Rasooli, Amirhossein, Turner, Jim, Varga-Atkins, Tünde, Pitt, Edd, Asgari, Shaghayegh and Moindrot, Will (2024) *Students' perceptions of fairness in groupwork assessment: validity evidence for peer assessment fairness instrument.* Assessment and Evaluation in Higher Education . pp. 1-16. ISSN 0260-2938.

Students' perceptions of fairness in groupwork assessment: validity evidence for peer assessment fairness instrument

Amirhossein Rasooli, Jim Turner, Tünde Varga-Atkins, Edd Pitt, Shaghayegh Asgari & Will Moindrot

Published online: 27 May 2024.

Submit your article to this journal

Article views: 344

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

Students' perceptions of fairness in groupwork assessment: validity evidence for peer assessment fairness instrument

Amirhossein Rasooli[a] , Jim Turner[b] , Tünde Varga-Atkins[c] , Edd Pitt[d] , Shaghayegh Asgari[e] and Will Moindrot[c]

[a]Learning Sciences and Assessment, National Institute of Education, Singapore, Singapore; [b]Teaching and Learning Academy, Liverpool John Moores University, Liverpool, UK; [c]Centre for Innovation in Education, University of Liverpool, Liverpool, UK; [d]Centre for the Study of Higher Education, University of Kent, Canterbury, UK; [e]BPP University, London, UK

ABSTRACT

Groupwork is a crucial aspect of work contexts and a key twenty first century skill. Assessment of groupwork provides a persistent challenge for educators in university contexts with students reporting experiences of unfairness from their peers during groupwork. This study developed a novel Peer Assessment Fairness Instrument to explore factors driving students' perceptions of fairness during groupwork processes. The results showed that students perceived fairness of groupwork in relation to (1) the grading outcomes they received (i.e. grade congruence), (2) the procedures based on which groups contributed to groupwork (i.e. performative group dynamics), and (3) the relationships based on which group members interacted (i.e. interpersonal treatment). The findings provide student-driven directions to promoting fairer peer assessments within groupwork contexts.

Introduction and literature review

In groupwork, students are expected to collaborate with peers to drive self-directed and collaborative learning to address the problems posed by the tasks and projects they have selected or been assigned. Implementing groupwork is not without its challenges for educators and students alike. Dysfunctional groups, which necessitate educator micromanagement to ensure all students experience a rewarding and enriched experience are often a source of their frustration (Svinicki and Schallert 2016). Such frustration can also be rooted in the culture of avoiding conflict in higher education (Barsky and Wood, 2005).

Groupwork assessment can facilitate positive relational dynamics to encourage students to learn from each other and enable learning potentials that might not be realized vis-à-vis individual learning (Davies 2009). Being able to work functionally within a group, solve problems collaboratively and understand individual responsibility to others are all key graduate skills (Pellegrino and Hilton 2012). While there is a vast body of literature investigating the strategies for designing effective groupwork to equip students with necessary skills across many disciplines, there is no

definitive set of best practices applicable across disciplines on how to design and manage group-work (Tumpa et al. 2022).

Educators sometimes form groups based on performance and ability levels (i.e. low, mid, and high-performing abilities) in order to allow lower-performing students to receive learning support from high-performing students (Williams, Cera Guy, and Shore 2019). Without such interventions, research reports that students tend to naturally select group members on similar achievement status to their own. This unfortunately then leads to the weaker students working with similarly identified students (Mellor 2012). From an outcome point of view, it is perhaps not surprising therefore that higher achieving students work more effectively in groups due to their initial perception of group work being positive and look to reinforce positive interdependence with similar achievement status students (Kwon, Liu, and Johnson 2014).

Students may also value the process of learning from and contributing to collaborative spaces offered in groupwork. However, the forceful presence of grading outcomes might loom large in directing their desires to work with friends and trusted classmates with demonstrated abilities (i.e. group formation) that can contribute to the share of the work promised (i.e. group dynamics) (Gweon et al. 2017; Panadero 2016; Panadero, Romero, and Strijbos 2013). Friendship groups do provide an initial positive effect upon the formation of a group and the effectiveness of early work discussions, but the introduction of graded judgements of peers' contributions (commonly through peer assessment or peer evaluation) can cause problems. Students are not comfortable making negative judgements about their friends (Williams 1992) and may intentionally award higher grades than deserved to their friends (Zhang, Johnston, and Kilic 2008). Similarly, many students do not always possess the requisite aptitude to make objective assessments of their peers' contributions when these sit at odds with their external friendships (Loughry, Ohland, and Moore 2007). This often leads to all students receiving identical marks from their peers which may not correlate with the educator-given grades for each individual student's achievement (Zhang and Ohland 2009).

Research has persistently noted both high and low-achieving students' frustrations with free-riding (i.e. when a group member does not contribute) (Chang and Brickman 2018). In Chang and Brickman (2018) study, regardless of achievement status, all students explained that the minimum expected standard of a successful group was collective responsibility and individual accountability. But when students are individually assessed as part of a group, this can cause maladaptive behaviours of not sharing work with others, failing to support each other, intragroup rivalry, or letting others do all of the work (Meijer et al. 2020). To counteract such issues, educators intentionally engineer strategies such as peer assessment to involve students in holding group members accountable in contributing. This can take the form of anonymous peer evaluations, which take away the worry of friends knowing what their peer has said about them, dissuades free-riding and increases the sense of individual responsibility to work towards the shared goal of the group (Aggarwal and O'Brien 2008).

Educators also aim to promote collaborative work as a significant skill authentic to student needs beyond university contexts. Studies have indicated that engaging in group activities has a beneficial impact on participants' collaboration, such as attending meetings, communication abilities, like expressing opinions and fostering positive involvement, such as motivation (Dijkstra et al. 2016; Forsell, Forslund Frykedal, and Hammar Chiriac 2020). Accordingly, educators design group formation, dynamics, and assessment processes to encourage deeper student engagement with learning tasks (Forsell, Forslund Frykedal, and Hammar Chiriac 2020; Fung, Hung, and Lui 2018).

There is an established body of literature espousing the benefits of groupwork; however, the same cannot be said about how to address students' perceptions of inequity or fairness which leads to scepticism about the learning potential of groupwork in higher education (Hall and Buzwell 2013; Healy, McCutcheon, and Doran 2014). Broadly speaking in recent years, the concern for a fairer and more ethical assessment practice has been seen as the primary target of

'new generation assessment environments' (Nieminen and Tuohilampi 2020). Some emerging research in teacher education has shown signs that students consider self-assessment and peer assessment to make the process of groupwork seem fairer and democratises the assessment due to the increased sense of responsibility and accountability (Ion, Díaz-Vicario, and Mercader 2023). Yet the efficacy of peer assessment processes in promoting more constructive groupwork spaces is yet to be fully empirically examined (Panadero 2016). Specifically, various dimensions contributing to the efficacy of peer assessment in groupwork spaces including students' perceptions of fairness needs additional investigations (Brookhart 2013; Rasooli, Zandi, and DeLuca, 2018).

