How do students perceive the educational value of peer assessment in relation to its social nature? A survey study in Flanders

Tijs Rotsaert[a,*], Ernesto Panadero[b], Eduardo Estrada[c], Tammy Schellens[a]

[a] *Department of Educational Studies, Faculty of Psychology and Educational Sciences, Ghent University, Ghent, Belgium*
[b] *Departamento de Psicología Evolutiva y de la Educación, Universidad Autónoma de Madrid, Madrid, Spain*
[c] *Department of Psychology, Camilo José Cela University, Madrid, Spain*

ARTICLE INFO

ABSTRACT

This study explores the relationship between students' perceptions of peer assessment (PA) and its social nature. A quantitative survey study (N = 3680) was conducted in secondary education in Flanders, examining the students' perceptions of PA interpersonal variables and their beliefs on the educational value of PA. The structural equation modeling (SEM) results show that the educational value students attribute to PA was positively predicted through trust in their own and their peers' evaluative capabilities, awareness of negative interpersonal processes (e.g. fear of disapproval and friendship marking), and beliefs about PA accuracy. The importance attributed to anonymity appeared to be a negative predictor of PA conceptions. Tests of mean latent differences were performed to explore the differences between educational levels, PA experience and gender.

1. Introduction

Current approaches to assessment have moved away from end-of-course (i.e. summative) testing to in-course (i.e. formative), improvement-oriented interactions between learners and instructors (Black & Wiliam, 1998; Havnes, Smith, Dysthe, & Ludvigsen, 2012). This formative view blurs the strict distinction between instruction and assessment (Cowie, Moreland, & Otrel-Cass, 2013). Formative assessment strategies stress that the active involvement of students in assessment processes is necessary. For this reason, peer assessment (PA) has been embraced as an innovative method of formative assessment and is often seen as a way to offer significant educational value for learning (e.g. Topping, 2010).

In a PA activity, a student generates feedback that might be useful to the assessee, and potentially gives the peer assessor insights as to how their own work might be improved (Dochy, Segers, & Sluijsmans, 1999; Panadero, 2016; Reinholz, 2015). In this study, PA is defined as an interpersonal, collaborative learning arrangement in which students assess their fellow peers' performance by providing peer feedback (PF), which can be quantitative (e.g. grades or ratings across assessment criteria) and/or qualitative (e.g. written or oral comments) (Bolzer, Strijbos, & Fischer, 2015;

Topping, 2010). Despite the benefits of PA, it remains a challenging assessment method to implement. Its social nature due to being a fundamentally interpersonal process has been relatively over-looked and has only been studied in relatively small-scale intervention studies, mainly within vocational and higher education (Panadero, 2016; van Gennip, Seger, & Tillema, 2010).

Furthermore, previous research shows that teachers and students' conceptions about the purpose of assessment largely influence its implementation (Brown, Lake, & Matters, 2011; Segers & Tillema, 2011). To this end, this study aims to explore the relationship between secondary student perceptions related to PA and its social nature.

1.1. Conceptions of peer assessment

Classroom assessment is always a social experience; it does not happen in isolation because students define and practice assessment from their own point of view, and in relation to others (e.g. teachers, peers) (Brown, McInerney, & Liem, 2009). These student conceptions represent ideas, beliefs, propositions, and preferences that fundamentally describe how students experience educational matters such as assessment practices (Brown et al., 2009). In formative assessment, students' conceptions are essential because students need to proces the assessment to learn (Cowie, 2009). Unfortunately, the fact that student conceptions have not yet been explored in sufficient detail, even in formative assessment

* Corresponding author at: Department of Educational Studies, Ghent University, Henri Dunantlaan 2, BE9000 Ghent, Belgium.
E-mail address: Tijs.Rotsaert@UGent.be (T. Rotsaert).

literature, provides a knowledge gap for the present study. Next we will present some of the evidence.

A small number of studies have been conducted on students' conceptions of assessment, indicating their significant contribution to students' learning behavior and (future) learning (e.g. Harris, Brown, & Harnett, 2014; Struyven, Dochy, & Janssens, 2005). In a series of studies, Brown and colleagues report on how to measure student and teachers' conceptions of the purpose of assessment (e.g. Brown & Hirschfeld, 2008). Building on Ajzen's (2005) theory of planned behavior, which suggests that personal intentions or beliefs about what others think shapes ones' behavior, Brown and colleagues argue that a student's response to assessment depends on their appreciation of the process and its aims. Additionally, drawing on Zimmerman and Schunk's (2001) self-regulation theory, they state it is important for students to have a "personally meaningful purpose of assessment," as self-regulated learners often need to use feedback from educational assessment (Brown & Harris, 2011, p. 46).

Therefore, more insight into how students perceive PA is essential. A recent large-scale survey study generated a robust picture of what teachers think about the use of PA (Panadero & Brown, 2017). This study came to the conclusion that, although teachers like the instructional use of PA overall, they struggle with its inherent difficulties (e.g. peer pressure), and only use PA if previous experiences have been positive, which included positive reasons for using PA, beliefs that students should participate in assessment, willingness to include PA in grading and prior use. Interestingly, primary and secondary teachers reported higher values of PA implementation and certainty about its educational value, in contrast with the higher education teachers. Similar results were found in two previous studies that investigated how higher and secondary teachers perceived PA with smaller sample sizes (Lynch & Golen, 1992; Noonan & Duncan, 2005). Nevertheless, the field lacks knowledge about the conceptions of students as they are the most important actors in this fundamentally interpersonal assessment process, and will thus be explored in the present study.

