebene, hier nicht weiter behandelte Items

1. Wie viele Sitzungen hast Du bei diesem Seminar gefehlt?

Antwortvorgaben: keine, eine, zwei, drei oder mehr Sitzungen

2. Wie viele Stunden hast Du das Seminar im Schnitt pro Woche vor- und nachbereitet?

Antwortvorgaben: Offenes Antwortfeld

3. Das Seminar war meine:

Antwortvorgaben: Erstwahl, Zweitwahl, Drittwahl oder geringer

4. Ich habe dieses spezielle Seminar gewählt (Mehrfachantworten möglich):

Antwortvorgaben: aus Zeitgründen, aus Interesse am Thema, wegen des/r Dozenten/in, kein alternativer Kurs, es waren bei der Wahl keine Informationen verfügbar, weil es eine Pflichtveranstaltung ist

5. Ich finde, die räumliche Ausstattung (Seminarraum, Gruppenräume, Experimentalräume) war angemessen.

6. Die Lautstärke war so, dass ich immer alles gut verstehen konnte.

7. Der Seminartermin passte gut in meine Zeitplanung.

23. Ich habe folgende Materialien zusätzlich zur Veranstaltung benutzt (Mehrfachantworten möglich):

Antwortvorgaben: keine, Folien, Skript, Literaturangaben, Webseite des Dozenten/der Veranstaltung, andere Webseiten, Handout, Sonstiges

24. Ich fand die Menge des Materials, das zu dieser Veranstaltung zur Verfügung gestellt wurde, war...

Antwortvorgaben: zu gering, angemessen, zu groß, nicht sinnvoll beantwortbar

25. Ich habe in der Veranstaltung viel gelernt.

26. Ich würde dieses Seminar anderen Studierenden weiterempfehlen.

Antwortvorgaben: Ja, Nein

27. Im Punktesystem der gymnasialen Oberstufe (0 [ungenügend] bis 15 [sehr gut +]) bewerte ich dieses Seminar mit folgender Punktzahl: ____

Antwortformat: Offenes Antwortfeld

28. Anmerkungen für den/die Lehrende/n (Vorschläge/Lob/konstruktive Kritik):

Antwortformat: Offenes Antwortfeld

THEORETISCHER HINTERGRUND

Die Evaluation der Lehre ist ein wichtiges Instrument zur Qualitätssicherung an Hochschulen. Die studentische Lehrveranstaltungskritik blickt dabei auf eine gut 50-jährige Geschichte zurück (Schmidt & Loßnitzer, 2010). Kontroverse Diskussionen in den 1980er und 1990er Jahren haben sich insbesondere mit der Validität dieser Art von Evaluation beschäftigt (Greenwald, 1997; Mutz, 2003; Rindermann, 1996, 2003 & 2009). Dabei hat sich gezeigt, dass studentische Beurteilungen für Lehrevaluationen geeignet sind. Dementsprechend sind sie zu einer zentralen, teilweise gesetzlich verankerten Standardmethode der Qualitätssicherung an Hochschulen geworden (Rindermann, 2009; Schmidt & Loßnitzer, 2010). Nach Rindermann (1996) kann Lehrevaluation in verschiedenen Bereichen Verbesserungen initiieren: So kann sie die Lehrqualifikation der Lehrenden verbessern, Stärken und Schwächen auf Veranstaltungs-, Fach- oder Universitätsebene aufdecken, Diskussionen zwischen Lehrenden und Studierenden unterstützen, die Mittelvergabe steuern und Weiterbildungsmaßnahmen veranlassen. Souvignier und Gold (2002) fassen die verschiedenen Ziele von Lehrevaluation in „Feedback“, „Steuerung“ und „Forschung“ zusammen.

Seit dem Wintersemester 2000/1 werden im Fach Psychologie an der Universität Münster strukturierte Evaluationsfragebogen zur studentischen Veranstaltungskritik eingesetzt. Dabei wurden zunächst nur die Vorlesungen evaluiert. Eine Seminarevaluation findet seit dem Wintersemester 2002/3 statt. Seit dem Wintersemester 2003/4 wird die Lehrevaluation über eine webbasierte Darbietungsplattform online durchgeführt (vgl. Haaser, Thielsch & Moeck, 2007). Vorteile dieser Online-Erhebungsform sind enorme Zeitersparnisse bei der Erhebung, Auswertung und Präsentation der Daten (vgl. z.B. Göritz, Soucek & Bacher, 2005; Haaser et al., 2007). Außerdem konnte durch sie die Anzahl fehlender Werte durch entsprechend programmierte Rückmeldungen deutlich reduziert und das Feedback für Dozenten durch frei wählbare Vergleichsveranstaltungen verbessert werden. Nach einer Ausweitung der evaluierten Veranstaltungen werden seit 2008 auch vereinzelt so genannte Mixed-Mode Erhebungen durchgeführt, d.h. einzelne Veranstaltungen werden auch mithilfe von Papier-Bleistift Fragebogen evaluiert.