What we do presently know is that the efficacy of peer assessment for facilitating fair groupwork processes can be defended if students perceive that they have (a) been treated respectfully by peers (i.e. interactional justice, Bies and Moag 1986), (b) received their deserved grades based on their group contributions (i.e. distributive justice, Adams 1965), and (c) had classmates enacting professional groupwork procedures (i.e. procedural justice, Leventhal 1980). These three justice foundations have already been shown to contribute to students' perceptions of fairness in various domains of assessment (Rasooli, Zandi, and DeLuca, 2019). However, a systematic review has recently demonstrated that prior research has only focused on leveraging these dimensions to explore students' perceptions of fairness of their teacher-based assessments (Rasooli, Zandi, and DeLuca, 2023a). Recent novel instruments in classroom assessment fairness also framed educators as the key players in shaping students' perceptions of fairness in assessment (Rasooli et al., 2023b; Sonnleitner and Kovacs 2020). Accordingly, this study constructed a novel Peer Assessment Fairness Instrument (PAFI) to empirically explore students' perceptions of fairness in groupwork assessment spaces. Examining students' perceptions of fairness in groupwork assessment contexts will broaden the evidence base to account for the different sources of fairness perceptions engendered in classroom assessment spaces. The remainder of this manuscript is a report of the validation and initial empirical results for further investigation of this important but underexplored area of research in assessment fairness.

Methodology

Peer assessment fairness instrument development

This study used multiple steps to develop and validate the Peer Assessment Fairness Instrument (PAFI). The Standards for Educational and Psychological Testing (American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME), 2014) advises five sources of evidence (i.e. content, response processes, internal structure, relationship to other variables, and consequences) for supporting the validity of an instrument. From these five sources, this study reports evidence based on content, response processes, and internal structure. The next subsections detail these sources of evidence.

Evidence of validity based on test content and response processes

Social psychology theory including distributive, procedural, and interactional justice dimensions was used in previous measurement instruments to gauge students' and teachers' perceptions of fairness in assessment contexts (Sonnleitner and Kovacs 2020; Rasooli, Zandi, and DeLuca, 2023a; Rasooli et al., 2023b). Given the well-established evidence for using this theory to gauge students' perceptions of fairness, we also relied on this theory to construct the PAFI. However, as we were not aware of any prior instrument specifically evaluating perceptions of fairness in peer assessment contexts, we built on organizational justice literature and instruments to construct the PAFI. In organizational justice literature, 'peer justice climate is defined as the collective perception that individuals—who work together within the same unit and who do not have formal

authority over each other—judge the extent to which they treat one another fairly' (Li, Cropanzano, and Bagger 2013, 3). While organizational justice research has historically explored perceptions of fairness associated with supervisor and managerial entities (Greenberg 1987; Rasooli, Zandi, and DeLuca, 2019), a recent focus on peer justice was aimed to encompass different sources of fairness perceptions (i.e. peers), especially with the heightened contemporary emphasis on group-work skills in the context of organizations (Fortin et al. 2020; Li, Cropanzano, and Bagger 2013; Li and Cropanzano 2009). Fortin et al. (2020) note 'There have been few attempts to investigate how far these classical norms [distributive, procedural, and interactional justice] represent fairness experiences and concerns in modern workplaces, especially in the context of working with peers' (1632–1633). This emphasis resonates with classroom assessment fairness research in broadening the lenses (e.g. educator and peer foci) through which students would perceive fairness in assessment contexts (Rasooli, Zandi, and DeLuca, 2023a).

Li and Cropanzano (2009) conceptualized peer justice as including three dimensions with related items: (1) distributive intraunit justice climate (five items e.g. Some of my teammates have received a better grade for the team projects than they would have deserved.'); (2) procedural intraunit justice climate (five items e.g. 'My teammates are able to express their views and feelings about the way decisions are made in the team.'); (3) Interactional intraunit justice climate (four items e.g. 'My teammates help each other out'). Overall, their instrument reflected three dimensions of distributive, procedural, and interactional justice in contributing to perceptions of peer justice. In a subsequent study, Li, Cropanzano, and Bagger (2013) provided multiple sources of validity evidence to support the measurement adequacy of this instrument.

Given that Li and Cropanzano (2009) relied on prior organizational justice theory and literature to build their instruments, Fortin et al. (2020) conducted exploratory qualitative research to identify additional justice rules that were neglected in prior research. Their qualitative research led to identifying 14 additional justice rules and including them in the new instrument. They conceptualized peer justice as encompassing three dimensions with related items: (1) relationship justice (14 items e.g. 'They are empathetic and supportive.'); (2) task justice (14 items e.g. 'In general, my colleagues apply procedures consistently.'); (3) distributive justice (five items e.g. 'Does your (outcome) reflect the effort you have put into your work?'). Overall, their instrument reflected three dimensions of interactional, procedural, and distributive justice with additional items vis-à-vis Li and Cropanzano's (2009) in contributing to perceptions of peer justice.

Following Li and Cropanzano (2009), Fortin et al. (2020), and social psychology theory in classroom assessment fairness (Rasooli, Zandi, and DeLuca, 2019), we conceptualized the PAFI instrument to include three dimensions: (1) distributive justice, (2) interactional justice, and (3) procedural justice. We combined the items from Li and Cropanzano (2009), Fortin et al. (2020) as the foundation for our new instrument. As a research team of six members, we met in several virtual meetings and communicated numerous asynchronous exchanges to finalize this instrument for our purpose. After removing the duplicates from two instruments for each dimension as well as considering the relevance of items for our assessment context, we retained two items for distributive justice, seven items for interactional justice, and 12 items for procedural justice dimensions. We made considerable revisions to the wording of these items and agreed over the item alignment with each dimension across a number of meetings. After considerable rewording over several iterations, think-aloud protocols were conducted with seven university students to streamline the reading and processing of the item contents for potential participants. For example, item 2 under procedural intraunit justice climate from Li and Cropanzano (2009) was 'The way my teammates make decisions is free from personal bias'. Item 2 under task justice from Fortin et al. (2020) was 'Decisions that they make are free of bias'. We revised these items and included an item in the PAFI 'Decisions that my group members made were free of conscious bias'. We added 'conscious' bias because we acknowledged that unconscious biases could exist all the time without students being aware of it. In our think-aloud protocols and using the student feedback, we revised this item in our final version to be 'Decisions that my group members made

were not biased'. One item was also removed from procedural justice ('My group members took full responsibility for their actions'.) as it overlapped with the following item ('My group members did their allocated tasks properly'.). Altogether, we retained 20 items and included them in our final instrument.

Evidence of validity based on internal structure

Data collection

After gaining ethical approval at all three institutions, the PAFI was distributed using an online link to three universities in the UK. 190 participants provided complete responses to the instrument. Upon clicking the online link, the participants were presented with a consent form outlining the purpose of this study and their rights concerning their data. Upon their consent to complete the PAFI, participants were first shown demographic questions. 60% of participants were female, attending university as a domestic student (83%), and were from White Caucasian background (78%). Most participants reported they had prior groupwork experience (78%), with 54% of them agreeing to have had positive groupwork experiences. However, 53% reported to prefer individual work over groupwork. 75% of participants had their groups formed by their tutors, with 10% reporting other (i.e. random assignment, some members were known and some not, choice based on topic of interest). The details of participants' demographic information are presented in Table 1.

Upon completion of demographic questions, students were presented with the 20 PAFI items. Participants were asked to read each item and identify the extent to which they agree with each item (1= Strongly Disagree, 2=Disagree, 3=Neutral, 4= Agree, 5= Strongly Agree).

Data analysis

We conducted exploratory factor analysis and reliability analysis to examine the internal structure of the PAFI. All the analyses were conducted in R (R Core Team 2020). We selected to use exploratory factor analysis rather than confirmatory factor analysis because this is an initial study to explore students' perceptions of peer assessment fairness in the context of their learning experience. Given that the items and underpinning constructs are borrowed from organizational justice literature and classroom provides another unique context (Pretsch and Ehrhardt-Madapathi 2018; Sabbagh and Resh 2016), the exploratory factor analysis seemed a more rational first step, providing initial evidence for the next studies with an intent to refine and consolidate the instrument with confirmatory factor analysis. Also, we had 190 participants for the PAFI items, resulting in the ratio of 9 respondents per item, which is considered adequate and less error-prone for conducting exploratory factor analysis (Osborne 2014).

