Effects of a teacher professional development program in formative assessment on teachers' conceptions of feedback and assessment and their self-reported feedback practices

Torulf Palm[1], Gavin Brown[2,3], and Björn Palmberg[4]

[1]Umeå University, Umeå Mathematics Education Research Centre (UMERC), Department of Science and Mathematics Education, Sweden; torulf.palm@umu.se

[2]Umeå University, Department of Applied Educational Science, Sweden; gavin.brown@umu.se

[3]The University of Auckland, Faculty of Education and Social Work, New Zealand; gt.brown@auckland.ac.nz

[4]Umeå University, Umeå Mathematics Education Research Centre (UMERC), Department of Science and Mathematics Education, Sweden; bjorn palmberg@umu.se

Formative assessment has been shown to have the power to improve student achievement. Therefore, many professional initiatives (PD initiatives) have been carried out to support teachers to developing their formative assessment practice. However, accomplishing such practices have been proven difficult. Among factors that are important for outcomes of PDs are teachers' beliefs and conceptions. This study examines the effects of a PD in formative assessment on teachers' conceptions of assessment and feedback, and self-reported feedback practices. These variables were measured through a surves in the beginning and at the end of the PD, and differences between the intervention and a control group were examined at both time points using factor analytic methods. For the intervention group, significant positive differences were found in both the means of important conceptions and in the strength of relationships, while this was not the case for the control group.

Keywords: Beliefs, conceptions, feedback, formative assessment, professional development.

Introduction

Formative assessment is a classroom practice in which teachers and/or students elicit evidence of students' learning needs through assessment and then adapt teaching and/or learning accordingly. It has been shown to have the potential to improve student achievement (e.g., Baird et al., 2014). Some professional development programs (PDs) have also succeeded in helping teachers accomplish formative assessment practices that improve learning in mathematics (e.g., Andersson & Palm, 2017). However, most commonly, PD initiatives have been unsuccessful in accomplishing substantial improvements in teachers' formative assessment (e.g., Bell et al., 2008; Jönsson et al., 2015). Teachers' conceptions of, or beliefs about, assessment and feedback are among factors that affect implementation of formative assessment components (Brown et al., submitted). This paper focuses on the effects of a professional development program in formative assessment on mathematics teachers' conceptions of assessment and feedback, and self-reported feedback processes.

Methods

Design

The professional development program ran as an experimental intervention with control group. A pre- and post-experiment survey was conducted in a northern provincial city in Sweden with a large control group and a small experimental group. Differences between intervention and control groups were examined at both time points using factor analytic methods.

Participants

A total of 461 teachers working between school years 1 and 9 responded to the survey. Among them, 257 teachers responded to the survey at both times. They were matched between time 1 (2021) and time 2 (2023) so the variance over time could be properly evaluated. Expectation maximation was used to impute the small amount of missing data.

Professional development intervention

The PDP was organized by a research team led by the first author. The researchers and the mathematics teachers met once a month during 3-6 hours for three years. The teachers also met by themselves once a month. The meetings included lectures about formative assessment and concrete activities for its implementation, as well as group discussions and analysis of the content and suggested activities. Time was also put aside for the teachers to plan for implementing formative assessment activities in their classrooms. Before the next meeting, the teachers carried out these activities with their students. In the following meeting the teachers evaluated the try-outs, shared experiences of success as well as discussed how they could overcome obstacles and develop the use of a particular activity. The researchers supported these discussions and intervened with suggestions when deemed useful. The teachers were also supported in their self-regulated learning of formative assessment by providing an evaluation tool, and time to use it, for evaluating and setting goals for their practices. Generally, the programme possessed a formative, process-oriented character and also provided support for the teachers to influence the program.

Instruments

The Swedish Teachers Conceptions of Assessment inventory (TCoA) measures three major constructs. For the purposes of this study, the conception that assessment formatively serves teaching and learning (Improvement) was selected as it was most sensitive to the impact of the formative assessment professional development program. This Improvement factor has 4 1st-order factors (i.e., assessment helps teachers improve teaching, assessment helps students improve learning, assessment is reliable, and assessment is diagnostic).