1.2. Peer assessment and its social nature: six interpersonal variables of interest

The majority of research has always been aware of the inherent social and emotional nature of assessment, and this is especially the case for PA (Boud, 1995). Previous research has questioned the accuracy and/or validity and/or reliability of PA (for a detailed discussion see Panadero, Romero, & Strijbos, 2013) due to the presence of possible reciprocity effects caused by interpersonal processes such as friendship marking or psychological risk (Harris & Brown, 2013; van Gennip et al., 2010; Vanderhoven, Raes, Montrieux, Rotsaert, & Schellens, 2015). The limited number of studies on the social nature of PA show that students' perceptions of these interpersonal processes might be related to their conceptions of PA (Cheng & Tsai, 2012; Harris & Brown, 2013; van Gennip et al., 2010). Attention to social and human factors is thus needed because well-implemented PA should decrease negative social problems, assure accuracy, and lead to positive learning outcomes (Panadero & Brown, 2017; Topping, 2010).

Six interpersonal variables frequently referred to in research are discerned in this study (for a detailed discussion see Panadero, 2016). These six variables were chosen because they are the most relevant when it comes to their possible effects on PA, and are as follows. (1) *Friendship marking*, due to friendship bonds, has been mentioned as a source of potential scoring bias. However, only a small number of studies have directly addressed this topic (Panadero et al., 2013). Recent research on the diminishing effect of rubrics on over- and underscoring by peers in PA shows that,

whilst a low or medium friendship reduces the friendship bias, for high-level friendship the rubric seems to amplify the potential friendship bias (Panadero et al., 2013). Additionally, Cheng and Tsai (2012) find that anonymity is preferred for the reason of avoiding the pressure of friendships. (2) *Fear of disapproval* refers to the assessors' fear of negative comments from the assessee if they give them a low score or negative feedback (recrimination) (Cartney, 2010). To decrease this type of fear, it has been argued that anonymity might play a role. For example, in the research of Vanderhoven et al. (2015), students in an anonymous, face-to-face PA setting experienced significantly less fear of disapproval compared to students in a non-anonymous setting. (3) *Psychological safety* refers to a situation in which students have a shared belief about taking interpersonal risks in a group. People that feel psychologically safe tend to perceive differences in opinions as opportunities rather than conflicts (Nicol, 2010; Yu & Sung, 2015). This is important as several authors state that creating a safe environment is a precondition for accurate, and thus valuable, PA activities (Harris & Brown, 2013; van Gennip, Segers, & Tillema, 2009). (4) *Value congruency* refers to the importance of unanimity on both the goals and criteria of the PA activity (Cheng & Tsai, 2012). Rubrics hold the potential to augment the value congruency within a PA activity as they provide the assessment criteria in a structured format and might thus enhance the perceived fairness and comfort with PA (Panadero et al., 2013). (5) *Trust in themself as an assessor* refers to the assessors' beliefs about their skills when assessing a peer (van Gennip et al., 2010). Previous research has indicated that the more the assessor trusts himself/herself, the deeper the learning from PA becomes (Cheng & Tsai, 2012), which has the potential to be increased through intensive practice and interaction (Panadero et al., 2016). (6) *Trust in the other as an assessor* refers to the confidence in the reliability and validity of the assessment and feedback received from a peer. Students will only act on the basis of trustworthy information: if they believe that comments are capricious, they will not act on the basis of them (Carless, 2013).

Two crucial PA aspects that affect the above mentioned variables are anonymity and accuracy. Regarding the first one, as this interplay of interpersonal variables influences the assessment outcome, it has often been stated that decreasing negative social effects via anonymity is desirable (Ballantyne, Hughes, & Mylonas, 2002; Vickerman, 2009) or should at least be explored (Howard, Barrett, & Frick, 2010). Topping (2010) indicates that privacy is an important structural feature of PA, in that disclosing the identity of the assessor or assessee seems to matter to students. Vanderhoven et al. (2015) find that students have more positive attitudes toward PA when anonymity for the assessor was assured, while the participating teacher reported that revealing his/her identity worked as a means to control any undesirable social effects. Yu and Sung (2015) state that anonymity might offer more psychological safety for students, but at the same it might lead to misbehavior, such positive marking toward friends. A recent survey study by Panadero and Brown (2017), which explores the reasons Spanish teachers gave for using PA, reveals that the majority used an anonymous version of PA. It is worth noting, however, that this verdict was not found to be a significant determinant of the frequency of PA use, except for by university teachers. Panadero (2016) in his review recommended to carefully consider anonymity in terms of learning benefits, and whether the positive effects on one hand (e.g. less peer pressure) could come with bigger setbacks (e.g. no interaction assessor and assessee). In conclusion, anonymity needs further research, especially with regard to students' conceptions of the different anonymity modes that can be manipulated in a PA setting (i.e. the anonymity of the assessor, the assessee and the teacher).