The Kaiser-Meyer-Olkin value showed 0.94, and Bartlett's Test of Sphericity was also significant ($p=0.001$), indicating that the dataset is suitable for running exploratory factor analysis. Given the five-point Likert-scale nature of our data, the polychoric correlation was run to further examine the correlations across items (See Figure 1 below). Overall, the items showed acceptable correlations for proceeding with factor analysis.

Next, we conducted the parallel test to determine the number of factors and associated eigenvalues fitting with our data. Parallel analysis was run using maximum likelihood and polychoric correlations with 20 iterations. The results suggested three factors with eigenvalues (13.23, 0.89, 0.67) higher than simulated factors (0.83, 0.50, 0.42). The fourth factor showed an eigenvalue (0.26) less than the simulated factor (0.37). Therefore, we retained three factors corresponding with our conceptual model (i.e. distributive, procedural, and interactional justice). Please see Figure 2 for the results of the parallel analysis.

Table 1. Overview of participants' demography (*n* = 190).

	N (Percentage)
Gender	
Female	114 (60.0%)
Male	69 (36.3%)
Genderqueer	2 (1.1%)
Non-binary	2 (1.1%)
Prefer not to say	3 (1.6%)
Student Status	
Home student	159 (83.7%)
EU student	9 (4.7%)
International	22 (11.6%)
Ethnicity Background	
Asian or Asian British	21 (11.1%)
Black, Black British, Caribbean or African	7 (3.7%)
Mixed or multiple ethnic groups	10 (5.3%)
White	149 (78.4%)
Other ethnic group	1 (0.5%)
Prefer not to say	2 (1.1%)
Prior Groupwork Experience	
Yes	148 (77.9%)
No	42 (22.1%)
Positive Prior Group Experience	
Strong disagree	7 (3.7%)
Disagree	22 (11.6%)
Neutral	57 (30.2%)
Agree	78 (41.3%)
Strongly agree	25 (13.2%)
Choice of Group Task over Individual Task	
Strong disagree	36 (18.9%)
Disagree	65 (34.2%)
Neutral	48 (25.3%)
Agree	29 (15.3%)
Strongly agree	12 (6.3%)
Choice of Group Members	
Tutor Allocated	142 (74.7%)
Self-selected based on knowing them beforehand	32 (16.8%)
Self-selected without knowing them beforehand	6 (3.2%)
Other	10 (5.3%)

Exploratory factor analysis was then run with three factors using Maximum Likelihood method with oblimin rotation to interpret factor solutions. The choice of oblimin was motivated to allow intercorrelation between factors. Overall, three factors explained 70% of the variance in the data. The reliability evidence for each factor was calculated using both Omega (McDonald 1999) and coefficient alpha (Cronbach 1951). The coefficient alpha for each dimension was 0.78, 0.95, and 0.94. Given that the first dimension had only two items, we calculated the omega reliability index for dimension 2 (0.96) and 3 (0.95). Altogether, the PAFI enjoyed high-reliability evidence across both internal consistency metrics.

Findings

Following our conceptual mode, we found empirical support for the PAFI items to map onto distributive, procedural, and interactional justice factors (Table 2). The first factor, grade congruence, included two items focusing on students' perceptions of the fairness of grades (i.e. distributive justice) they received during groupwork. The second factor, performative group dynamics, represented nine items showing students' perceptions of peer assessment fairness if the group procedures (i.e. procedural justice) followed (a) allocated tasks properly and consistently, (b) were communicated timely, thoroughly, and frankly, (c) remained professional, and (d) were considered to be accurate with group member agreements. Please note that items 10 ('My group members remained professional in doing their groupwork properly at all times.') and 15 ('My group
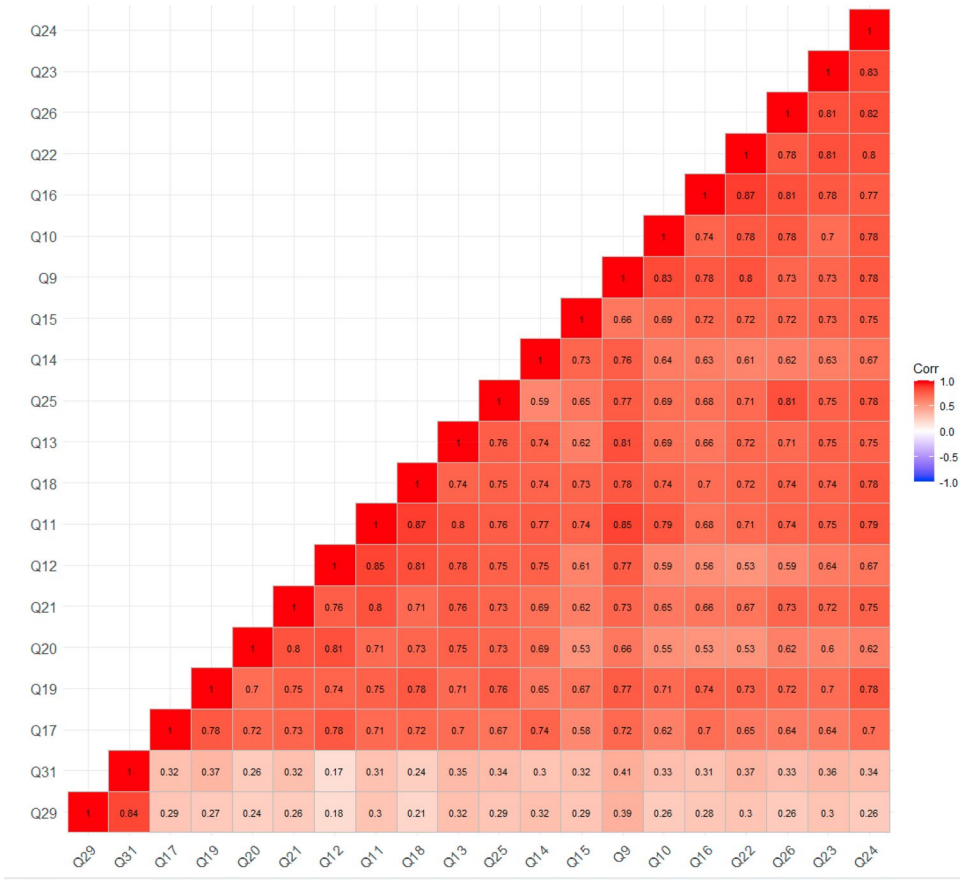
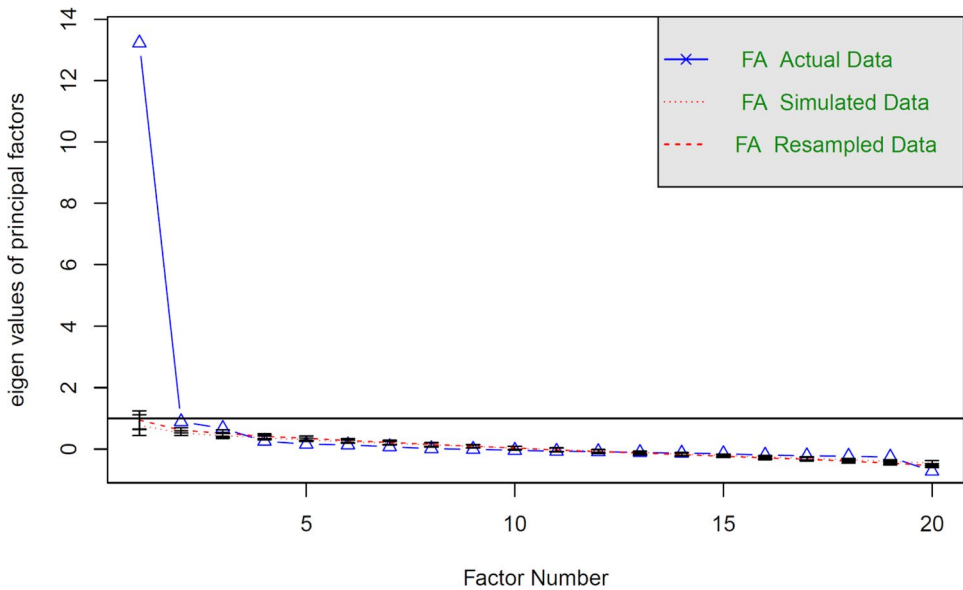**Figure 1.** Polychoric correlation for the PAFI items.

