International Society of the Learning Sciences

Evaluating Peer Collaboration in Higher Education: Behaviorally Anchored Rating Scales

Verena Schürmann, Rhine-Waal University of Applied Sciences, verena.schuermann@hsrw.eu
Theresa Spahn, Rhine-Waal University of Applied Sciences, theresa.spahn@hsrw.org
Nicki Marquardt, Rhine-Waal University of Applied Sciences, nicki.marquardt@hsrw.eu
Daniel Bodemer, University of Duisburg-Essen, bodemer@uni-due.de

Abstract: Today's challenges are better faced together than alone, highlighting that collaborative working and learning are of particular importance in the 21st century. However, research shows that graduates often do not feel well-prepared in this regard. Appropriate measurement approaches might be helpful for students to reflect their collaborative learning processes and to monitor their skill development. The current study describes the process of constructing an instrument for evaluating the quality of peer collaboration in higher education. Based on theoretical conceptions and in-depth observations of student groups, ten initial Behaviorally Anchored Rating Scales (BARS) were developed. Containing unequivocal behavioral anchors for high, medium and low levels of collaboration, the instrument has the potential to support evaluation processes in higher education. The contribution discusses enriching aspects of the approach but also what is still missing, and depicts potential fields of application.

Introduction

Collaboration is one of the most desired skills in the 21st century. Therefore, students transitioning from higher education into the workforce will be expected to collaborate with others in order to solve complex problems. However, the study by Wilson et al. (2018) highlights that many undergraduate students do not feel well prepared to work collaboratively and do not develop appropriate skills to do so. The existing soft skills gap between graduates' competencies and organizational demands (Abbasi et al., 2018) supports these findings. With regard to this, suitable measurement approaches would help students to reflect their own as well as the groups' learning processes and to monitor their skill development. Behaviorally Anchored Rating Scales (BARS) might be a promising approach to this end. BARS include behavioral anchors as response options that represent different quality levels of a construct with high values representing highly effective behaviors (Debnath et al., 2015). While they are particularly developed and used in terms of job performance ratings (Bernardin et al., 1976), their application in the context of performance in higher education courses also led to promising findings in terms of utility and applicability (e.g., McIntyre & Gilbert, 1994). Moreover, BARS have been found beneficial in similar areas of application, namely for the assessment of team member effectiveness (Ohland et al., 2012) and team adaptation processes (Georganta & Brodbeck, 2020). Respecting the promising findings on their feedback potential (Hom et al., 1982) and their use for self- and peer-evaluation purposes (Ohland et al., 2012), BARS might be a beneficial complementation to existing rating schemes primarily used by trained observers (e.g., the rating scheme for assessing collaboration processes by Meier et al., 2007). Hence, the aim of the current study is to construct an initial system of BARS for evaluating peer collaboration in higher education.

Method

Different procedures and formats for developing BARS have already been used (Bernardin et al., 1976) indicating that the development process is very flexible (Debnath et al., 2015). The procedure described in the current study is comparable to the one used by Georganta and Brodbeck (2020, Study 1). Based on a systematic literature review, we identified four core facets of collaboration including several sub-facets (Schürmann et al., 2022). Second, this coding scheme was pre-tested and used to derive and classify behavioral anchors from observations of student collaboration. The pre-study showed that the facets were suitable and applicable. Additionally, the findings enabled a deeper understanding of favorable behaviors and led to the enrichment and refinement of example indicators for each facet. In line with the method used by Georganta and Brodbeck (2020) the definitions and examples guided the following development stages and serve as description of each sub-facet in the BARS. Next, critical incident analysis (Flanagan, 1954) and frequency analysis were combined in order to derive and classify behavioral anchors for each of the ten sub-facets. Therefore, the video recordings of two collaborating student groups were observed. The recordings stemmed from an experimental field study where 38 undergraduate psychology students (in 14 teams of two to three) participated in a collaborative problem-solving task in a digital setting. In order to maximize variance, a group with a low level of collaboration and a group with a high level of

collaboration (based on self-reports) were selected. Comparing two contrasting cases aimed to identify example behaviors from poor to excellent. The observation was systematically conducted by one rater (second author) using a combination of bottom-up and top-down approach to integrate quantitative and qualitative insights on the student peer collaboration. Critical incident analysis (Flanagan, 1954) was used to develop behavioral examples reflecting a spectrum from desirable to undesirable behaviors for each sub-facet. This process was supported by using frequency analysis to examine the prevalence of behaviors in both groups.