Für den deutschsprachigen Raum liegt bereits eine Reihe von Instrumenten zur Lehrevaluation vor. Eine Übersicht geben Schmidt und Loßnitzer (2010). Sie sind jedoch zum Teil vergleichsweise lang und somit nur bedingt für eine internetgestützte Darbietung und eine gleichzeitige Evaluation unterschiedlicher Veranstaltungen geeignet. Speziell für Online-Erhebungen und Rückmeldungen werden eher kurze und ökonomische Skalen benötigt. Da Lehrevaluationen zudem typischerweise gegen Semesterende stattfinden, fallen diese zeitlich in den neuen Studiengängen oftmals in die angespannte Phase der Prüfungsvorbereitung der Studierenden (vgl. Bechler & Thielsch, 2012). Hohe Ökonomie, d.h. eine möglichst geringe Belastung von Studierenden durch die Befragung, war deshalb ein zentrales Ziel bei der Konstruktion der hier beschriebenen Itematterie. Zusätzlich wurde die Erhebung verschiedener Biasvariablen, wie z. B. Vorinteresse oder räumliche Ausstattung, gewünscht. Dazu wurden in die aktuelle Revision des MFE-S zum MFE-Sr mehrere Items neu aufgenommen. Für die meisten existierenden Lehrevaluationsinstrumente wurden Daten für eine Prüfung ihrer psychometrischen Eigenschaften nur beim ersten Einsatz erhoben und sie wurden später nicht anhand neuer Daten auch konfirmatorisch geprüft. Im Unterschied dazu wird hier die auf Basis früherer explorativer und konfirmatorischer Faktorenanalysen sowie theoretischer Überlegungen zu den gegenüber dem MFE-S neu aufgenommenen Items die für dem MFE-Sr erwartete vier-dimensionale Strukturierung der Antworten konfirmatorisch getestet.

INSTRUMENTENTWICKLUNG

ITEMKONSTRUKTION UND -AUSWAHL

Die vorliegende, gegenüber dem MFE-S erweiterte und revidierte Version des Münsteraner Fragebogens zur Evaluation von Seminaren (MFE-Sr), basiert auf einer Reihe von Vorgängerversionen und einem umfangreichen Modifikationsprozess: Im ersten Schritt wurde auf Basis einer Analyse der Merkmale guter Lehre und ausgehend von empirischen Voruntersuchungen ein erstes Instrument am Fachbereich Psychologie der Universität Münster konstruiert (Grabbe, 2003). Es wurde zwischen dem Wintersemester 2002/3 und dem Sommersemester 2005 jedes Semester eingesetzt. Die Originalversion sollte mit 17 Items drei Dimensionen der Lehrqualität möglichst detailliert erfassen (Grabbe, 2003). Erste Item- und Hauptkomponentenanalysen der im Sommersemester 2005 erhobenen Daten veranlassten eine Verkürzung des Fragebogens auf 14 Items, von denen sich neun drei Hauptkomponenten zuordneten (Haaser, 2006). Für das abschließend beibehaltene Instrument wählte Haaser für diese drei Subskalen bewusst Items aus, die unterschiedliche Aspekte der entsprechenden Konstrukte erfassen sollen. Dies spiegelte sich in teilweise niedrigen Werten für die interne Konsistenz nach Cronbachs Alpha wieder. Dieses Instrument wurde in 2008 einer konfirmatorischen Prüfung unterzogen mit positiver Bestätigung der erwarteten dreifaktoriellen Struktur (MFE-S; Hirschfeld & Thielsch, 2009). Danach zeigte sich jedoch weiterer Revisionsbedarf, resultierend aus den Anforderungen, zum einen

mögliche Biasvariablen zu erheben (inwieweit beispielsweise die studentischen Lehrevaluationen durch die räumliche Ausstattung oder die Terminierung der Veranstaltung beeinflusst sind), zum anderen aber auch den Fokus der Lehrevaluation zu erweitern. Ausgehend von der Literatur und vorliegenden Itementwürfen wurde daher die hier vorgestellte revidierte Version zum MFE-Sr erstellt. Ihre psychometrischen Eigenschaften wurden mit Daten aus dem Sommersemester 2009 explorativ sowie mit Daten aus dem Wintersemester 2009/10 konfirmatorisch geprüft. Ziel war dabei insbesondere, Itemformulierungen zu optimieren und relevante Aspekte möglichst handlungsnah zu erfassen, Biasvariablen zu erheben sowie die Feedbackmöglichkeiten sinnvoll zu erweitern.

