PERSONALITY’S EFFECT ON PEER ASSESSMENT ABILITY IN CASE METHOD CONTEXT

Tommi Mahlamäki, Teija Valkeinen, Jussi Mylläriemi, Pasi Hellsten, Santeri Repo
Tampere University of Technology (FINLAND)

Abstract

Peer assessment is a current and controversial topic, especially in the educational context. The rapid development of information technology has boosted its use in recent years. One of the most important questions regarding the use of peer assessments concerns their validity. This paper aims to tackle the validity issue by identifying the possible influence of personality types on the ability of students to assess their peers. The research data was gathered from a course that uses the case method. During the course, 62 students made 738 peer assessments, which were then compared with the teachers’ assessments. At the beginning of the course, the students also completed a five-factor model personality trait questionnaire. Finally, the accuracies of the assessments were analysed in relation to the personality traits using a structural equation model. The results show that the conscientiousness and openness to experience had a negative effect on the assessments. Extroversion, agreeableness, and emotional stability had no statistically significant relationship with performance.

Keywords: peer assessment, validity, personality.

1 INTRODUCTION

Peer assessment is a current and controversial topic in the field of education. While the concept of peer assessment has been around for years, the rapid development of information technology has boosted its use in recent years [1]. The increase in the use of peer assessment has been witnessed in higher education and educational settings in general [2, 3]. Similarly, research into peer assessment has grown rapidly during the past twenty years, and a review of the literature reveals that most of the recent studies have dealt with acceptability, effectiveness, fairness, or reliability [4].

Peer assessment can be defined as ‘an arrangement in which individuals consider the amount, level, value, worth, quality, or success of the products or outcomes of learning of peers of similar status’ [5]. The work that is assessed is often written work, but there are many other possibilities as peer assessment is also commonly used in the evaluation of exhibits, presentations, and portfolios [6]. Previous research on peer assessment seems to lack studies that explore how personality traits influence a student’s ability to conduct peer assessments. The aim of this paper is to tackle this research gap; therefore, the goal of the research is to identify the possible influence of different personality types on the ability to assess peers.

The paper continues by introducing the peer assessment method and a description of the benefits and challenges associated to peer assessment. Next, we will explain the research methods and the context of the research. This is followed by the results, and finally, conclusions are drawn.

2 THE USE AND ADVANTAGES OF PEER ASSESSMENT

Peer assessment may have a range of goals, and Gielen et al. [4] state that the quality of peer assessment is highly dependent on these goals. Each goal has a different quality criterion, and therefore, reference must be made the goal of implementing peer assessment in each particular case. One of the most common goals in peer assessment is to use it as an assessment tool (e.g., [4]), although the goal may also be for it to act as complement to the teacher’s evaluation. Among others, Tsai and Liang [3], McGourty [7], and Sivan et al. [8] have shown that student peers display valid scoring that is consistent with the marks of an expert or a teacher. When peer assessment is used responsibly, it can save teachers time [9].

Another advantage of peer assessment is that students can observe their peers during the entire learning process, which means that the students may have better knowledge of their peers’ work than the teacher has [10]. When peers interact to assess one another’s work, the purpose is almost always
formative. Often, the quality of the work done by both parties will improve due to the thinking involved and feedback provided [6]. According to Topping [5], peer assessment can also encourage teachers to provide better guides for the learning process including clarity of objectives and grading scales used in the assessment.

Peer assessment can also be seen as a way of learning (e.g., [11]). In her study, Stefani [12] asked students to evaluate the learning benefits, and the results showed that 85% of the students felt that peer assessment helped them learn more. Using peer assessment assists the students in developing skills, such as communication skills, self-evaluation skills, observation skills, writing skills, and self-criticism [2, 13].

3 CHALLENGES OF PEER ASSESSMENT

While peer assessments provide many kinds of benefits, there are challenges to be considered. One of the most significant challenges has been its reliability and validity (e.g., [14, 15]). For example, many teachers fear that the students’ evaluations will differ significantly from their own [12], and this concern is noteworthy because teacher assessments are often considered to be the best in terms of reliability [14].

Another important challenge is the potential for various biases. Much of the criticism focuses on the relationships between students in a group and the influence of these relationships on a student’s ability to assess his or her peers without bias (e.g., [16]). A common trend is for students to inflate marks when assessing others’ work [17].

Hulsman et al. [16] explored the impact of personality and social reputation as sources of bias in the assessment of communication skills and discovered that peer assessments are vulnerable to the student’s perceived personality and reputation, which is also the case when the assessors are teachers.

4 METHODOLOGY

The goal of this research was to explore how personality traits influence students’ performance in conducting peer assessments. The data was gathered from master’s-level students taking part in an intensive case course covering different management topics ranging from marketing and strategy to accounting. The small group size and lengthy duration of the course were considered ideal for a peer assessment exercise. The enrolment for the course was 62 students; they were divided into five smaller groups, each containing a maximum of fifteen students. At the beginning of the course, the students completed a personality inventory.

