

Running head: A CLOSER LOOK ON THE EVALUATION OF AN OBLIGATORY LECTURE
IN RESEARCH METHODS

DO UNDERGRADUATE PSYCHOLOGY STUDENTS EVALUATE AN OBLIGATORY LECTURE IN RESEARCH METHODS UNBIASED FROM OVERALL SATISFACTION WITH THEIR STUDIES?

Achim Elfering¹, Simone Grebner², Silke Wehr³
^{1,2,3}University of Bern, Switzerland

Abstract

Student evaluation of courses is common practice in academic psychology. This study tests whether overall study satisfaction at the beginning of the lecture predicts evaluation results of an obligatory lecture on research methods. Using structural equation modelling judgments from 131 undergraduates resulted in five interrelated scales. Judgments on teaching skills, teacher social competencies, and quality of course material were unrelated to overall satisfaction. Moreover, initial satisfaction was unrelated to the judgments whether the course has increased study motivation. Whether students reported the teacher had increased their interest in the topic, however, was strongly related to initial overall study satisfaction ($\gamma = .46, p < .001$). The lecture on research methods was less interesting to students who initially report lower overall satisfaction with their studies. Hence, initial study satisfaction should be assessed and considered in analysing evaluation data.

Keywords: *undergraduate teaching, course evaluation, study satisfaction, bias*

Do undergraduates evaluate research method lectures unbiased
from overall satisfaction with their studies?

Course evaluation has become standard in higher education. Students evaluate their overall satisfaction with courses using questionnaires. Moreover, they evaluate facets of courses such as didactics, preparation, used teaching methods, relevance to practice, teacher social skills, teacher ability to increase student interest and teacher ability to motivate students. Although increasing information on course quality and feedback to the teachers should facilitate teaching, the reliability, validity, and usefulness of student evaluations is under debate (Gold, 1996; Helmke, 1996; Kromrey, 1996; Rindermann, 1996a, b; 1997 a, b; Rindermann, Kohler, & Meisenberg, 2007).

First, there is concern about construct validity of questionnaires. For instance, can undergraduate students in their second term already evaluate relevance of course content to practice? Second, there is bias when participation in course evaluation is low. Third, course evaluations might be influenced by optional versus required participation. Optional participation is likely to be related to higher interest in course content and more favourable evaluations. Required participation is plausibly less favourable evaluated. The literature shows a consistent association of evaluation scores with interest for the topic (Spiel & Gössler, 2000; Peus et al., 2005; Hoffmann 2009). Low interest therefore, might reduce mean evaluation levels. One should also expect that evaluation of required courses is more affected by overall study satisfaction. These courses or lectures are often very large. Spiel and Gössler (2000) investigated the association of bias and student evaluation at the University of Vienna. Their results from six faculties, 66 courses and lectures and 756 students revealed that bias was stronger in lectures compared to seminars, especially when participation was required. We expect satisfaction with studies prior to the beginning of the course to predict later student evaluations in required undergraduate lectures. Bias should be especially strong in lectures on research method because most interest in methods among is lower than in other topics in particular among undergraduate students. In this study, therefore we tested whether initial overall study satisfaction predicts later evaluation of a research method lecture.

We expect that judgments on teaching skills, teacher social skills, lecture material, teachers ability to increase students interest, and teachers ability to motivate students in a undergraduate method lecture are related to undergraduates overall study satisfaction.

Method

Sample

This questionnaire study took place in University teaching of undergraduate students in their first (70%) and second semester (30%). 131 German-speaking Swiss psychology undergraduates (102 women and 29 men, mean age = 24.1 years, $SD = 4.1$ years) volunteered to participate and filled out a standard evaluation questionnaire and the additional form asking for overall study satisfaction.

Procedure

The lecture on research methods included 15 weeks with two hours per week. The contents were (1) Introduction to scientific thinking, (2) causality, (3) hypotheses and measuring, (4) variables and designs, (5) control of bias variables, (6) the experiment, (7) validity and threats to validity, (8) multifactor designs, (9) repeated measurement designs, (10) correlation, (11) intervening variables, (12) qualitative methods, (13) documentation of research, (14) ethics and evaluation, (15) test. The questionnaire was administered in the mid of 14th lecture, before having a 15 min break in between the 90 min lecture. The Questionnaire on student satisfaction was administered at the beginning of first lecture. The course evaluation questionnaire was administered in the mid of 14th lecture, after a break of 15 min between the two parts of the lecture.