DURCHFÜHRUNGSMETHODE

STICHPROBE

Im Wintersemester 09/10 wurden im Rahmen der Lehrevaluation insgesamt 2709 Seminarbewertungen erhoben. Aus diesen wurden 1646 ausgeschlossen, entweder wegen zu vieler fehlender Werte oder weil eine oder mehrere Fragen mit „nicht sinnvoll beantwortbar“ bewertet wurden. Ferner wurden die Daten aus weiteren 406 Fragebogen nicht berücksichtigt, weil sie eine identische Session-ID hatten, d.h. Beurteilungen mehrerer Veranstaltungen durch identische Studierende enthielten. Der so bereinigte und hier verwendete Datensatz umfasst somit noch Evaluationen von 657 Studierenden. 172 von ihnen waren männlich, 459 weiblich. 26 machten keine Angabe zu ihrem Geschlecht. Die Befragten waren zwischen 18 und 48 Jahren alt ($M = 23.08$; $SD = 3.41$). Die Mehrzahl (52.8%) studierte Psychologie im Bachelorstudiengang, die übrigen im Diplomstudiengang (20.7%), im Nebenfach Erziehungswissenschaften (12.5%), im Bachelorstudiengang Bildungsarbeit mit Kindern und Jugendlichen oder im Studiengang Diplom-Pädagogik (9.3%). Rund fünf Prozent beantworteten die entsprechende Frage nicht (siehe Appendix A).

DURCHFÜHRUNG DER STUDIE

Die Lehrevaluation am Fach Psychologie der Universität Münster wird seit dem Wintersemester 2002/3 jeweils am Ende eines Semesters durchgeführt. Für diese Erhebungen wurde eine Online-Plattform, basierend auf PHP kombiniert mit MySQL-Datenbanken, entwickelt (Haaser et al., 2007). Einzelne Veranstaltungen im Nebenfach werden seit 2008 auch mit Papier-Bleistift Fragebogen beurteilt. In alle hier berichteten Analysen wurden Daten aus dem Wintersemester 2009/10 einbezogen.

AUSWERTUNGSMETHODEN

Ausgehend von den Ergebnissen früherer Analysen wurde die Dimensionalität der hier berücksichtigten Antworten konfirmatorisch mit linearen Faktoranalysen und unter Rückgriff auf den konventionellen Maximum Likelihood Schätzer geprüft. Alle entsprechenden Analysen erfolgten mit AMOS.

DESKRIPTIVE STATISTIKEN

Deskriptive Statistiken für die Items und ihre Subskalen (siehe Tabelle 2 und Tabelle 1) liegen vor.

DIMENSIONALITÄT

Ein konfirmatorisches Modell (siehe Tabelle 2) mit vier korrelierten Faktoren erzielt eine akzeptable Passung ($\chi^2 = 301.0$, $df = 84$; $TLI = .95$, $CFI = .96$, $RMSEA = .06$). Dies bestätigt die auf der Basis früherer Analysen der Daten aus dem Sommersemester 2009 erwartete vier-dimensionale Strukturierung der

Antworten zu den 15 Items. Die vier Dimensionen lassen sich interpretieren als Dozent & Didaktik, Überforderung, Teilnehmer und Material. Dozent & Didaktik vereinigt dabei hoch korrelierte Items zum Dozentenverhalten und der Lehrweise in einem Faktor. Als getrennter – aber mit der ersten Skala hoch korrelierter – Faktor wird die Qualität zusätzlicher Materialien bewertet. Die Items auf dem Faktor Teilnehmer behandeln Vorbereitung, Interesse und Engagement der Studierenden, während der Faktor Überforderung die Schwierigkeit der Lehrinhalte, Zeitaufwand und Tempo umfasst. Während weitgehend gute bis sehr gute Itemkennwerte vorliegen, zeigt sich bei Item 16 (Zeitaufwand für das Seminar) eine etwas geringere Ladung auf dem Faktor Überforderung. Womöglich fällt dieser Aspekt qualitativ etwas aus dem Rahmen, da er auch von anderen Faktoren außerhalb der Veranstaltung beeinflusst sein kann. Dennoch ist dieses Item aus Gründen der Rückmeldung interessant und soll beibehalten werden. Die Korrelation zwischen den Skalenmittelwerten variieren zwischen $-.23$ und $.76$ (siehe Tabelle 3).