At the end of the course, the students were asked to complete a peer assessment of all the other students in their small group, based on the 11 weekly case sessions where all the group members were present. The students were given scoring guidelines, and their task was to score their peers on a scale of 1 to 100. This process resulted in 738 student assessments. Both the assessments and personality tests were completed online. At the same time, the teachers evaluated the students based on the same guidelines and on the same scale.

4.1 Measures

Personality was measured using an established 40-item five-factor model (FFM) personality inventory. The inventory includes items to measure extroversion, agreeableness, conscientiousness, emotional stability, and openness to experience. To analyse the validity of the measures, Cronbach’s alpha and composite reliability statistics were calculated (Table 1). Alpha values ranged from 0.72 to 0.83, which were above the suggested 0.70 level. Similarly, the composite reliabilities were above the suggested 0.70. Validity of the measures were also analysed with average variance explained (AVE) statistics, which were all above the suggested 0.50 level. Correlations between the measures are presented in Table 2.
Assessment performance was identified as the difference between student’s peer assessment and the teacher’s assessment. Assessment performance score was calculated as an average from all the assessments an individual student made. The number of assessments an individual student made varied from 10 to 14 depending on the group size. The lower the assessment performance score, the better the performance, meaning that the student’s peer assessment is closer to the teacher’s assessment.

### 4.2 Structural equation model

The effects of personality traits were analysed using a structural equation model (SEM) where paths were drawn from individual traits to assessment performance. The partial least squares (PLS) method SEM was used with a SmartPLS software. During the analysis some personality trait items were trimmed off from the model. The number of items in the final model ranged from four to six for each of the personality traits. As mentioned earlier, the assessment performance measure was an average measure from 10 to 14 assessments done by the student and the teacher.

### 5 RESULTS

The results show that two of the five personality factors had a statistically significant relationship to assessment performance (Fig. 1). Both conscientiousness ($\beta = -0.27, p = 0.04$) and openness to experience ($\beta = -0.34, p = 0.003$) had a negative effect on assessment performance. Agreeableness had a weak positive effect ($\beta = 0.15, p = 0.325$), but the effect was not statistically significant. Neither extraversion nor emotional stability had a significant relationship with assessment performance. Table 3 presents the path coefficients and p-values for the relationships between the personality traits and assessment performance.

### Table 1. Cronbach’s Alphas and Composite Reliabilities

<table>
<thead>
<tr>
<th></th>
<th>Cronbach’s Alpha</th>
<th>Composite Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>0.72</td>
<td>0.81</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.83</td>
<td>0.88</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>0.70</td>
<td>0.80</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>0.78</td>
<td>0.84</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>0.75</td>
<td>0.83</td>
</tr>
</tbody>
</table>

### Table 2. Construct Correlations

<table>
<thead>
<tr>
<th>E</th>
<th>A</th>
<th>C</th>
<th>ES</th>
<th>OE</th>
<th>AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-0.05</td>
<td>0.11</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>0.62</td>
<td>0.38</td>
<td>-0.13</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>-0.17</td>
<td>-0.08</td>
<td>-0.29</td>
<td>-0.29</td>
<td>1</td>
</tr>
<tr>
<td>Assessment Performance</td>
<td>0.18</td>
<td>0.18</td>
<td>0.30</td>
<td>0.31</td>
<td>-0.38</td>
</tr>
</tbody>
</table>

### Table 3. Path Coefficients and p-values

<table>
<thead>
<tr>
<th></th>
<th>Path Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>0.06</td>
<td>0.975</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>0.15</td>
<td>0.325</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>-0.27</td>
<td>0.040</td>
</tr>
<tr>
<td>Emotional Stability</td>
<td>0.11</td>
<td>0.383</td>
</tr>
<tr>
<td>Openness to Experience</td>
<td>-0.34</td>
<td>0.003</td>
</tr>
</tbody>
</table>
6 CONCLUSIONS

A surprising result is that the conscientiousness trait, which is commonly associated with performance in a range of tasks, negatively affected assessment performance. One possible explanation could be that the students scoring high in conscientiousness were too strict in their evaluations. Another reason for the result could be explained by the nature of the course. The students are evaluated primarily on their participation in the classroom and not on homework. It could be argued that the conscientiousness students are better at case analysis and written tasks than they are on the discussions in the classroom (where the students scoring high on extraversion typically excel). This might cause conscientiousness students to be more critical on the assessments of other students. Either way, deeper analysis is needed to draw definite conclusions.

Openness to experience also negatively affected the ability to assess peers. Openness to experience could manifest itself by giving a different perspective for the assessing. Students scoring high on openness to experience may have a different perception of what performance is, or they might be more open to seeing good performances in situations where others don’t.

In any case, the strong negative effect of openness to experience should be taken into consideration, especially when peer assessments are used in grading. An easy way to deal with the problem is to get a large number of assessments on each student, thus keeping the negative effects relatively low. A more difficult way would be to identify the students scoring high on openness to experience and then keep their assessments from affecting the final grading.

These results raise questions about the validity of peer assessment and whether or not the assessments should be used without any considerations, especially when grading students. Yet, it must still be acknowledged that this research is based on a relatively small sample. It is therefore imperative that more research on the topic be conducted.

---

**Figure 1. Structural Model and Results**
REFERENCES