Measures

The scale satisfaction with the studies in general was adopted from a scale that measures job satisfaction used by Baillod and Semmer (1994), based on Oegerli (1984). It measures satisfaction with the studies in general. The original scale has four items, one of which is a Kunin-item asking "How satisfied are you in general with your studies?" This faces scale is widely used to anchor evaluations in measures of overall job satisfaction (Cook, Hepworth, Wall, & Warr, 1981; Kunin, 1955). It ranges from 1 (*exceedingly unsatisfied*) to 7 (*exceedingly satisfied*). The other items ask people to indicate how often they had the following thoughts about their studies: "I hope my academic studies will always remain as good as they are now", and "After days-off, I'm really happy to return to studies", ranging from 1 (*never*) to 7 (*always*). These three items yielded an internal consistency of .66.

The evaluation questionnaire consisted of 15 items from 5 dimensions. The German questionnaire was an adapted version of the questionnaire introduced by Gralki (1991). The five dimensions were (1) "teaching skills" with three items "the teacher is able to explain difficult problems very well", "the teacher has a clear leitmotif", and "uses vivid and concise examples". Cronbach alpha of the scale was .71. The second scale was "social competencies of the teacher" with items "creates a pleasant teaching atmosphere", "has a good sense of humour" and "has a good relationship with students" Cronbach's alpha was .75. The third scale of the evaluation questionnaire was "course material" with items "distributes good teaching material", "work-sheets and scripts etc. were helpful for understanding the course content" and "figures on blackboard and slides were clearly arranged and well structured" Cronbach's alpha was .81. The fourth scale was "teacher increases interest in the topic" with items "I think the course content represents an important part of my education", "the course content was worth knowing and interesting", and "the course offered the possibility to learn a lot" (Cronbach alpha was .82). The fifth scale was "teacher motivates" with items "the teacher showed enthusiasm and engagement for his/her subject", "the teacher motivated me to occupy myself intensively with the topic", and "the teacher is able to motivate students to contribute to the course". Cronbach alpha was .64. The response format to all items was (1) "not at all", to (5) "fully true".

Data Analysis

AMOS 7.0 (Arbuckle & Wothke, 2006) was used to model the latent path structure with all scales as latent variables and test the influence of satisfaction with the studies in general as an exogenous latent variable on “teaching skills”, “social competencies of the teacher”, “course material”, “teacher increases interest in the topic”, and „teacher motivates“ as endogenous latent variables. Correlations between endogenous latent variables were modelled. No other paths were specified. Alpha was set to .05, two-tailed.

Results

Complete data were received from 131 undergraduate students, resulting in a participation rate of 95%. Table 1 shows mean values and correlations between study variables. Mean values in all 5 dimensions of the evaluation questionnaire were around 4 in a response format range between 1 and 5. The lowest value was in an item on motivation with 2.93, the item was “the teacher is able to motivate students to contribute to the course” was typical for a lecture, where interaction with students is restricted. Mean values of items representing general study satisfaction were also high in the first item asking for general satisfaction but lower in the more cognitive second and third item.

Student satisfaction as a latent variable

The standardized path coefficients were high for the three indicators (Kunin-item: $\lambda = .93$); item “I hope my studies will always remain as good as it is now“: ($\lambda = .59$); item “After days-off, I’m really happy to return to studies“: ($\lambda = .50$).

Dimensions of course evaluation as latent variables

Standardized path coefficients were satisfactory for all 15 indicators. Indicators for teaching skills were between .60 and .70 ($\lambda = .69, .68, \text{ and } .64$). Teacher social skills showed also high factor loadings ($\lambda = .72, .79, \text{ and } .62$). Path coefficients in “course material” were $\lambda = .73, .80, \text{ and } .79$; in “teacher increases interest in the topic” path coefficients were $\lambda = .73, .90, \text{ and } .70$. Indicators in „teacher motivates to contribute“ were moderate ($\lambda = .55, .64, \text{ and } .66$). Intercorrelations between latent variables of evaluation ranged between .22 and .87 between teaching skills and social skills. The only significant standardized path from undergraduates overall satisfaction with studies to dimensions of evaluation was to “teacher increases interest in the topic” ($\gamma = .46, p < .001$, cf. Figure 1). There was also no significant path from overall satisfaction with studies on judgments whether the lecture was increasing study motivation ($\gamma = .17, p = .137$). The fit of the structural equation model was acceptable ($\chi^2(120) = 193.42; p = .01; \chi^2/df = 1.61, \text{ RMSEA} = .07$).

