



Developing a team effectiveness, peer-rating rubric for sports journalism ethics classes

Sada Reed
Arizona State University

Abstract

Though scholars have concluded small-group learning has multiple benefits for students, many students strongly dislike working in groups. They resent when everyone in a group receives the same grade, regardless of their individual contribution. According to Social interdependence theory, groups' structure determines (a) how individuals interact with each other, and (b) the type of outcomes individuals generate (Johnson & Johnson, 2005). Cooperative learning studies have found that students more conscientiously participate when they know their peers will assess them. This has resulted in numerous measures for forming and assessing types of teams. Through the lens of social interdependence theory, this study combines Oakley et al.'s (2004) Team Member Evaluation Measures, Felder and Brent's (2007) self-assessment of team function guidelines, and Brown's (1995) autorating system into one framework for creating and supporting effective teams and showing the role substitutability, cathexis, and inducibility (Deutsch, 1949, 1962) have on effective team membership. These measures were piloted in three sections of an Ethics in Sports Media class at a southwestern American university. Results showed this study's substitutability, cathexis, and inducibility measures predicted about 80% of students' experience working in groups, and about 70% of how peers assessed each other. These measures, however, could not predict the team case study grade groups received.

The Spring 2022, Fall 2023, and Fall 2025 semesters marked the 11th, 12th, and 13th times, respectively, this study's author taught Ethics in Sports Media at a southwestern American university. Prior to these semesters, the author noted multiple instances when small groups presented high-quality case studies, but members later reported negative experiences working with their groups. Someone had often done most of the project by themselves, yet everyone came away with the same score. By allowing "free riders," or students who make little to no contribution to the group

project score, earn the same grade as the team's main contributors, the instructor/author was essentially rewarding academic misconduct (Dix, 2024). Educators must help students create and maintain effective groups, and give students the ability to hold each other accountable.

In response to Halfhill et al.'s (2005) and Levine and Moreland's (2012) call for more research devoted to assessing small group effectiveness, this study combines Oakley et al.'s (2004) Team Member Evaluation measures, Felder and Brent's (2007) self-assess-

ment of team function guidelines, and Brown's (1995) autorating system into one framework for creating, supporting, and grading group projects. These measures were also used to assess the role substitutability, cathexis, and inducibility have on students' experiences working in groups, and how they assess each other's contributions. These measures were piloted during the Spring 2022, Fall 2023, and Fall 2025 sections of Ethics in Sports Media class, where students interpret ethical issues through the lens of the Society for Professional Journalists' (SPJ) Code of Ethics. The semester's summative assessment is to research an example of something a sports journalist or sports media organization did that students considered unethical. In small groups, students examine these situations through the lens of the SPJ principles and present the case study to the class. Oakley et al.'s (2004) Team Member Evaluation measures and Felder and Brent's (2007) self-assessment of team function guidelines were used to structure groups and measure substitutability, cathexis, and inducibility. Students used Brown's (1995) autorating system to assess each other's contributions to the group.

The following paper contains a review of relevant social interdependence theory, cooperative learning, and autorating systems literature. Research questions, methods, and findings sections follow, concluding with a discussion section.

Literature Review

Though there are four major theories that have been used to conceptualize cooperative learning, social interdependence has generated the most research and procedures (Johnson & Johnson, 2015). Social interdependence theory asserts that individual's outcomes depend on both their actions and the actions of their group members (Johnson & Johnson, 1989). Morton

Deutsch (1949) published the theory, but its conceptual roots hail from early 20th-century Gestalt Psychology. Gestalt Psychology posits that the human mind perceives experiences as organized wholes, instead of collections of individual parts. Koffka (1935), for example, theorized that groups were wholes onto themselves, with members having various degrees of interdependence. Lewin (1948) developed the concept further, proposing that a group's interdependence is essentially what makes a group a group. Changing one group member influences the other group members, Lewin (1948) said, and group members experi-

ence an "intrinsic state of tension" that motivates them to accomplish the group's common goals (Johnson & Johnson, 2015, p. 7).

After serving in the U.S. Air Force as a psychologist and navigator during World War II, Deutsch studied at the Massachusetts Institute of Technology (MIT) under Lewin as part of Lewin's Research Center for Group Dynamics. Deutsch (1949) went on to develop the theory of cooperation and competition, which states that positive interdependence exists when individuals' goals positively correlate, and individuals perceive that reaching their own goals depends on other group members reaching their goals, too. Positive interactions occur "when members help and assist each other, exchange resources, give and receive feedback, challenge each other's reasoning, and encourage increased effort" (Johnson & Johnson, 2015, p. 7).