Results and discussion

In line with Ohland et al. (2012), we decided to use a 5-point scale with anchors for high, medium, and low levels of the respective sub-facets of collaboration. The "5", "3" and "1" ratings each comprise four behavioral anchors, reflecting excellent, satisfactory and poor behaviors, respectively. Level "2" and level "4" do not contain behavioral anchors but can be chosen if behaviour is demonstrated which fits to both adjacent levels. In total, ten initial BARS (one for each of the sub-facet) for measuring student peer collaboration were developed. These scales are both based on theoretical concepts and derived from observations of natural and typical collaborative behavior. With respect to the second, combining critical incident analysis and frequency analysis turned out to be a suitable and fruitful approach as qualitative and quantitative data could be integrated and synthesized. While the frequency analysis provided indications of the differences between groups in terms of the distribution of facets, the critical incident technique helped to clarify these analysis results more precisely. Besides these strengths, there are also noteworthy limitations with regard to the development process. The observation and the assignment of the behavioural examples to the sub-facets were primarily done by one rater limiting objectivity and reliability. Although the current version needs to be further developed and evaluated to prove its psychometric characteristics, it already has the potential to be used beneficially. For example, it could be given to students as an instrument to support joint reflective processes about their collaboration (i.e., groups use it as self-evaluation instrument to detect areas of improvement). Thus, students might feel better prepared to work collaboratively and to develop appropriate skills to do so. Supplementary, this approach would allow to get insights into whether students are comfortable or having problems with using the BARS format. Taken together, future research is encouraged to investigate the potential of BARS in higher education context.

References

Abbasi, F. K., Ali, A., & Bibi, N. (2018). Analysis of skill gap for business graduates: managerial perspective from banking industry. *Education and Training*, *60*(4), 354–367.

Bernardin, H. J., LaShells, M. B., Smith, P. C., & Alvares, K. M. (1976). Behavioral expectation scales: Effects of developmental procedures and formats. *Journal of Applied Psychology*, *61*(1), 75–79.

Debnath, S. C., Lee, B. B., & Tandon, S. (2015). Fifty Years and Going Strong: What Makes Behaviorally Anchored Rating Scales So Perennial as an Appraisal Method? *International Journal of Business and Social Science*, *6*(2), 16–25.

Flanagan, J. C. (1954). The Critical Incident Technique. *Psychological Bulletin*, *51*(4), 327–358.

Georganta, E., & Brodbeck, F. C. (2020). Capturing the Four-Phase Team Adaptation Process with Behaviorally Anchored Rating Scales (BARS). *European Journal of Psychological Assessment*, *36*(2), 336–347.

Hom, P. W., DeNisi, A. S., Kinicki, A. J., & Bannister, B. D. (1982). Effectiveness of performance feedback from behaviorally anchored rating scales. *Journal of Applied Psychology*, *67*(5), 568–576.

McIntyre, F. S., & Gilbert, F. W. (1994). Improving Performance in Case Courses: An Argument for Behaviorally Anchored Rating Scales. *Marketing Education Review*, *4*(1), 51–58.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer Supported Collaborative Learning*, *2*, 63–86.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., Layton, R. A., Pomeranz, H. R., & Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self- and peer evaluation. *Academy of Management Learning and Education*, *11*(4), 609–630.

Schürmann, V., Marquardt, N., & Bodemer, D. (2022, August 30 - September 1). *Measuring peer collaboration in higher education and beyond: An integrative framework [Paper presentation].* EARLI SIG 27*, Online Measures at the Crossroad of Ethical and Methodological Challenges,* Southampton, Great Britain.

Wilson, L., Ho, S., & Brookes, R. H. (2018). Student perceptions of teamwork within assessment tasks in undergraduate science degrees. *Assessment and Evaluation in Higher Education*, *43*(5), 786–799.