Figure 2. Parallel analysis scree plots.

Table 2. Standardized factor loadings for the PAFI items.

Items	ML 1 (Grade Congruence)	ML 2 (Performative Group Dynamics)	ML 3 (Interpersonal Treatment)	Item Complexity
Q29. The grade I received reflected the effort I put in.	**1.02**			1.0
Q31. The grade I received reflects what I contributed to the group work.	**0.78**			1.1
Q22. My group members did their allocated tasks properly.		**0.96**		1.0
Q16. In general, my group members consistently followed their allocated roles and schedules		**0.93**		1.0
Q26. My group members communicated details in a timely manner.		**0.85**		1.0
Q10. My group members remained professional in doing their groupwork properly at all times.		**0.73**		1.1
Q24. My group members explained the task decisions they make thoroughly.		**0.71**		1.2
Q23. My group members were open and frank in their communications with me.		**0.69**		1.1
Q15. My group members didn't exploit me.		**0.57**		2.1
Q19. My group members based decisions on accurate information.		**0.45**	0.39	2.0
Q25. I agreed with the decisions my group members made on the task.		**0.47**	0.39	1.9
Q12. I was able to express my views to my group members.			**1.01**	1.0
Q20. I could discuss and appeal the decisions that my group members made.			**0.79**	1.0
Q11. My group members treated me with respect.			**0.63**	1.4
Q17. Decisions that my group members made were not biased			**0.56**	1.3
Q14. I was treated the same as other group members.			**0.62**	1.2
Q18. My group members treated me in a polite manner.		0.31	**0.60**	1.5
Q13. My group members gave me recognition and praise		0.33	**0.53**	1.8
Q21. My group members were able to accommodate my specific needs.		0.33	**0.53**	1.7
Q9. My group members were empathetic and supportive.		0.48	**0.41**	2.1
Proportion of Explained Variance	**9%**	**34%**	**27%**	
Coefficient Alpha	**0.78**	**0.95**	**0.94**	
Omega reliability		**0.96**	**0.95**	

members didn't exploit me'.) were initially conceptualized to represent interactional justice. Finally, the third factor, interpersonal treatment, centred on students' perceptions of fairness of interactional justice if students were recognized, treated respectfully, politely, and empathetically, treated equally as others, accommodated based on needs, and were able to express views. Please note that items 12, ('I was able to express my views to my group members'.), 20 ('I could discuss and appeal the decisions that my group members made'.), 17 ('Decisions that my group members made were not biased'.), and 21 ('My group members were able to accommodate my specific needs'.) were initially conceptualized to reflect procedural justice.

Inter-factor correlations showed that grade congruence had $r = 0.34$ linkage with performative group dynamics, and $r = 0.25$ with interpersonal treatment. Performative group dynamics demonstrated $r = 0.69$ with interpersonal treatment.

The following histogram shows the frequency of students' responses to each option in the PAFI items (Figure 3). Students largely agreed that their group climate was distributively, procedurally, and interactionally fair. However, over 20 participants disagreed that their group members were professional, open, timely, thorough, frank, and empathetic in their communications. Also, over 20 participants disagreed that they received the grades they deserved.

Discussion

Fairness is empirically investigated as a significant factor influencing peer assessment within and beyond groupwork contexts (Panadero, Romero, and Strijbos 2013; Panadero et al. 2023). A recent systematic review identified 17 studies that used fairness as a factor to examine its relationship with peer assessment (Panadero et al. 2023). The findings were mixed with some studies showing students feeling peer assessment as a positive experience, with others feeling it negatively, and studies with students feeling both positive and negative (Carvalho 2013; Güler 2017; Wilson, Diao, and Huang 2015). The additional studies considered the conditions conducive for
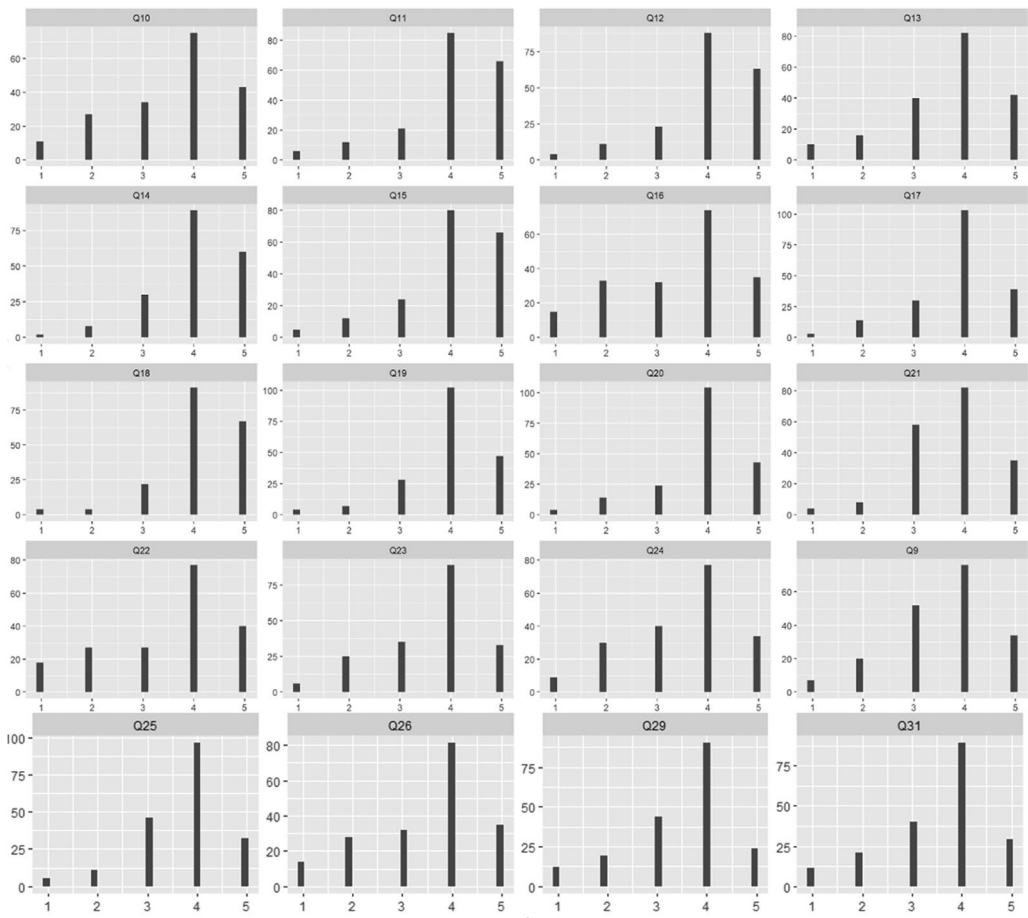
Figure 3. Frequency of students' responses to the PAFI items.

perceptions of fairness in peer assessment. Anonymity was found (Lin 2018) or not (Güler 2017) as contributing to perceptions of fairness. While peer feedback elaboration contributed to perceptions of fairness (Strijbos, Narciss, and Dünnebier 2010), the use of rubrics for assessment did not have an impact (Panadero, Romero and Strijbos, 2013). Across the studies reviewed in Panadero et al. (2023), fairness has been measured differently (see Table 3). While all studies used the word 'fairness' in the discussion sections, they measured fairness with different foci; one examining students' perceptions of fairness of peer assessment, another fairness of peer feedback, another fairness of peer grades, and yet another fairness as honesty.