Psychological processes: Substitutability, cathexis, and inducibility

Positive interdependency influences group members' psychological processes (Deutsch, 1949, 1962). There are three processes. The first is substitutability. Substitutability is the extent to which one person's actions can be substituted for another (Johnson & Johnson, 2015). The second process is cathexis. According to the Chicago Psychoanalytic Institute (2024, para. 1), cathexis (kuh-THEK-suhs) is "the mental energy a person invests in themselves, others, or objects. ... it explains how we become attached to people, ideas, or goals, and how that investment impacts our emotional well-being." Measuring cathexis requires assessing group members' perception of interpersonal support, encouragement, and affective relationship quality. The third variable is inducibility, or a person's openness to influence. This openness includes a person's willingness to accept input from others and to adjust their behavior accordingly (Deutsch, 1949).

In the current study, Oakley et al.'s (2004) Team Member Evaluation measures and Felder and Brent's (2007) self-assessment of team function guidelines will be used to structure groups and to measure substitutability, cathexis, and inducibility. A composite score, based on internal validity of substitutability, cathexis, and inducibility, will be used to represent these three concepts in a multiple linear regression measuring how these concepts predict students' group experience.

Cooperative Learning

According to Johnson and Johnson (1998), cooperative learning involves students working in teams to accomplish a common goal that includes positive social interdependence, individual accountability, face-to-face promotive interaction, appropriate use of collaborative skills, and group processing. Scholars began applying social interdependence to U.S. educational settings in the 1980s. Johnson et al. (1981), for example, published a meta-analysis on cooperative, competitive, and individualistic goal structures' effects on student achievement and productivity. The authors confirmed Gillies' (2016) findings: that cooperation is more effective on individual achievement and productivity than interpersonal competition and individualistic goals. Compared to teaching through lectures and individual assignments, cooperatively taught students, according to Felder and Brent (2009):

Tend to exhibit higher academic achievement, greater persistence through graduation, better high-level reasoning and critical thinking skills, deeper understanding of learned material, greater time on task and less disruptive behavior in class, lower levels of anxiety and stress, greater intrinsic motivation to learn and achieve, greater ability to view situations from others' perspectives, more positive and supportive relationships with peers, more positive attitudes toward subject areas, and higher self-esteem. (p. 1)

Struggling students are less likely to give up, and successful students find gaps in their own knowledge by having to explain concepts to other students (Felder & Brent, 2009). Overall, small-group learning can result in longer information retention, deeper levels of understanding, and enhanced communication skills among student participants (Springer et al. 1999; Terenzini et al., 2001).

These benefits, however, are not automatic. In fact, many students strongly dislike working in groups. Compared to other assessment options, students rate group work lowest in accuracy and fairness (Neto et al., 2023).

Brown autorating system

Cooperative learning environment scholarship has produced numerous measures geared toward forming and assessing specific types of teams, such as manufacturing, advertising, or sports (Ancona, 1990; Ancona & Caldwell, 1992a, 1992b; Cohen & Bailey,

1997). Despite being applied in different industries, these measures are alike in that they recognize the importance of context, relationships, and external environment on team success. For example, students more conscientiously participate when they know their peers will assess them (Brown, 1995; Felder & Brent, 2009). Peer reviewing enhances students' sense of ownership and responsibility, requiring them to assess how their own thinking and behavior contributes to a team's goal (Lax & Fentiman, 2016; Nicol & MacFarlane-Dick, 2009; Ocampo et al., 2023). When trained to assess their own work and that of their classmates, students' understanding of a project's learning outcomes have just as good, if not better, effects on students' development than an instructor's assessment (Lax & Fentiman, 2016; Sharp & Dyrud, 2009).

Brown (1995) proposed a system for assigning an individual score for each team member, based on their group members' assessment of their contribution and that of each team member. For example, there are four people in a group, and the instructor gave the group a team score of 80%. If one student awarded themselves 87.5%, and their teammates gave them votes of 100%, 87.5%, and 87.5%, the instructor would take the average of these four votes (90.62%) and divide it by the team score (in this instance, 80%) to calculate what Brown (1995) called an adjustment factor. For this hypothetical student, $90.62/80 = 1.13$. Then, this adjustment factor is subtracted from the average vote to calculate the student's individual grade. In this example, $90.62 - 1.13 = 89.23\%$. So, though the team received 80% for its project, this student would receive 89.23%. This same process is repeated for all team members.