Discussion

Student evaluation of courses is common practice in academic psychology. Continuous monitoring of education with standardized evaluation questionnaires may serve to survey quality standards of education. The reliability and validity of such evaluation data is under discussion (Hoffmann, 2009), but most researchers agree that legitimacy and screening for outlying negative teaching are important goals. More important, evaluation should help teachers to improve their teaching via student feedback. The usefulness of student feedbacks in teaching obligatory undergraduate courses in methods, however, may be restricted when student feedbacks are systematically biased by general study satisfaction. Study results showed a significant bias occurred for overall study satisfaction to the dimension “teacher increases interest in the topic”. There was no relation between overall study satisfaction and other dimensions like teacher social skills. Students who were unsatisfied with their studies were less interested in the research method lecture while satisfied students were more interested in the lecture. Further studies should investigate the specificity of this correlation. It is likely that teachers tend to think that undergraduates who are unsatisfied with their studies at the beginning of the semester are less interested in most obligatory courses including their lecture, while undergraduates might expect that method lectures in fact do (further) decrease their satisfaction with studies.

Taken together, bias from satisfaction with studies was limited to teachers interest as one important dimension of course evaluation, but total bias was smaller than expected for a method

lecture that has to be passed by everyone. The high participation rate of 131 out of 150 participants might have reduced potential bias because extreme judgments might be more prominent when participation rates are lower, and in smaller samples extreme values have more influence on correlations.

Limitations

The high participation rate in this study is uncommon and did prevent restriction in variation both in overall study satisfaction and interest in the method lecture. Therefore, further studies should test whether bias from satisfaction depends on participation rate.

Conclusion

In evaluation of obligatory method lectures on undergraduate level general satisfaction with studies should be measured at the beginning of the lecture and considered as a bias variable.

References

- Arbuckle, J. L., & Wothke, W. (2006). *Amos 7.0 user's guide*. Chicago: SPSS Inc.
- Baillo, J. & Semmer, N. (1994). Fluktuation und Berufsverläufe bei Computerfachleuten [Turnover and career paths of computer specialists]. *Zeitschrift für Arbeits- und Organisationspsychologie*, 38, 152–163.
- Cook, J. D., Hepworth, S. J., Wall, T. D. & Warr, P. B. (1981). *The experience of work. A compendium and review of 249 measures and their use*. London: Academic Press.
- Gold, A. (1996). Können Studierende die Qualität der Lehre beurteilen? Einige Anmerkungen zu Rindermanns Antwort an seine Kritiker [Are University students able to judge the quality of teaching? Some comments on Rindermanns' response to his critics]. *Zeitschrift für Pädagogische Psychologie* 10, 147–150.
- Gralki, H.O. (1991). *Studentische Lehrevaluation* [Course evaluation]. Freie Universität Berlin, Hochschuldidaktische Fortbildung und Beratung.
- Helmke, A. (1996). Studentische Evaluation der Lehre - Sackgassen und Perspektiven [Students' evaluations of teaching - dead ends and perspectives]. *Zeitschrift für Pädagogische Psychologie*, 10, 181–186.
- Hoffmann, A. (2009). Studentische Lehrveranstaltungsevaluation – Ein Zerrbild der Qualität akademischer Lehre? [Student's evaluation of university courses – a disturbed picture of the quality of academic teaching?] *Empirische Pädagogik*, 23, 19–49.
- Kromrey, H. (1996). Qualitätsverbesserung in Lehre und Studium statt sogenannter Lehrevaluation. Ein Plädoyer für gute Lehre und gegen schlechte Sozialforschung [Improving the quality of teaching and learning instead of implementing so-called students' evaluations. A plea for better teaching and against poor social research]. *Zeitschrift für Pädagogische Psychologie*, 10, 153–166.
- Kunin, T. (1955). The construction of a new type of job satisfaction measure. *Personnel Psychology*, 8, 65–77.
- Oegerli, K. (1984). Arbeitszufriedenheit. Versuche einer qualitativen Bestimmung [Job satisfaction. A qualitative approach]. Unpublished doctoral dissertation. University of Berne, Switzerland.
- Peus, V., Valerius, G., Schäfer, L., Freyer, T., Berger, M. & Voderholzer, U. (2005). Lehrevaluation in der Medizinischen Fakultät Freiburg, Teil II: Formative Lehrveranstaltungsevaluation. [Teaching evaluation at the Medical Faculty of Freiburg, part II: formative teaching evaluation] *GMS Zeitschrift für Medizinische Ausbildung*, 22, 1–6.
- Rindermann, H. (1996a). *Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen* [Studies on the usefulness of student course evaluation]. Landau: Verlag Empirische Pädagogik.
- Rindermann, H. (1996b). Zur Qualität studentischer Lehrveranstaltungsevaluation: Eine Antwort auf Kritik an der Lehrevaluation [On the quality of students' evaluations of University teaching: An answer to evaluation critique]. *Zeitschrift für Pädagogische Psychologie*, 10, 129–145.
- Rindermann, H. (1997a). Die studentische Beurteilung von Lehrveranstaltungen. Forschungsstand und Implikationen für den Einsatz von Lehrevaluationen [Student course evaluation: Evidence and