Rasooli (2023a) has shown that the measurement of students' and teachers' perceptions of fairness in educational research is also varied with serious validity challenges. The peer assessment studies also used a single item to measure perceptions of fairness. The use of a single item for measuring perceptions of fairness would be useful in terms of research feasibility and advancement of peer assessment theory using statistical machinery. However, given the complexity of students' perceptions of fairness, it would be difficult to discern what drives a student's agreement with an item such as 'my peer assessment is fair'. Fair in terms of consistency, accuracy, giving voice, etc. or a combination of them? Without outlining such underpinning meanings in students' item responses, not only the validity of the single items is at stake (Embretson 1983), but also the potential translation of their findings into practical strategies for preparing educators would be undermined. Without delineating what fairness means to students when responding to an item, the success of future research to address the following direction stated by Panadero et al. (2023) will also be undermined: 'The challenge is then to explore the

Table 3. Studies and related items focusing on fairness as a variable in peer assessment review (Panadero et al. 2023).

Studies	Items
Zou et al. 2017	'I doubt the fairness of peer review'
Carvalho 2013	6. My team's final mark was fair; 7. The peer assessment in my team was adjusted
Panadero, Romero, and Strijbos 2013	"Do you believe that your peer will conduct a fair assessment of your concept map?"
Li and Steckelberg 2004	7. The peers' comments on my work were fair; 8. Peers can assess fairly.
Güler 2017	used three items from (Vanderhoven et al. 2015); Assessment by means of [mode of peer assessment] happened in an honest way; Assessment by means of [mode of peer assessment] happened in a correct way; Assessment by means of [mode of peer assessment] happened in a subjective way.
Seifert and Feliks 2019	To what extent did you think that your peers' comments were fair?; To what extent do you think that you assessed your peers' works fairly?
Vander Schee and Birrittella 2021	I believe the peer group grading process was fair
Wilson, Diao, and Huang 2015	One item measuring student perception of fairness (not reported in the text).
Lin 2018	"Peer comments I received from peer assessment were not fair enough."
Wen and Tsai 2006	24. I think PA is fair to assess students' performance; 36. Online PA activities are fair when assessing students' performance
Strijbos, Narciss, and Dünnebier 2010	I would be satisfied with this feedback; I would consider this feedback fair; I would consider this feedback justified

conditions under which students' perceptions of fairness and comfort would be negatively harmed, and what peer assessment design elements can alleviate this effect' (p. 1069). It is therefore significant to define peer assessment characteristics as well as meanings of fairness for valid conclusions to inform positive peer assessment practices.

To contribute to such research directions, it is beneficial to expand the current definitions and conceptualizations used in the measurement of fairness in peer assessment. Panadero et al. (2023) synthesized the definition of fairness in their review: 'students' perceptions that the PA [peer assessment] is free from bias, dishonesty or injustice' (1057). The definitions of fairness in educational assessment have been identified to be vague (Tierney 2013; Rasooli, Zandi, and DeLuca, 2019). Recently, Ion, Díaz-Vicario, and Mercader (2023) drew on the literature synthesis of fairness in the health profession (Valentine et al. 2021) to conceptualize and measure students' perceptions of fairness in peer assessment. The survey included four factors and 19 items: (a) fairness in human judgement and credibility; (b) multiplicity of opportunities and assessors; (c) procedural fairness and transparency; and (d) fitness for purpose. The study did not report collection and analysis of validity evidence for the survey and its four-factorial dimensionality.

In this study, we developed a novel instrument (i.e. PAFI) for measuring fairness in peer assessment contexts. We examined the validity evidence for the PAFI based on social psychology theory, students' response processes, and statistical analysis of internal structure. Overall, the initial evidence reported in this study contributes sound validity for the PAFI in measuring students' perceptions of fairness in assessment contexts. The PAFI drew on organizational justice theory and surveys (i.e. an application of social psychology theory of justice in organizational contexts) to examine students' perceptions of fairness alongside three dimensions: distributive, procedural, and interactional justice. The collated validity evidence supports that these three dimensions account for students' responses to the PAFI.

First factor: grade congruence

Distributive justice (i.e. grade congruence) is interpreted in this study in relation to grading outcomes given our focus on classroom assessment, e.g. received grade reflecting the contribution and effort of students. Similarly, Fortin et al. (2020) identified distributive justice items focusing on outcome distributions. However, they presented distributive justice items with more open-ended 'outcomes' (in a parenthesis) due to a variety of valued outcomes in work contexts ('Is your (outcome) appropriate for the work you have completed?'). That said, the loading of the recognition item on distributive justice in Fortin et al. could also be viewed as a classroom

assessment outcome, but students in this study interpreted that item in relation to interactional justice (i.e. interpersonal treatment).

Second factor: performative group dynamics

Parallel with the second factor of this study (i.e. performative group dynamics), Fortin et al. (2020) showed that the task justice factor (procedural justice) included items examining accurate information (item 19), consistent procedure (item 16), professionalism in groupwork (item 10), proper conduct of allocated tasks (item 22), timely and detailed communication (item 26), frankness in communication (item 23), reasonable decision-making (item 24), and thorough explanation (item 25). However, item 15 (i.e. 'my group members did not exploit me') from the performative group dynamics factor in this study contributed to the relationship justice factor in Fortin et al.

Third factor: interpersonal treatment

Similar to the third factor of this study (i.e. interpersonal treatment), Fortin et al. (2020) showed that the relationship justice factor included items focusing on respect, politeness, equal treatment, empathy. While items 12 on student voice, 17 on unbiased treatment, and 21 on accommodating needs loaded on distributive justice in Fortin et al.'s, they loaded on the interpersonal treatment factor in our study. Notably, voice and unbiased treatment were identified as procedural justice items within peer justice climate survey (Cropanzano, Li and Benson, 2011; Li, Cropanzano, and Bagger 2013; Li and Cropanzano 2009).

Overall, these results show similarities and differences within and across organizational and classroom assessment contexts. While continued studies are needed to make some reliable conclusions, it seems reasonable to argue that the similarity of results across these two contexts signifies preparing students in colleges and universities with fairer mechanisms for group work assessment that they can also employ within the future workplace spaces. Groupwork is a crucial aspect of work contexts and a key twenty first century skill (Maxwell 2023). We can leverage the three factors in this study and related items as a basis to prepare students with skills that they can implement in their future groupwork tasks. The findings of this study coupled with Fortin et al. (2020) shows that procedural justice (i.e. performative group dynamics) is the strongest factor, followed by interactional justice (i.e. interpersonal treatment) and distributive justice (i.e. grade congruence) in perceptions of fairness in peer assessment contexts. This finding implies that fairer groupworks need the members to do their fair share in following procedures in completing the tasks and projects so that everyone gets their deserved grades. Equally important, members need to work in groups that are safe relationally with interpersonal warmth, empathy, and recognition. It is hard to imagine one can continue positive contributions in groups that they do not find relationally fair (i.e. respectful, recognized, polite, empathetic and supportive, accommodating needs) and where procedures are not followed professionally and properly.