Research Questions

The instructor/author adapted Oakley et al.'s (2004) Team Member Evaluation measures,

Felder and Brent's (2007) self-assessment of team function guidelines, and Brown's (1995) autorating system to create one tool for forming and supporting effective groups, for allowing group members to assess each other, and for determining the extent to which substitutability, cathexis, and inducibility predict student outcomes. Combining social interdependence theory with effective team-building and peer-review assessment literature shows the extent to which students' perception of substitutability, cathexis, and inducibility influence team outcomes. This is import-

ant because many students are resistant to working in groups, citing negative past experiences, free riders, and unfair labor divisions (Wong et al., 2022). This resistance can undermine cooperative learning. To determine students’ experience working in this group, the following research question was proposed:

RQ1: To what extent do substitutability, cathexis, and inducibility predict students’ overall experiences working in groups?

This RQ was explored through descriptive statistics and a multiple linear regression based on a questionnaire adapted from Oakley et al. (2004) that students completed at the end of the semester, after they presented their case studies.

Cooperative learning is effective in higher education (Smith et al., 2005; Springer et al., 1999) but group work does not always reflect students’ contributions (Khuzwayo, 2018; Willmot & Crawford, 2007; Zhang & Ohland, 2009). Cooperative learning entails students being held individually accountable for the complete content of an assignment (Felder & Brent, 2009). When students are not held accountable, they can become free riders. Instructors are not necessarily able to identify free riders by observation alone (Willmot & Crawford, 2007). This means students’ peer assessment of one another’s contributions is a crucial component of team effectiveness (Durdubas & Koruc, 2023; Kim & Kim, 2012; Tassi et al., 2023). To determine the degree to which Brown’s (1995) autorating system allowed students to assess each other and, in turn, hold each other accountable, the following RQ is proposed:

RQ2: What is the difference between the team score individuals earned and the peer review score they received from their group members?

This RQ was examined through descriptive statistics, an ANOVA, t-tests, and Brown’s (1995) autorating system.

According to social interdependence theory research, interdependence is measured through substitutability, cathexis, and inducibility. To determine the extent to which these variables can predict students’ group and individual outcomes, the following RQs are proposed:

RQ3a: How do substitutability, cathexis, and inducibility predict groups’ team scores?

RQ3b: How do substitutability, cathexis, and inducibility predict peer-given scores?

These RQs will be explored through a multilinear regression based on a questionnaire adapted from Oakley et al. (2004)’s Team Member Evaluation Measures, Felder and Brent’s (2007) self-assessment of team function guidelines, and Brown’s (1995) autorating system. Students completed these questionnaires after their first- and second in-class meetings, and after they presented their case studies.

Method

Ethics in Sports Journalism is a three-credit, in-person class for sports journalism undergraduate majors and masters students. For the Spring 2022 undergraduate and fall 2025 masters classes, the author/instructor put students in three-person groups. For the Fall 2023 undergraduate class, students self-selected into pairs. This was done to explore differences between how two- and three-person groups responded to the adapted measures. Regardless of group size, students chose which case study they wanted and when they wanted to present. Groups could choose their own case study or select one from a list of 28 case studies the instructor/author provided. The rubric is included in Appendix A.

Measures

After teams formed, groups met in class to have their first team meeting. During this meeting, teams completed an open-ended questionnaire adapted from Oakley et al. (2004).

- Team mission. What is the reason for this group being together?
- Team goals. What are some specific team goals or indicators of performance that we are striving to obtain for this project?
- Action plan. What do we need to do as a team and as individuals to reach that goal?
- Strengths. What are our strengths as a group, and how can we best utilize these strengths to be successful? For example, what project-specific skills do each person bring that, in combination, will benefit us (e.g., good at organizing; public speaking; building PowerPoints, podcasts, videos, or other media; research)?