ITEMKENNWERTE

Für eine Beurteilung der psychometrischen Eigenschaften der Items liegen Itemkennwerte (siehe Tabelle 2) und Interitemkorrelationen (siehe Tabelle 4) vor.

GÜTEKRITERIEN

RELIABILITÄT

Cronbachs Alpha beträgt $.89$ für die sechs Items der Subskala Dozent & Didaktik, $.77$ für die drei Items der Subskala Überforderung, $.82$ für die drei Items der Subskala Teilnehmer und $.83$ für die drei Items der Subskala Materialien. Die interne Konsistenz der Subskalen ist danach als ausreichend bis gut anzusehen, insbesondere bei Berücksichtigung ihrer jeweils geringen Itemzahl. Diese Ergebnisse entsprechen denen aus Studien zu umfangreicheren Lehrevaluationsinstrumenten (z.B. FEVOR, Staufenbiel, 2000; HILVE, Rindermann, 2009; KIEL, Gediga et al., 2000 oder TRIL, Gollwitzer & Schlotz, 2003).

VALIDITÄT

Bereits Marsh (1984) weist auf die grundsätzliche Schwierigkeit hin, Lehrevaluationsinstrumente angemessen zu validieren, da keine allgemein anerkannten Kriterien für gute Lehre existieren. Die meisten Lehrevaluationsinstrumente, wie auch die Itemsammlung von Grabbe (2003), auf der die hier vorgestellte Itembatterie aufbaut, können als Versuch angesehen werden, alle möglichen relevanten Aspekte zu erfassen. Die Reduktion der Itemzahl während der Instrumentenentwicklung hat beim MFE-Sr gegenüber umfangreicheren Instrumenten zu einer Einschränkung der betrachteten Bereiche geführt – mit einer Konzentration auf Auftreten und Didaktik von Lehrenden, das Verhalten der Studierenden, das von den Lehrenden verwendete Material und eine eventuelle Überforderung der Studierenden.

Nach dem multimodalen Bedingungsmodell des Lehrerfolgs stellen die Verhaltensweisen des Dozenten den wichtigsten Bedingungsfaktor des Lehrerfolgs dar (Rindermann, 2009). Dies spricht für die Inhaltsvalidität der hier vorgeschlagenen Subskala Dozent & Didaktik des MFE-Sr. Für die konvergente Validität aller vier Subskalen auf Veranstaltungsebene sprechen nach einer Analyse der Daten aus 93 Veranstaltungen mittlere bis sehr hohe Korrelationen ihrer Summenwerte mit einer Gesamtbewertung der jeweiligen Veranstaltung ($.89$ für die Skala Dozent & Didaktik, $-.33$ für die Skala Überforderung, $.81$ für die Skala Teilnehmer und $.81$ für die Subdimension Materialien). Ihre divergente Validität belegen z.B. nicht signifikante Rangkorrelationen zwischen der ursprünglichen Präferenz für eine Veranstaltung (Erst-, Zweit- oder Drittwahl) und den Summenwerten für

Überforderung, Teilnehmer und Materialien, Obwohl die entsprechende Korrelation von $-.13$ mit dem Summenwert für die Subskala Dozent & Didaktik als statistisch signifikant ausgewiesen wird, ist auch sie jedoch substantiell eher unbedeutend. Ein multivariater Vergleich, mit Veranstaltungsart als unabhängiger Variable und den Summenwerten für die vier MFE-Sr Subskalen zeigt schließlich, dass verschiedene Veranstaltungen signifikant unterschiedlich mithilfe dieser Itembatterien beurteilt werden ($F = 3.06$, $df = 368$, $p < .01$, η -Quadrat = $.33$). Dies spricht für ihre diskriminative Validität.

Ein zentrales Ziel der Konstruktion der hier dokumentierten Subskalen war auch ihre Ökonomie, die mit nur sechs bzw. drei Items pro Subskala sicherlich als zufriedenstellend beurteilt werden kann, auch wenn das Instrument gegenüber dem MFE-S zur Erfassung einer weiteren Facette mehr Items umfasst. Sehr hohe Rücklaufquoten (im Mittel über alle Seminarveranstaltungen rund 70%, Thielsch et al., 2010) und regelmäßige Metaevaluations (Thielsch et al., 2010) zeigen zudem, dass Lehrende und Studierende die Befragung mit diesen Items als wenig belastend erleben, das Instrument akzeptieren und es positiv beurteilen.