The Swedish Teachers Conceptions of Feedback inventory (TCoF) (Brown et al., 2023) consists of six conceptions and a Formative Feedback Practices factor that is predicted by Improvement and Students Ignore Feedback factors. Of these factors, only four were retained in this study as they were most likely to be sensitive to the professional development (i.e., Students Ignore Feedback, Feedback Improves performance, Feedback Involves Students in Peer and Self-feedback, and Formative Feedback Practices).

Data analysis

The model we used had assessment conceptions predicting feedback conceptions and practices on the assumption that feedback generally occurs after assessment events (Hattie & Timperley, 2007). To account for the repeated measures design, a cross-lagged, bivariate path model with autoregressive paths (Curran & Bollen, 2001) was tested. Within each time point, the assessment conceptions factor with four dependent scores was regressed onto the feedback factor, which had four scores. A path from Student Involvement in Feedback to Formative Feedback Practices was added. Autoregressive paths from each variable at Time 1 were added to the matching variable in Time 2. Unfortunately, no cross-lag paths from Assessment or Feedback factors at Time 1 to Time 2 could be identified. Hence, the model could be described as structural path model within time with auto-regressive paths across time. To compare the model between the two groups, nested invariance testing was conducted (Brown et al., 2017). Path analysis and invariance testing were conducted with AMOS v29.0.0 (IBM, 2022).

Results

Prior to multi-group analysis, the assessment to feedback model with autoregression was found to have acceptable to good fit for the whole group. Also, the input model for the two-group analysis had acceptable fit. Invariance testing showed that measurement weights were not equivalent between groups. Hence, the two groups differed at the unconstrained level, indicating that they were drawn from two separate populations.

The intervention group differed from the control group in significant and substantial ways that were most notable after the intervention itself. Based on a t-test of differences on the change score from pre- to post-intervention time point, the intervention group gained substantially (Cohen's $d \geq .50$) for Assessment Helps Students Improve; Assessment is Reliable, Feedback Formative Practices, and Feedback Improvement variables. These changes indicate that the intervention group's conceptions of assessment and feedback moved substantially in favour of formative assessment, while the means for the control group fundamentally remained constant. Equally notable, there was a positive shift in the strength of the relationships from the assessment and feedback latent factors to their respective items only in the intervention group. Throughout, the control group, as would be expected without any focused professional development, did not change in means or model path values.

Discussion

This study provides an example of a successful attempt to improve teachers' beliefs that would be favorable for implementing formative assessment. Which features of the PD that were decisive for the effects cannot be determined from the study. However, it is possible that the features giving the teachers both time and support for planning implementation in their classes together with support to overcome difficulties may have played a role. Also, the rather substantive length of the PD may have been a factor since belief change often takes time and occurs gradually. Finally, the formative character, and the support for teachers to influence, the PD together with support for taking individual and collective responsibility for their own learning through self-regulated learning processes may have contributed to the positive outcomes.

References

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction, 49*, 92–102. doi:10.1016/j.learninstruc.2016.12.006

Baird, J., Hopfenbeck, T., Newton, P., Stobart, G., & Steen-Utheim, A. (2014). *State of the field review: Assessment and learning*. Oslo: Report for the Norwegian Knowledge Centre for Education, case number 13/4697. Retrieved from http://forskningsradet.no

Bell, C., Steinberg, J., Wiliam, D., & Wylie, C. (2008, March). *Formative assessment and teacher achievement: Two years of implementation of the Keeping Learning on Track Program.* Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY.

Brown, G. T. L., Andersson, C., Winberg, M., & Palm, T. (2023). Predicting formative feedback practices: Improving learning and minimising a tendency to ignore feedback. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1241998

Brown, G. T. L., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: identifying and understanding non-invariance. *International Journal of Research & Method in Education*, *40*(1), 66-90. https://doi.org/10.1080/1743727X.2015.1070823

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107–135). American Psychological Association. https://doi.org/10.1037/10409-004

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81e112. http://dx.doi.org/10.3102/003465430298487.

Jönsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of assessment for learning in Sweden. *Assessment in Education: Prinicples, Policy & Practice, 22*(1), 104–121. doi:10.1080/0969594x.2014.970612