- Areas for improvement. What do I/we need to improve to be successful (e.g., tendency to procrastinate, to not communicate, to overwork without asking for help)?
- Team support. What can I do to help support my team?
- Team check-in. How often will I check in with my group to let them know my progress? By what means will I do this check-in (e.g., e-mail, text, in-person)?
- Interpersonal conflict. When conflicts arise within our group, what will I/we do to resolve the issues?

Two weeks later, teams met again during class, completing a questionnaire adapted from Felder and Brent (2007) asking them to rate their level of agreement on a 5-point Likert scale (1 meaning *strongly disagree* and 5 meaning *strongly agree*) on the following statements:

- Our discussion was productive.
- We accomplished what I wanted us to accomplish.
- Everyone performed their role effectively.
- We communicated well with each other.
- We supported each other well.
- What does the group need to do to be successful? (Write in.)

Social Interdependency Theory

After completing their presentation, students completed the third and final questionnaire adapted from Oakley et al. (2004). They were asked to rate their level of agreement on a 5-point Likert scale (1 being strongly disagree, 5 being strongly agree) on a series of statements that referred to either their group as a whole or individual team members. For three-person groups, students repeated the statements geared toward specific individuals twice – once for each team member. This study is unique in that these exit surveys also measured substitutability, inducibility, and cathexis. Though these three concepts are interrelated, they were coded based on their dominant conceptual focus, which is consistent with prior research on social interdependence theory (Johnson et al., 1998). For example, “We communicated well with each other” supports accomplishing tasks and group members. It was coded as inducibility because it reflects openness to influence and responsiveness (Deutsch, 1949; Johnson & Johnson, 1989).

The statements, adapted from Oakley et al.

(2004), are listed by whether they were measuring substitutional, inducibility, or cathexis. Their Cronbach alphas, which indicate measures’ internal consistency, are also listed:

Substitutability ($\alpha = .850$).

- One person usually did most of the talking.
- Each member did what they were supposed to do to prepare for the presentation.
- This person attended team meetings, either in person or virtually.
- This person made a serious effort to do the work we agreed upon between team meetings.
- This person made an effort to fulfill their team role responsibilities.
- This person attempted to contribute to group meetings.

Cathexis ($\alpha = .719$).

- The team used its time together wisely.
- We were prepared for our presentation.
- I enjoyed working with this group.

Inducibility ($\alpha = .812$).

- Team members understood what they wanted to accomplish.
- Team members often interrupted each other without listening to the person who had the floor.
- Any issues we had were resolved.
- This person notified a teammate if they could not attend a meeting or fulfill a responsibility.
- This person listened to teammates ideas’ and opinions respectfully and gave them careful consideration.
- This person cooperated with the group.

Students were also asked whether they had any other comments or concerns they wanted to share with the instructor/author, and a text box in which they could record their comment or concern was included. They repeated this process for each member of their team. For the sake of privacy, students are identified by letters in this study (e.g., Student A, Student B). The study received exemption from the author/instructor’s university’s Institutional Review Board, STUDY00003655.

Findings

There were 104 students in the study. Overall, 43 people (41.7%) were second-year students, 38 (36.9%) were masters students, 18 (17.5%) were third-year students, and 4 (3.9%) were fourth-year students. There were 60 students spread between 20

Table 1. Regression Coefficients for Predicting Students’ Experience Working in Group

Variable	B	SE B	B	t	p
Substitutability	.006	.005	.064	1.365	.175
Cathexis	.294	.018	.865	16.178	<.001
Inducibility	.016	.026	.034	.623	.534

Note. R² = .795, Adjusted R² = .789, F(3,97) = 125.488, p < .001

three-person groups. There were 44 students spread between 22 pairs. The Spring 2022 and Fall 2025 classes were predominantly three-person groups, with the exception of a two-person group in each section. The fall 2023 class was all pairs, with the exception of one three-person group. Overall, there were 77 male students (74%) and 27 (26%) female students. Chi-square test results showed statistically significant differences between year in school and the three classes, $\chi^2(6, N = 103) = 135.158, p < .001$. The majority of students in the Spring 2022 group were third-year students (13 of 23 students), most in the Fall 2023 class were second-year students (36 of 42 students), and all 38 of the students in the Fall 2025 class were masters students. Two students dropped out at the end of the term and did not take the final exit survey.