AUSWERTUNGSHINWEISE

Angesichts der offensichtlichen Eindimensionalität der Items der vier Subskalen können ihre Antwortwerte aufsummiert oder gemittelt werden. Dabei ist zu beachten, dass auf der Skala Überforderung niedrige bis mittlere Werte erwünscht sind. Das Evaluationssystem im Fach Psychologie an der Universität Münster erlaubt den Dozenten einen Vergleich der Ergebnisse für verschiedene Veranstaltungstypen.

Wurde eine Veranstaltung von mehreren Lehrenden abgehalten, so wird die Skala Dozent & Didaktik jeweils für jeden Lehrenden dargeboten. Die Studierenden werden dann instruiert, mit dieser Skala jede/n Veranstalter/in einzeln zu evaluieren. Alle übrigen Skalen und Items beziehen sich auf die Gesamtveranstaltung und werden entsprechend nur einmal abgefragt.

Ein Evaluationsdatensatz wird nur dann für die Auswertung verwendet, wenn der Evaluierende dem am Ende der Online-Befragung explizit zustimmt, ein freiwilliger Selbstausschluss der Daten wird so ermöglicht (vgl. Thielsch & Weltzin, 2012).

LITERATUR

- Bechler, O. & Thielsch, M. T. (2012). Schwierigkeiten bei der Vorbereitung auf schriftliche Prüfungen. *Zeitschrift für Hochschulentwicklung*, 7 (3), 137-156.
- Gediga, G., von Kannen, K., Schnieder, F., Köhne, S., Luck, H., & Schneider, B. (2000). *Kiel: Ein Kommunikationsinstrument für die Evaluation von Lehrveranstaltungen*. Bissendorf: Methodos.
- Göritz, A. S., Soucek, R. & Bacher, J. (2005). Webbasierte Lehrevaluation. In J. Wiemeyer (Hrsg.). *Education, research and new media. Chances and challenges for science* (S. 218-222). Hamburg: Czwalina.
- Gollwitzer, M. & Schlotz, W. (2003). Das „Trierer Inventar zur Lehrveranstaltungsevaluation“ (TRIL): Entwicklung und erste testtheoretische Erprobungen. In: Krampen, G.; Zayer, H. (Hrsg.). *Psychologiedidaktik und Evaluation IV* (S. 114-128). Bonn: Deutscher Psychologen Verlag.
- Grabbe, Y. (2003). *Zwischenbericht: Lehrevaluation. Konstruktion eines Fragebogens für Vorlesungen und Seminare*. Westfälische Wilhelms-Universität Münster. Online verfügbar unter <http://miami.uni-muenster.de/servlets/DocumentServlet?id=5034>

- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction (special issue: student ratings of professors). *American Psychologist*, 52, 1182-1186.
- Haaser, K. (2006). *Revision zweier Fragebögen zur studentischen Lehrveranstaltungsbeurteilung am Fachbereich Psychologie der Westfälischen Wilhelms-Universität Münster*. Unveröffentlichte Diplomarbeit, Westfälische Wilhelms-Universität Münster.
- Haaser, K., Thielsch, M. T. & Moeck, R. (2007). Studentische Lehrveranstaltungsevaluation online: Erfahrungen, Empfehlungen und Standards der Prozessgestaltung. In: M. Krämer, S. Preiser & K. Brusdeylins (Hrsg.). *Psychologiedidaktik und Evaluation VI* (S. 337-346). Göttingen: V&R unipress.
- Hirschfeld, G. & Thielsch, M. T. (2009). Münsteraner Fragebogen zur Evaluation von Seminaren (MFE-S). In A. Glöckner-Rist (Hrsg.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. ZIS Version 13.0. Bonn: GESIS.
- Marsh, H. W. (1984). Students evaluations of university teaching - dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology*, 76(5), 707-754.
- Mutz, R. (2003). Multivariate Reliabilitäts- und Generalisierbarkeitstheorie in der Lehrevaluationsforschung. *Zeitschrift für Pädagogische Psychologie*, 17(3/4), 245-254.
- Rindermann, H. (1996). *Untersuchung zur Brauchbarkeit studentischer Lehrevaluation*. Landau: Verlag Empirische Pädagogik.
- Rindermann, H. (2003). Lehrevaluation an Hochschulen: Schlussfolgerungen aus Forschung und Anwendung für Hochschulunterricht und seine Evaluation. *Zeitschrift für Evaluation*, 2(2), 233-256.
- Rindermann, H. (2009). *Lehrevaluation – Einführung und Überblick zu Forschung und Praxis der Lehrveranstaltungsevaluation an Hochschulen. Mit einem Beitrag zur Evaluation computerbasierter Unterrichts* (2. Aufl.). Landau: Empirische Pädagogik.
- Schmidt, B. & Loßnitzer, T. (2010). Lehrveranstaltungsevaluation: State of the Art, ein Definitionsvorschlag und Entwicklungslinien. *Zeitschrift für Evaluation*, 9(1), 49-72.
- Souvignier, E. & Gold, A. (2002). Fragebögen zur Lehrevaluation: Was können sie leisten? *Zeitschrift für Evaluation*, 1(2), 265-280.
- Staufenbiel, T. (2000). Fragebogen zur Evaluation von universitären Lehrveranstaltungen durch Studierende und Lehrende. *Diagnostica*, 46(4), 169-181.
- Thielsch, M. T., Hirschfeld, G., Dusend, T., Steinhoff, A., Pöttker, J. & Fischer, S. (2010). *Evaluationsbericht Psychologie 2009: Gemeinsamer Bericht über die Evaluationen im Fach Psychologie im WS 08/09 und SoSe 09*. Westfälische Wilhelms-Universität Münster. Online verfügbar unter <http://miami.uni-muenster.de/servlets/DocumentServlet?id=5330>
- Thielsch, M. T. & Weltzin, S. (2012). Online-Umfragen und Online-Mitarbeiterbefragungen. In M. T. Thielsch & T. Brandenburg (Hrsg.), *Praxis der Wirtschaftspsychologie II: Themen und Fallbeispiele für Studium und Praxis* (S. 109-127). Münster: MV Wissenschaft.

QUELLE

ZITIERUNG DIESES ZIS-BEITRAGES

Thielsch, M. T. & Hirschfeld, G. (2012). Münsteraner Fragebogen zur Evaluation von Seminaren – revidiert (MFE-Sr). In A. Glöckner-Rist (Hrsg.). *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. ZIS Version 15.0. Bonn: GESIS.

ORIGINALPUBLIKATION

Vorgängerversionen finden sich dokumentiert bei:

Grabbe, Y. (2003). Zwischenbericht: Lehrevaluation. Konstruktion eines Fragebogens für Vorlesungen und Seminare. Westfälische Wilhelms-Universität Münster. Online verfügbar unter <http://miami.uni-muenster.de/servlets/DocumentServlet?id=5034>

Haaser, K. (2006). *Revision zweier Fragebögen zur studentischen Lehrveranstaltungsbeurteilung am Fachbereich Psychologie der Westfälischen Wilhelms-Universität Münster*. Unveröffentlichte Diplomarbeit, Westfälischen Wilhelms-Universität Münster.

Hirschfeld, G. & Thielsch, M. T. (2009). Münsteraner Fragebogen zur Evaluation von Seminaren (MFE-S). In A. Glöckner-Rist (Hrsg.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. ZIS Version 13.0. Bonn: GESIS.

ADRESSE DER AUTOREN

Adresse der Autoren:

Meinald Thielsch, Dr., Dipl.-Psych., e-mail: thielsch@uni-muenster.de, Westfälische Wilhelms-Universität Münster, Psychologisches Institut 1, Fliednerstr. 21, 48149 Münster, Telefon: 0251 - 83 34154.

Gerrit Hirschfeld, Dr., Dipl.-Psych., e-mail: gerrit.hirschfeld@uni-muenster.de. Vodafone Stiftungsinstitut und Lehrstuhl für Kinderschmerztherapie und Pädiatrische Palliativmedizin Vestische Kinder- und Jugendklinik Datteln, Private Universität Witten/Herdecke, Dr.-Friedrich-Steiner-Str. 5, 45711 Datteln, Telefon: 02363-975-183.

Aufnahme der Dokumentation:

ZIS 14.00, 2010.

Änderungen:

ZIS 15.0, 2012

Korrekturen von Itemnummern und ergänzend vorgelegter Items; Erweiterte Auswertungshinweise. Kleinere Textergänzungen in weiteren Feldern.