While grade outcomes (distributive justice) seem to give the impression as to be the most significant contributor to fairness of peer assessment, the findings of this study highlight the more significant influences of procedural and interactional justice dimensions. This means that a student may still get the best mark but feel unfairness due to procedural and interactional justice principles. More research is needed to investigate the respective influence of the three justice dimensions on students' perceptions of fairness in peer assessment. Further, existing research has focused heavily on peer assessment as a separate episode worthy of investigation; this focus may artificially separate the relational aspects of groupwork activities that happen naturally during a course. This is also noted by Panadero et al. (2023, 1067): 'the current findings do not clarify the role of interdependence in peer assessment since this factor relies on the assumption that peer assessment is an interactive activity which is still not the case in the

majority of peer assessment studies'. In recognition of the procedural and interactional factors for groupwork activities (where peer assessment is a part of), Fortin et al. (2020) suggest offering the empowerment of employees to 'set priorities and build their own systems of justice' (27) to ensure fairness in a diverse environment. This is akin to educators building capacity for students by explicitly teaching about interactional and procedural rules for operating fairly at the start of their groupwork activities. The PAFI can be used for teaching about as well as evaluation of fairness in peer assessment.

Implications for practice

Building on these insights and findings of this study, it seems in line with student assessment literacy (Hannigan, Alonzo, and Oo 2022) and student voice (Elwood and Hanna 2023) to prepare students with fairer skills for peer assessment and group work in response to the needs of work contexts. An explicit education with students about fairness in groupwork activities will support this agenda. The significant element of such education should also include perceptions of fairness in groupwork as engendered by cultural backgrounds. With the growing rise of classrooms and work contexts with group members of various cultural backgrounds, research needs to investigate the interactive spaces where cultures and values underpinning groupwork might be in harmony or disharmony. Such values will inevitably affect student assessment of fairness in peer assessments, group dynamics, interpersonal relationships, and students' willingness to work together.

Limitations

This study is not without limitations. For instance, it was only possible to sample three HEIs in the UK. Future research should continue validating the Peer Assessment Fairness Instrument (PAFI) with additional collection of evidence. Specifically, further research needs to leverage confirmatory factor analysis as well as measurement invariance to provide evidence of the stability of the PAFI within and across cultural contexts. It would be insightful to explore the impact of cultural groups and profiles on students' perceptions of peer assessment fairness. Qualitative research is also needed to explore the content validity of the survey as the items have been borrowed from organizational contexts. The evidence from qualitative data would provide avenues for potential examination of the adequacy and appropriacy of the current dimensions and items for the peer assessment contexts.

Conclusion

This study intended to add to understanding fairness of assessment from peers' lens as significant contributors to assessment contexts alongside teachers. Despite the growing focus on teachers' perceptions of fairness (Tierney 2013), and students' perceptions of teacher assessment including fairness (Brown 2008), more explicit attention on students' perceptions of fairness in peer assessment is needed. This study constructed a novel instrument to measure students' perceptions of fairness in peer assessment contexts. The results showed that students' perceptions of fairness in peer assessment contexts focused on grade congruence, performative group dynamics, and interpersonal treatment. While grade congruence is the first thing that would come to mind in examining fairness of peer assessment, students in this study showed that they also attend to the procedural and relational aspects of group work in judging the fairness of peer assessments. The instrument in this study and the findings can thus potentially be used as frameworks and tools to guide training educators and students for fairer group work and peer assessment practices. The emphasis in such training may need attention to whether peer grades and

feedback are distributed fairly, whether procedures for productive and professional group work assessment are followed fairly, and whether interpersonal relationships are valued during the group work and peer assessment processes.

Acknowledgments

Disclosure statement

Notes on contributors

Amirhossein Rasooli (PhD) currently holds the lecturer position in Educational Assessment at the Learning Sciences and Assessment Academic group, National Institute of Education, Nanyang Technological University. Amir's scholarship focuses on examining what notions of fairness and justice underpin assessment practices and what justice ideals assessment practices aim to propagate in society. Through his research, Amir draws on perceptions of fairness as a basis to improve assessment experiences and practices.

Jim Turner has over 20 years of experience driving innovation in digital education. His career began at International Centre of Digital Content, which was established in 2000 to explore the impact of digital technology on society and business. More recently, at Liverpool John Moores University, as a Senior Learning Technologist he has been instrumental in implementing learning technologies and supporting staff development to enhance teaching and learning.

Tünde Varga-Atkins is a Senior Educational Developer at the Centre for Innovation in Education, University of Liverpool, UK and a Principal Fellow of the Higher Education Academy. Her specific areas of research encompass areas in digital capabilities, signature pedagogies, curriculum design and evaluation, assessment and feedback, learner experience research and scholarship of learning and teaching. Tünde is editor of Research in Learning Technology.

Edd Pitt is the Programme Director for the Post Graduate Certificate in Higher Education and Reader in Higher Education and Academic Practice in the centre for the study of higher education at the University of Kent, UK. His principal research field is Assessment and Feedback with a particular focus upon student's use of feedback.

Shaghayegh Asgari specialises in blended learning and digital pedagogy. She has successfully developed and implemented innovative design strategies, enhancing online learner engagement through Technology-Enhanced Learning tools. Tya has led digital education projects and contributed to teaching policy development. Her research interests include the design and implementation of group work activities, where she focuses on innovative pedagogical approaches to improve educational practices.

Will Moindrot works as an Educational Developer for the Centre for Innovation in Education at the University of Liverpool, having worked within the area of educational technology at UK Higher Education Institutes for approx. 20 years with specific experience in supporting learning design and the adoption of new teaching technologies and pedagogies to support student engagement. He undertook his masters degree at the University of Manchester (Digital Technology, Communication and Education), and has several pieces of published work.

ORCID

Amirhossein Rasooli http://orcid.org/0000-0001-8888-8400
Jim Turner http://orcid.org/0000-0002-1823-6412
Tünde Varga-Atkins http://orcid.org/0000-0002-9492-4284
Edd Pitt http://orcid.org/0000-0002-7475-0299
Shaghayegh Asgari http://orcid.org/0000-0003-3497-2433
Will Moindrot http://orcid.org/0000-0001-7583-4833

References

Adams, J. S. 1965. "Inequity in Social Exchange." In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, Vol. 2, 267–299. New York: Academic Press.

Aggarwal, Praveen, and Connie L. O'Brien. 2008. "Social Loafing on Group Projects: Structural Antecedents and Effect on Student Satisfaction." *Journal of Marketing Education* 30 (3): 255–264. doi:10.1177/0273475308322283.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, and NCME). 2014. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Barsky, A. E., and L. Wood. 2005. "Conflict Avoidance in a University Context." *Higher Education Research & Development* 24 (3): 249–264. doi:10.1080/07294360500153984.

Bies, R. J., and J. S. Moag. 1986. "Interactional Justice: Communication Criteria of Fairness." In *Research on Negotiation in Organizations*, edited by R. J. Lewicki, B. H. Sheppard, and M. H. Bazerman, Vol. 1, 43–55. Greenwich, CT: JAI Press.