RQ1 asked to what extent substitutability, cathexis, and inducibility predict students’ overall experiences working in groups. On a scale of 1 to 5, with 5 being *strongly agree* and 1 being *strongly disagree*, students’ average for the adapted Oakley et al.’s (2004) Team Member Evaluation measures was 3.72 (SD = .299). There were no statistically significant differences between male and female students’ answers, or between semesters. The statements intended to measure substitutability, cathexis, and inducibility each had acceptable alphas, so a composite score was made for each concept based on students’ average scores for said concept. A multiple linear regression was used to test how substitutability, cathexis, and inducibility predicted students’ experiences working in groups. The results of an ANOVA were statistically significant, $F(3,97) = 125.48, p < .001$. The R² was .795, indicating that these measures could explain about 80% of the variance in students’ experience working in groups. Cathexis (i.e., “The team used its time together wisely,” “We were prepared for our presentation,” “I enjoyed working with this group”) made a statistically

significant contribution to the model.

However, there were statistically significant differences between semester, and levels of substitutability, cathexis, and inducibility. The Fall 2025 masters student class had a statistically significant higher substitutability average (M = 5.32, SD = 3.62) than the Fall 2023 group (M = 4.26, SD = .571), and significantly higher cathexis (M = 4.75, SD = .537) and inducibility (M = 4.89, SD = .388), respectively, than the Spring 2022 group (M = 4.39, SD = .630, and M = 4.60, SD = .583, respectively). There were no statistically significant differences between the fall 2023 and spring 2022 undergraduate groups.

RQ2 asked what the difference was between the team score individuals earned and the peer-review score they received from their group members. Overall, there was a statistically significant difference between the team scores groups received and semester, $F(2,100) = 4.524, p = .013$. Students in the fall 2023 course (M = 89.52%, SD = 19.58) produced statistically significant lower quality case studies than students in the spring 2022 (M = 97.74%, SD = 5.96) and fall 2025 (M = 97.89%, SD = 8.19) courses. A t-test found no statistically significant difference between students’ team scores (M = 94.44%, SD = 14.25), and the score students earned from their peers (M = 94.37, SD = 9.45), $t(101) = -.678, p = .499$. However, results of an ANOVA showed this varied significantly depending upon the semester, $F(2,100) = 3.69, p = .028$. Students in the Spring 2022 class earned peer-assessed scores that were 6.52 points lower than the group case study grade their team earned. For the Fall 2023 and 2025 classes, however, individuals’ peer assessed scores were 2.98 points and 1.82 points higher, respectively, than the team scores they earned. For example, Students G, N, and Q were in a group together during the Spring 2022 semester. As a group, the team received a perfect 100% on their case study. But in the exit question-

naire, Student G gave below-average ratings on several measures, saying:

[Student N] was engaged the whole time. [Student Q] only seemed to care when the project was coming up and did not want to engage until we said he had done nothing. He seemed more keen on proving to us he was not lazy rather than actually helping the team.

Student N strongly agreed with Student G on the measures, but did not give comments to contextualize the experience as Student G did. The average peer-assessed scores they gave Student G was 2.42 and 2.14 on a 5-point scale. This resulted in Student G receiving a final score of 45%. Without this peer-assessed score, Students G, N, and Q would have all received 100%, regardless of contribution.

In another example, Students R, K, and E were in a group together during the spring 2022 semester. They earned a team grade of 82.69% after choosing their own topic, but it was one not specific to a sports journalist’s or sports media outlet’s ethical mistake, as was required. This led to the group missing points throughout the presentation. In their exit questionnaire, Student R said Student E “hardly did anything for the project. He kind of just rode our coattails. In all honesty, [Student K] did the majority of the work.” Student R went on to say that Student K wrote the talking points and conducted most of the research, and Student R added graphics and edited. But Student E rarely opened the PowerPoint, Student R said – adding that when confronted about it, Student E told Students R and K he was “too busy.” Student R found this to be an unacceptable excuse because they, as students, are all busy. “Every time we were doing a Zoom or talking about [the project], he was either not there or checked out,” Student R said. This was especially damaging to the group because when they assigned tasks in the Week 1 team meeting (Oakley et al., 2004), Student E was going to be responsible for the PowerPoint. In the Week 3 check-in, Student E said “work on the slideshow” as the top thing the group needed to do to be successful (Felder & Brent, 2007).

