Chapter 4

Prejudicial Evaluation: Bias in Self-and-Peer-Assessments of Teamwork Contributions to Design

Richard Tucker
Deakin University, Australia

ABSTRACT

This chapter considers a simple but important question: can students fairly assess each other’s individual contribution to team designs? The chapter focuses on a key problem when using online self-and-peer assessment to individualising design grades for team assignments, namely rater bias – the possibility of students being biased when assessing their own and their peers’ contributions. Three rater-bias issues are considered in depth: (1) self-overmarking; (2) gender bias and gender differences; and (3) out-group bias in the peer assessment of international students in multicultural cohorts. Each issue is explored via the analysis of eight years of quantitative data from the use of an online self-and-peer assessment tool. Evidence is found of self-overmarking and of out-group bias in nonhomogeneous cohorts. However, no evidence is found of gender bias. The chapter concludes with recommendations for design teachers around the assessment of individual contributions to teamwork using self-and-peer assessment.

INTRODUCTION

In teamwork assignments it is possible to assess students as individuals or to award a team mark, and whichever of these two assessment models is used can have, especially for design cohorts, a significant impact on the student learning experience. Assessment recognising individual contribution is pivotal in this context because of the highly emotive nature of designing, and because of the difficulty of assigning authorship to a creative work meaning ‘free-loading’ is difficult to detect. Thus, for students designing in teams it is crucial that they are assessed ‘fairly’ to recognise individual effort. This chapter considers in depth therefore a simple but important question: can we assess individual contributions to team designs through self-and-peer assessment (SAPA)? Specifically, the chapter focuses on a key problem when using self-and-peer assessment to individualise design grades for team assignments, namely rater bias – the possibility of students being biased when assessing their own and their peers’ contributions.

DOI: 10.4018/978-1-5225-0726-0.ch004
The chapter reviews ten year’s of research by the author on the use of self-and-peer-assessment in higher education contexts, specifically in built environment teaching, and focused on students learning to design in teams. Three rater-bias issues are considered in depth that teachers should be aware of: (1) self-overmarking; (2) gender bias and gender differences; and (3) out-group bias in multicultural cohorts. Each issue is explored via the analysis of eight years of quantitative data from the use of an online SAPA tool. The chapter concludes with recommendations for design teachers around the assessment of individual contributions to teamwork using SAPA.

**BACKGROUND**

The impetus for the research presented in this chapter comes from nearly 15 years of teaching architecture students how to design in teams. Early experiences during this period were not always positive for the author or his students. It became increasingly clear that many students did not enjoy team- or group- work, and felt that far too many of their assignments across their curricula required them to work with their peers. The common complaint from students was not about the relevancy of teamwork learning as preparation for professional practice, but that they felt that most teamwork in their courses was for the benefit and convenience of their teachers rather than the students. This opinion reflected the fact that with rapidly increasing student-to-staff ratios, many teachers were attempting to reduce assessment and face-to-face time demands by teaching and assessing groups rather than individuals. As face-to-face time demands are high for design teachers, were one-to-one feedback has always been a common pedagogy, these problems were acute in studio education. Thus, rather than spreading their time thinly over a large number of individual projects, an increasing number of design teachers were setting group projects to allow for longer and more in-depth review sessions on a smaller number of assignment submissions. Commonly in such contexts student teams are required to submit one design per team, and each team member receives the same grade. Furthermore, students communicated that they were more often than not taught very little about how to design in teams.

Such negative feedback from students prompted the author to research if their discontent was supported by data. Early research participated in by students from the courses enrolled in by students from the author’s institution clearly evidenced that unfair assessment was the key complaint, closely followed by poorly designed pedagogies failing to reflect the demands of teamwork and the skills and capabilities students needed to learn to successfully work in teams (R. Tucker, 2008; R. Tucker & Reynolds, 2006; R. Tucker & Rollo, 2006). As one student complained, ‘it is easy to free-ride in a group and, unfairly, it is us, the hard workers, that have to carry the lazy ones’. It became clear to the author that a mechanism would have to be built into the assessment of team-design assignments that rewarded those working hard while penalizing those who were not. During the next couple of years software was developed to individualise grades for team assignments using self-and-peer-assessment conducted anonymously online. The tool, a particular type of SAPA known as Self-and-Peer-Assessment-Continuous-Assessment (SAPCA)), requires students to rate each other’s weekly contributions to team design assignments. Rather than the category-based approach that rates teamwork on explicit multiple dimensions, SAPCA uses the holistic approach to peer assessment, for research has demonstrated that students are more supportive of this method and that their teamwork may be improved using it (Lejk & Wyvill, 2002).

As discussed elsewhere (Tucker, Fermelis, & Palmer, 2009), in a two-year study SAPCA was shown to be robust under the most testing of educational conditions. It was used and evaluated by around two