TABELLEN

Tabelle 1: Deskriptive Statistiken für Subskalen

Skala	n	Med	M	SD	Schiefte	Kurtosis
Dozent & Didaktik	657	6.17	5.86	1.06	-1.58 (.10)	3.10 (.19)
Überforderung	657	2.00	2.36	1.16	1.09 (.10)	1.06 (.19)
Teilnehmer	657	5.67	5.56	1.04	-0.96 (.10)	1.28 (.19)
Materialien	657	6.00	5.88	0.99	-1.64 (.10)	4.36 (.19)

Anmerkung: Die Berechnungen erfolgten mit SPSS

Tabelle 2: Deskriptive Statistiken für Items (657 Bewertungen)

Item	M	SD	$r_{i, s-i}$	α	Alpha if deleted
Dozent & Didaktik					
Item 8	5.94	1.18	0.74	0.79	0.87
Item 9	5.85	1.28	0.77	0.82	0.87
Item 10	6.09	1.29	0.71	0.76	0.87
Item 11	5.69	1.41	0.79	0.86	0.86
Item 12	5.87	1.31	0.65	0.68	0.88
Item 13	5.71	1.38	0.64	0.67	0.89
Überforderung					
Item 14	2.22	1.27	0.68	0.86	0.62
Item 15	2.19	1.33	0.69	0.87	0.60
Item 16	2.67	1.61	0.48	0.52	0.86
Teilnehmer					
Item 17	5.49	1.26	0.58	0.61	0.85
Item 18	5.58	1.18	0.74	0.82	0.68
Item 19	5.61	1.20	0.71	0.90	0.71
Materialien					
Item 20	6.02	1.08	0.75	0.90	0.70
Item 21	5.96	1.11	0.75	0.88	0.70
Item 23	5.66	1.25	0.57	0.62	0.88

Anmerkungen: $r_{i, s-i}$ = Trennschärfe, α = Ladung; die Berechnungen erfolgten mit SPSS bzw. AMOS

Tabelle 3: Korrelationen der Skalenmittelwerte (657 Bewertungen)

Skala	Dozent&Didaktik	Überforderung	Teilnehmer	Materialien
Dozent&Didaktik	1.00	-.23	.66	.76
Überforderung	-.23	1.00	-.22	-.19
Teilnehmer	.66	-.22	1.00	.65
Materialien	.76	-.19	.65	1.00

Anmerkungen: Alle $ps < .01$; die Berechnungen erfolgten mit SPSS

Tabelle 4: Interitemkorrelationen (657 Bewertungen)

	I_8	I_9	I_10	I_11	I_12	I_13	I_14	I_15	I_16	I_17	I_18	I_19	I_20	I_21	I_23
I_8	1.00	0.64	0.58	0.68	0.59	0.51	-0.20	-0.15	-0.18	0.33	0.38	0.51	0.51	0.52	0.39
I_9	0.64	1.00	0.69	0.70	0.54	0.54	-0.19	-0.15	-0.15	0.31	0.39	0.49	0.55	0.54	0.43
I_10	0.58	0.69	1.00	0.64	0.46	0.53	-0.12	-0.12	-0.24	0.27	0.32	0.47	0.47	0.51	0.37
I_11	0.68	0.70	0.64	1.00	0.58	0.57	-0.15	-0.13	-0.18	0.32	0.45	0.58	0.61	0.57	0.41
I_12	0.59	0.54	0.46	0.58	1.00	0.51	-0.18	-0.16	-0.13	0.33	0.34	0.40	0.49	0.48	0.34
I_13	0.51	0.54	0.53	0.57	0.51	1.00	-0.11	-0.09	-0.22	0.28	0.27	0.40	0.46	0.47	0.34
I_14	-0.20	-0.19	-0.12	-0.15	-0.18	-0.11	1.00	0.75	0.44	-0.08	-0.16	-0.17	-0.14	-0.11	-0.15
I_15	-0.15	-0.15	-0.12	-0.13	-0.16	-0.09	0.75	1.00	0.46	-0.08	-0.16	-0.15	-0.13	-0.14	-0.17
I_16	-0.18	-0.15	-0.24	-0.18	-0.13	-0.22	0.44	0.46	1.00	-0.09	-0.14	-0.21	-0.18	-0.15	-0.15
I_17	0.33	0.31	0.27	0.32	0.33	0.28	-0.08	-0.08	-0.09	1.00	0.55	0.52	0.41	0.37	0.41
I_18	0.38	0.39	0.32	0.45	0.34	0.27	-0.16	-0.16	-0.14	0.55	1.00	0.74	0.45	0.43	0.32
I_19	0.51	0.49	0.47	0.58	0.40	0.40	-0.17	-0.15	-0.21	0.52	0.74	1.00	0.54	0.51	0.39
I_20	0.51	0.55	0.47	0.61	0.49	0.46	-0.14	-0.13	-0.18	0.41	0.45	0.54	1.00	0.79	0.55
I_21	0.52	0.54	0.51	0.57	0.48	0.47	-0.11	-0.14	-0.15	0.37	0.43	0.51	0.79	1.00	0.54
I_23	0.39	0.43	0.37	0.41	0.34	0.34	-0.15	-0.17	-0.15	0.41	0.32	0.39	0.55	0.54	1.00