For the Fall 2023 group, students’ overall assessment of their group experience was statistically significantly lower than the Spring 2022 group. But all comments were positive, with the exception of Student H4. Student H4 and V2 were partners in the Fall 2023 class. On the day of their presentation, Stu-

dent V2 did not attend class. Student H4 ended up giving the presentation alone. According to Student H4’s exit survey, “[Student V2] had texted the day of our presentation, after I had already presented, saying that he thought we presented on Monday so he ‘was not in a rush to look at the PowerPoint.’”

The Fall 2025 students reported the highest overall assessment of their group experience (M = 3.839, SD = .217). Its substitutability, cathexis, and inducibility were higher than the other groups also.

Student N5’s overall assessment was the second-lowest of the entire sample, with a mean of 2.81. He was in a group with Students M5 and O5, and had contacted the instructor/author the night before their presentation and asked for an extension. The student recanted the email the following day, saying they could present after all. Student N5 contextualized the relatively low 2.81 experience:

I do think that this is a great assignment. However, you will have some groups run into the issue of one or more members not pulling their weight, in which our case was [Student O5]. He did his slides quite literally the morning of the presentation, even though he said he would do them the night before.

Student O5 said in his own questionnaire that he was not “the best person to work with”:

I definitely procrastinated my work on the presentation. I was able to get in everything at the time of submission but I made my workmates concerned by how late it was that I did my part. I do think the work I did was good it was just the timeline of my own submission that I was concerned about.

Though the group received a 100%, Student O5’s peer-reviewed scores resulted in a final grade of a 71.40%. Students M5 and N5, received final grades of 89.1% and 97.35%, respectively, based on their peers’ assessment.

RQ3a asked how substitutability, cathexis, and inducibility predicted groups’ team scores. Results of a multiple linear regression were not statistically significant, $F(3,97) = 1.97, p = .123$. These three psychological processes could not predict the case study team grade groups earned.

RQ3b, on the other hand, asked how substitutability, cathexis, and inducibility could predict the scores students’ peers ultimately gave them. Results of a multiple linear regression was statistically significant, $F(3,96) = 75.59, p < .001$. The R² was .703. This

Table 2. Regression Coefficients for Scores Students Received From Peers

Variable	B	SE B	B	T	p
Substitutability	.341	.234	.082	1.455	.149
Cathexis	2.677	.970	.169	2.759	.007
Inducibility	15.545	1.31	.736	11.837	<.001

Note. R² = .703, Adjusted R² = .693, F(3,96) = 75.593, p < .001

means the extent to which students perceived their group members’ contributions to the group’s substitutability, cathexis, and inducibility predicted the individual score they gave each other.

Discussion

As stated in the introduction, the Spring 2022, Fall 2023, and Fall 2025 semesters denoted the 11th, 12th, and 13th times the instructor/author taught this course. On multiple occasions, the instructor/author learned that teams that appeared to be high functioning were anything but functional. This dynamic helps explain RQ3a: The degree of substitutability, cathexis, and inducibility group members perceived their group to have did not influence the team score groups received for their case studies, but it did predict the scores they gave each other. This study shows the role substitutability, cathexis, and inducibility have on team dynamics and outcomes. Groups could earn a high-quality grade, even if individual members weren’t pulling their weight. However, students with poor experiences are less likely to want to work in groups in the future (Abernethy & Lett, 2005). A benefit of the Brown (1995) autorating is that students exhibiting these behaviors will be held accountable for them, and thus, given the opportunity to self-reflect on how these traits influenced their group. This is especially important as students prepare to enter the workforce where teamwork is essential.

Also of practical importance is that, despite choosing their own partners, students in the Fall 2023 semester had a worse overall experience than their Spring 2022 and Fall 2025 counterparts. This echoes past research on self-selected versus instructor-selected groups, finding instructor-selected groups to be more enjoyable (e.g., Feichtner & Davis, 2015; Oakley et al., 2004; Obaya, 1999). What this study contributes, though, is that case study quality is overall lower in self-selected groups than in instructor-se-

lected groups. This relates to Oakley et al.’s (2004) conclusions that self-selected groups are more likely to be comprised of friends, and therefore, more likely to cover for each other than to inform instructors about failed participation or cheating.