Brookhart, S. M. 2013. *Grading and Group Work: How Do I Assess Individual Learning When Students Work Together?* Alexandria, VA: ASCD.

Brown, G. T. L. 2008. *Conceptions of Assessment: Understanding What Assessment Means to Teachers and Students*. Hauppauge, NY: Nova Science Publishers.

Carvalho, A. 2013. "Students' Perceptions of Fairness in Peer Assessment: Evidence from a Problem-Based Learning Course." *Teaching in Higher Education* 18 (5): 491–505. doi:10.1080/13562517.2012.753051.

Chang, Y., and P. Brickman. 2018. "When Group Work Doesn't Work: Insights from Students." *CBE Life Sciences Education* 17 (3): ar42. doi:10.1187/cbe.17-09-0199.

Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika* 16 (3): 297–334. doi:10.1007/BF02310555.

Cropanzano, R., A. Li, and L. Benson. 2011. "Peer Justice and Teamwork Process." *Group & Organization Management* 36 (5): 567–596. doi:10.1177/1059601111417990.

Davies, W. M. 2009. "Groupwork as a Form of Assessment: Common Problems and Recommended Solutions." *Higher Education* 58 (4): 563–584. doi:10.1007/s10734-009-9216-y.

Dijkstra, J., M. Latijnhouwers, A. Norbart, and R. A. Tio. 2016. "Assessing the 'I' in Group Work Assessment: State of the Art and Recommendations for Practice." *Medical Teacher* 38 (7): 675–682. doi:10.3109/0142159X.2016.1170796.

Elwood, J., and A. Hanna. 2023. "Assessment Reform and Students' Voices." In *International Encyclopedia of Education (Fourth Edition)*, edited by R. J. Tierney, F. Rizvi, and K. Ercikan, 4th ed., 119–128. Cambridge, UK: Elsevier. doi:10.1016/B978-0-12-818630-5.09029-1.

Embretson, S. E. 1983. "Construct Validity: Construct Representation versus Nomothetic Span." *Psychological Bulletin* 93 (1): 179–197. doi:10.1037/0033-2909.93.1.179.

Forsell, J., K. Forslund Frykedal, and E. Hammar Chiriac. 2020. "Group Work Assessment: Assessing Social Skills at Group Level." *Small Group Research* 51 (1): 87–124. doi:10.1177/1046496419878269.

Fortin, M., R. Cropanzano, N. Cugueró-Escofet, T. Nadisic, and H. Van Wagoner. 2020. "How Do People Judge Fairness in Supervisor and Peer Relationships? Another Assessment of the Dimensions of Justice." *Human Relations* 73 (12): 1632–1663. doi:10.1177/0018726719875497.

Fung, D., V. Hung, and W-m Lui. 2018. "Enhancing Science Learning through the Introduction of Effective Group Work in Hong Kong Secondary Classrooms." *International Journal of Science and Mathematics Education* 16 (7): 1291–1314. doi:10.1007/s10763-017-9839-x.

Greenberg, J. 1987. "A Taxonomy of Organizational Justice Theories." *Academy of Management Review* 12 (1): 9–22. doi:10.5465/amr.1987.4306437.

Güler, Çetin. Ç 2017. "Use of WhatsApp in Higher Education: What's up with Assessing Peers Anonymously?" *Journal of Educational Computing Research* 55 (2): 272–289. doi:10.1177/0735633116667359.

Gweon, G., S. Jun, S. Finger, and C. P. Rosé. 2017. "Towards Effective Group Work Assessment: Even What You Don't See Can Bias You." *International Journal of Technology and Design Education* 27 (1): 165–180. doi:10.1007/s10798-015-9332-1.

Hall, D., and S. Buzwell. 2013. "The Problem of Free-Riding in Group Projects: Looking beyond Social Loafing as Reason for Non-Contribution." *Active Learning in Higher Education* 14 (1): 37–49. doi:10.1177/1469787412467123.

Hannigan, C., D. Alonzo, and C. Z. Oo. 2022. "Student Assessment Literacy: Indicators and Domains from the Literature." *Assessment in Education: Principles, Policy and Practice* 29 (4): 482–504. doi:10.1080/0969594X.2021.1879203.

Healy, M., M. McCutcheon, and J. Doran. 2014. "Student Views on Assessment Activities: Perspectives from Their Experience on an Undergraduate Programme." *Accounting Education* 23 (5): 467–482. doi:10.1080/09639284.2014.949802.

Ion, G., A. Díaz-Vicario, and C. Mercader. 2023. "Making Steps towards Improved Fairness in Group Work Assessment: The Role of Students' Self-and Peer-Assessment." *Active Learning in Higher Education*. doi:10.1177/14697874231154826.

Kwon, K., Y. H. Liu, and L. P. Johnson. 2014. "Group Regulation and Social-Emotional Interactions Observed in Computer Supported Collaborative Learning: Comparison between Good vs. Poor Collaborators." *Computers & Education* 78: 185–200. doi:10.1016/j.compedu.2014.06.004.

Leventhal, G. 1980. "What Should Be Done with Equity Theory? New Approaches to the Study of Justice in Social Relationships." In *Social Exchange: Advances in Theory and Research*, edited by K. Gergen, M. Greenberg, and R. Willis, Vol. 9, 27–55. New York: Plenum Press.

Li, A., and R. Cropanzano. 2009. "Fairness at the Group Level: Justice Climate and Intraunit Justice Climate." *Journal of Management* 35 (3): 564–599. doi:10.1177/0149206308330.

Li, A., R. Cropanzano, and J. Bagger. 2013. "Justice Climate and Peer Justice Climate: A Closer Look." *Small Group Research* 44 (5): 563–592. doi:10.1177/104649641349811.

Lin, G.-Y. 2018. "Anonymous versus Identified Peer Assessment via a Facebook-Based Learning Application." *Computers & Education* 116: 81–92. doi:10.1016/j.compedu.2017.08.010.

Li, L., and A. Steckelberg. 2004. "Using Peer Feedback to Enhance Student Meaningful Learning." Paper presented at the Conference of the Association for Educational Communications and Technology, Chicago, USA.

Loughry, M. L., M. W. Ohland, and D. D. Moore. 2007. "Development of a Theory-Based Assessment of Team Member Effectiveness." *Educational Psychological Measurement* 67: 505–524.

Maxwell, G. S. 2023. "Assessment and Accountability Global Trends and Future Directions in 21st Century Competencies." In *International Encyclopedia of Education (Fourth Edition)*, edited by R.J. Tierney, F. Rizvi, and K. Ercikan, 4th ed., 313–320. Cambridge MA. Elsevier. doi:10.1016/B978-0-12-818630-5.09069-2.

McDonald, R. P. 1999. *Test Theory: A Unified Treatment*. Mahwah, NJ: Taylor and Francis.

Meijer, H., R. Hoekstra, J. Brouwer, and J.-W. Strijbos. 2020. "Unfolding Collaborative Learning Assessment Literacy: A Reflection on Current Assessment Methods in Higher Education." *Assessment & Evaluation in Higher Education* 45 (8): 1222–1240. doi:10.1080/02602938.2020.1729696.

Mellor, T. 2012. "Group Work Assessment: Some Key Considerations in Developing Good Practice." *Planet* 25 (1): 16–20. doi:10.11120/plan.2012.00250016.

Nieminen, J. H., and L. Tuohilampi. 2020. "Finally Studying for Myself' – Examining Student Agency in Summative and Formative Self-Assessment Models." *Assessment & Evaluation in Higher Education* 45 (7): 1031–1045. doi:10.1080/02602938.2020.1720595.