Anmerkungen: I = Item, die Skala Didaktik & Dozent besteht aus den Items 8 – 13, die Skala Überforderung aus den Items 14 – 16, die Skala Teilnehmer aus den Items 17 - 19 und die Skala Materialien aus den Items 20, 21 und 23. Die Berechnungen erfolgten mit SPSS.

Zu Beginn der Evaluation werden den Studierenden folgende Informationen auf der Startseite online präsentiert:

Herzlich Willkommen zur Lehrevaluation im Semester [Sommer- oder Wintersemester + Jahresangabe]

1. Alter
Antwortvorgabe: Freies Antwortfeld
2. Geschlecht
Antwortvorgaben: weiblich, männlich, keine Angabe
3. Studiengang
Antwortvorgaben: Bachelor Psychologie, Master Psychologie, Diplom Psychologie, Magister Psychologie, Diplom Pädagogik, Nebenfach Erziehungswissenschaften, BA-KJ, Master of Education, Sonstige
4. Fachsemester:
Antwortvorgabe: Freies Antwortfeld
5. ZIV-Kennung (freiwillig)
Antwortvorgabe: Freies Antwortfeld

Anonymität

Die Evaluation erfolgt anonym und deine Daten werden vertraulich behandelt. Soziodemographische Angaben wie z.B. dein Alter und Geschlecht werden nicht zusammen mit deinen Antworten an die Dozenten weitergegeben. Diese Daten werden lediglich für statistische Auswertungen durch das Evaluationsteam verwendet.

Verlosung

Die Informationen, die bei der Verlosung eingegeben werden, werden nicht mit den Daten aus der Evaluation kombiniert.

ZIV-Kennung

Wenn du deine ZIV-Kennung angibst, können wir Eingaben von dir zu verschiedenen Zeitpunkten nachvollziehen. Wir werden nie deine ZIV-Kennung speichern, sondern nur einen Hash (siehe Wikipedia-Erklärung „Hash“: <http://de.wikipedia.org/wiki/Hash>) davon. Wir können dadurch nur nachvollziehen, dass du dieselbe Person bist, jedoch nicht, wer du bist.

weitere Fragen

Solltet ihr noch weitere Fragen haben, so könnt ihr in unseren FAQs (siehe PsyEval-FAQ: <http://www.uni-muenster.de/PsyEval/wordpress/?cat=7>) Antworten finden oder auf den allgemeinen PsyEval-Informationsseiten. Beachte bitte auch unsere PsyEval-Netiquette [Link auf PsyEval-Netiquette: <http://www.uni-muenster.de/PsyEval/wordpress/?cat=14>]!

Cookies

Damit du an der Evaluation teilnehmen kannst, müssen in deinem Browser Cookies aktiviert sein. Cookies sind zwar standardmäßig in den üblichen Browsern aktiviert, solltest du jedoch nach dem Start der Evaluation immer wieder auf dieser Seite landen, so aktiviere bitte Cookies.

Wegen der Cookies müsst ihr leider für jede neue Person das Browser-Fenster einmal schließen, damit ihr eine Demographie eingeben könnt.

Cookies in Firefox

Klicke im Menü "Extras" auf "Einstellungen...". In dem neuen Fenster klicke auf "Datenschutz". Hier kannst du einstellen, wie lange Cookies gespeichert werden. Damit du an der Evaluation teilnehmen kannst, müssen Cookies akzeptiert werden und mindestens bis zur aktuellen Sitzung gespeichert werden.

Cookies in Internet Explorer

Klicke im Menü "Extras" auf "Internetoptionen...". Danach auf den Reiter "Datenschutz", hier sollte entweder für die Internetzone die Sicherheitseinstellung Mittel gelten oder du solltest unter "Erweitert" die "Automatische Cookiebehandlung aufheben" und "Cookies von Erstanbietern annehmen", sowie Sitzungscookies immer annehmen.