- implications for use]. In Jäger, R. S. Lehmann, R.H. & Trost, G. (Hrsg.). *Tests und Trends*. Weinheim: Beltz, S. 12–53.
- Rindermann, H. (1997b). Generalisierbarkeit studentischer Veranstaltungsbeurteilungen [Generalizability of student evaluations]. *Psychologie in Erziehung und Unterricht*, 44, 216–234.
- Rindermann, H., Kohler, J. & Meisenberg, G. (2007). Quality of instruction improved by evaluation and consultation of instructors. *International Journal for Academic Development*, 12, 73–85.
- Spiel, C. & Gössler, P. M. (2000). Zum Einfluß von Biasvariablen auf die Bewertung universitärer Lehre durch Studierende [The effects of biasing variables on students' ratings of courses]. *Zeitschrift für Pädagogische Psychologie*, 14, 38–47.

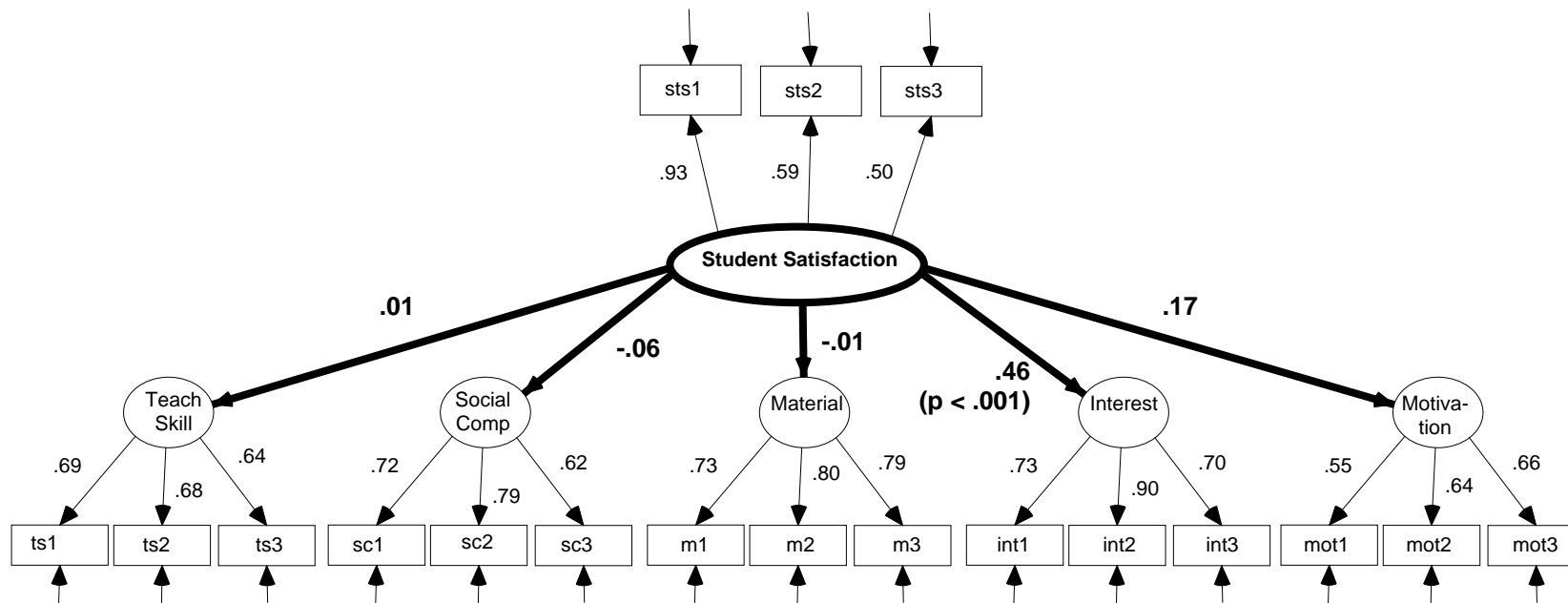
Table 1. Mean values, standard deviations, and correlations among study variables.