Osborne, J. W. 2014. *Best Practices in Exploratory Factor Analysis*. Scotts Valley, CA: CreateSpace Independent Publishing Platform.

Panadero, E. 2016. "Is It Safe? Social, Interpersonal, and Human Effects of Peer Assessment: A Review and Future Directions." In *Handbook of Human and Social Conditions in Assessment*, edited by G. T. L. Brown and L. R. Harris, 247–266. New York: Routledge.

Panadero, E., M. Alqassab, J. Fernández Ruiz, and J. C. Ocampo. 2023. "A Systematic Review on Peer Assessment: Intrapersonal and Interpersonal Factors." *Assessment & Evaluation in Higher Education* 48 (8): 1053–1075. doi:10.1080/02602938.2023.2164884.

Panadero, E., M. Romero, and J. Strijbos. 2013. "The Impact of a Rubric and Friendship on Peer Assessment: Effects on Construct Validity, Performance, and Perceptions of Fairness and Comfort." *Studies in Educational Evaluation* 39 (4): 195–203. doi:10.1016/j.stueduc.2013.10.005.

Pellegrino, J. W., and M. L. Hilton, eds. 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC: The National Academies Press.

Pretsch, J., and N. Ehrhardt-Madapathi. 2018. "Experiences of Justice in School and Attitudes towards Democracy: A Matter of Social Exchange?" *Social Psychology of Education* 21 (3): 655–675. doi:10.1007/s11218-018-9435-0.

R Core Team. 2020. R: A Language and Environment for Statistical Computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rasooli, A., C. DeLuca, L. Cheng, and A. Mousavi. 2023b. "Classroom Assessment Fairness Inventory: A New Instrument to Support Perceived Fairness in Classroom Assessment." *Assessment in Education: Principles, Policy and Practice* 30 (5–6): 372–395. doi:10.1080/0969594X.2023.2255936.

Rasooli, A., H. Zandi, and C. DeLuca. 2018. "Re-Conceptualizing Classroom Assessment Fairness: A Systematic Meta-Ethnography of Assessment Literature and Beyond." *Studies in Educational Evaluation* 56: 164–181. doi:10.1016/j.stueduc.2017.12.008.

Rasooli, A., H. Zandi, and C. DeLuca. 2019. "Conceptualising Fairness in Classroom Assessment: Exploring the Value of Organisational Justice Theory." *Assessment in Education: Principles, Policy and Practice* 26 (5): 584–611. doi:10.1080/0969594X.2019.1593105.

Rasooli, A., H. Zandi, and C. DeLuca. 2023a. "Measuring Fairness and Justice in the Classroom: A Systematic Review of Instruments' Validity Evidence." *School Psychology Review* 52 (5): 639–664. doi:10.1080/2372966X.2021.2000843.

Sabbagh, C., and N. Resh. 2016. "Unfolding Justice Research in the Realm of Education." *Social Justice Research* 29 (1): 1–13. doi:10.1007/s11211-016-0262-1.

Seifert, T., and O. Feliks. 2019. "Online Self-Assessment and Peer-Assessment as a Tool to Enhance Student-Teachers' Assessment Skills." *Assessment & Evaluation in Higher Education* 44 (2): 169–185. doi:10.1080/02602938.2018.1487023.

Sonnleitner, P., and C. Kovacs. 2020. "Differences between Students' and Teachers' Fairness Perceptions: Exploring the Potential of a Self-Administered Questionnaire to Improve Teachers' Assessment Practices." *Frontiers in Education* 5: 1–17. doi:10.3389/feduc.2020.00017.

Strijbos, J.-W., S. Narciss, and K. Dünnebier. 2010. "Peer Feedback Content and Sender's Competence Level in Academic Writing Revision Tasks: Are They Critical for Feedback Perceptions and Efficiency?" *Learning and Instruction* 20 (4): 291–303. doi:10.1016/j.learninstruc.2009.08.008.

Svinicki, M. D., and D. L. Schallert. 2016. "Learning through Group Work in the College Classroom: Evaluating the Evidence from an Instructional Goal Perspective." In *Higher Education: Handbook of Theory and Research*, edited by M. B. Paulsen, Vol. 31, 513–558. New York: Springer.

Tierney, R. D. 2013. "Fairness in Classroom Assessment." In *SAGE Handbook of Research on Classroom Assessment*, edited by J. H. McMillan, 125–144. Thousand Oaks, CA: SAGE Publications.

Tumpa, Roksana Jahan, Samer Skaik, Miriam Ham, and Ghulam Chaudhry. 2022. "A Holistic Overview of Studies to Improve Group-Based Assessments in Higher Education: A Systematic Literature Review." *Sustainability* 14 (15): 9638. doi:10.3390/su14159638.

Valentine, Nyoli, Steven Durning, Ernst Michael Shanahan, and Lambert Schuwirth. 2021. "Fairness in Human Judgement in Assessment: A Hermeneutic Literature Review and Conceptual Framework." *Advances in Health Sciences Education: Theory and Practice* 26 (2): 713–738. doi:10.1007/s10459-020-09976-5.

Vander Schee, B. A., and T. D. Birrittella. 2021. "Hybrid and Online Peer Group Grading." *Marketing Education Review* 31 (4): 275–283. doi:10.1080/10528008.2021.1887746.

Vanderhoven, Ellen, Annelies Raes, Hannelore Montrieux, Tijs Rotsaert, and Tammy Schellens. 2015. "What If Pupils Can Assess Their Peers Anonymously? a Quasi-Experimental Study." *Computers & Education* 81: 123–132. 10.1016/j.compedu.2014.10.001.

Wen, M. L., and C.-C. Tsai. 2006. "University Students' Perceptions of and Attitudes toward (Online) Peer Assessment." *Higher Education* 51 (1): 27–44. doi:10.1007/s10734-004-6375-8.

Williams, E. 1992. "Student Attitudes towards Approaches to Learning and Assessment." *Assessment & Evaluation in Higher Education* 17 (1): 45–58. doi:10.1080/0260293920170105.

Williams, J. M., J. N. Cera Guy, and B. M. Shore. 2019. "High-Achieving Students' Expectations about What Happens in Classroom Group Work: A Review of Contributing Research." *Roeper Review* 41 (3): 156–165. doi:10.1080/02783193.2019.1622165.

Wilson, M. J., M. M. Diao, and L. Huang. 2015. "I'm Not Here to Learn How to Mark Someone Else's Stuff': An Investigation of an Online Peer-to-Peer Review Workshop Tool." *Assessment & Evaluation in Higher Education* 40 (1): 15–32. doi:10.1080/02602938.2014.881980.

Zhang, B., L. Johnston, and G. B. Kilic. 2008. "Assessing the Reliability of Self- and Peer Rating in Student Group Work." *Assessment & Evaluation in Higher Education* 33 (3): 329–340. doi:10.1080/02602930701293181.

Zhang, Bo., and Matthew W. Ohland. 2009. "How to Assign Individualized Scores on a Group Project: An Empirical Evaluation." *Applied Measurement in Education* 22 (3): 290–308. doi:10.1080/08957340902984075.

Zou, Y., C. D. Schunn, Y. Wang, and F. Zhang. 2017. "Student Attitudes That Predict Participation in Peer Assessment." *Assessment & Evaluation in Higher Education* 43 (5): 800–811. doi:10.1080/02602938.2017.1409872.

Appendix

A full version of the Peer Assessment Fairness Instrument (PAFI) is available online at this location dx.doi.org/10.6084/m9.figshare.25623771