	Mean	SD	Range	age	sex	Ts1	Ts2	Ts3	Sc1	Sc2	Sc3	M1	M2	M3	Int1	Int2	Int3	Mot1	Mot2	Mot3	Sts1	Sts2	Sts3	
age	24.11	4.03	20-40																					
sex	1.23	0.42	1,2	.06																				
Ts1	3.68	1.01	1-5	.05	.06																			
Ts2	4.01	0.94	1-5	.06	-.04	.42																		
Ts3	3.99	0.90	1-5	-.04	.08	.49	.45																	
Sc1	3.80	0.94	1-5	.06	-.02	.35	.42	.38																
Sc2	3.68	0.94	1-5	-.01	.05	.51	.46	.51	.54															
Sc3	4.00	0.69	1-5	-.01	.06	.45	.27	.29	.53	.46														
M1	4.02	0.87	1-5	-.02	.01	.40	.51	.35	.30	.24	.27													
M2	4.07	0.88	1-5	-.01	-.10	.35	.42	.31	.22	.23	.24	.60												
M3	4.16	0.79	1-5	.00	-.10	.43	.50	.30	.38	.26	.27	.53	.66											
Int1	4.11	0.93	1-5	.15	.02	.10	.08	.20	.12	.27	.26	.12	.10	.02										
Int2	3.60	0.99	1-5	.11	.10	.17	.19	.17	.20	.25	.26	.25	.10	-.03	.66									
Int3	3.68	0.92	1-5	.15	-.01	.27	.27	.26	.31	.34	.39	.34	.35	.21	.53	.61								
Mot1	3.84	0.96	1-5	.11	.06	.29	.26	.30	.41	.52	.28	.22	.15	.21	.21	.25	.33							
Mot2	3.18	0.94	1-5	.10	.12	.36	.27	.35	.30	.36	.31	.26	.16	.15	.39	.52	.45	.32						
Mot3	2.93	0.90	1-5	.01	.02	.44	.43	.36	.55	.42	.34	.41	.29	.32	.13	.32	.24	.36	.45					
Sts1	5.38	0.91	1-7	-.07	-.05	.03	-.06	.06	-.09	-.06	.04	.03	.03	-.05	.29	.41	.24	-.05	.23	.07				
Sts2	4.07	1.60	1-7	-.11	.04	.01	-.02	.04	-.06	-.09	.04	.03	-.03	-.05	.11	.19	.16	-.07	.22	-.05	.55			
Sts3	4.22	1.22	1-7	.20	.07	.10	.05	.05	.01	.13	.17	-.05	-.06	-.05	.19	.32	.17	-.01	.29	.13	.46	.31		

Notes. . N = 131 participants. Sex (1 = women, 2 = men), Ts1 – Ts3 = Indicators teaching skills, Sc1 - Sc3 = Indicators social competencies, M1 – M3 = Indicators course material, Int1 – Int3 = Indicators interest, Mot1 – Mot3 = Indicators motivation, Sts1 – Sts3 = Indicators student overall satisfaction with studies.

Figure Caption

Figure 1. Structural equation model with satisfaction with the studies in general as exogen latent variable predicting dimensions of method lecture evaluation as endogen latent variables.



Article received: 2010-12-03