Overall, this means the autorating dropped students’ final grades during the spring 2022 semester – despite students in this group producing, on average, higher quality projects than the fall 2023 class. In fact, only one three-person group in the Spring 2022 class did not receive a perfect 100% as a team score. This resulted in significantly higher average team scores (M = 97.74, SD = 5.96) than the Fall 2023 class (M = 89.52, SD = 19.58). Overall, the difference between the team score groups received and students’ individual, peer-reviewed grade was less than a point. The Brown (1995) autorating raised students’ scores in the Fall 2023 semester, despite Fall 2023 case studies being of lower quality. Twenty-seven of the 42 students received a perfect 100% as a team score. However, 15 students’ team score was lower due to failing to meet required elements, such as naming who all was affected by the situation, not mentioning which SPJ value or values were violated (i.e., seek truth and report it, minimize harm, act independently, and/or be accountable and transparent), or not creating a presentation that was at least 20 minutes. Five students’ grades would have risen by as much as 31%. One student’s grade would have dropped by 47%.

For the Brown (1995) autorating to do its job, three-person groups in this sample were more efficient than two-person groups. The instructor/author recommends selecting groups for students, as opposed to allowing them to self-select.

For the groups whose exit questionnaires suggested group members did not contribute evenly, the Week 1 and Week 3 questionnaires gave insight into the reasons. The instructor/author recommends instructors read the Week 1 and Week 3 question-

naires in real time in order to troubleshoot. For example, one of the questions during the Week 1 meeting was, “What is the reason for this group being together?” Student G, who was later critical of Student Q’s contribution to the project, answered, “To love and support Paul Pierce through his many bad decisions.” This is referring to a case study about ESPN firing Pierce, a former NBA player, after videos of Pierce under the influence with exotic dancers on Instagram went viral in 2021 (Scott, 2023). By reading students’ answers soon after their completion, instructors can possibly spot signs of a floundering group early and redirect them.

In addition, if using Brown’s (1995) autorating as students’ official case study grade, the instructor would need to be able to follow up with students immediately. For example, Student P reported a cathexis score of 3.33. This was in the bottom 5% of reported scores. However, in the comments at the end of the questionnaire, she said, “I enjoyed the case study presentation!” This comment did not contextualize these scores. In such instances, an instructor would need to contact students such as Student P for clarification before administering a final grade.

Though piloting these measures were insightful, this study has limitations. The classes in which these measures were piloted spanned three years, and included both undergraduate and masters students. However, these were the last three times the instructor/author was assigned to teach these classes. Also, the raw numbers of how students responded to the Oakley et al. (2004) measures can be misleading. Regardless of group quality, some students still preferred to work alone. Student F, for example, had a cathexis rating of 4.0, which fell in the lowest 10% of reported scores. This was due to him agreeing with several positive statements instead of strongly agreeing. He contextualized this in the questionnaire:

I loved working the case study with my teammates. They were great people to work with, but as [a] sports junkie and competitor at heart, I love doing the work myself. I would like to maybe see some case studies being researched and presented individually.

A third limitation was that this study had only one evaluation point after the initial meeting but before the final project. Scholars recommend having multiple evaluation points as a way of addressing free riders early (Brooks & Ammons, 2003; Sridharan et al., 2018). According to Brooks and Ammons (2003),

a drop in evaluation source’s variance over two consecutive peer evaluations suggests an under-performing student has improved. The instructor/author will compensate for this by incorporating peer evaluation measures at both the mid-task and post presentation stages in the next iteration of this course (Dix, 2024). A fourth limitation is that students did not rate themselves. The instructor/author did not have students self-assess because past research on the effectiveness of self-assessment has found mixed results. When self-rating, students may inflate their own scores (Goldfinch, 1994; Willmot & Crawford, 2007; Zang et al., 2008). If a generous student gives higher scores to others than what their group members gave them, their own score will be downsized (Dix, 2024; Goldfinch, 1994). However, student self-rating is also a component of self-reflection (Dix, 2024). In the next iteration of the course, the instructor/author plans to add self-assessment, modifying Oakley et al.’s (2004) measures so that they read in the first person.

A future area of research this study did not take into consideration was a more recent development in peer evaluation, which is to give students the ability to “fire” an underperforming group member (e.g., Abernethy & Lett, 2005; Dix, 2024). In this method, performing students send a notice (e.g., e-mail) to a nonperforming group member and to the course instructor, outlining what the nonperforming group member is required to do and by which date (Abernethy & Lett, 2005; Dix, 2024). If the nonperforming student fails to meet this requirement, they are “fired” and receive a zero as a grade. The instructor/author did not use this method because the purpose of the Ethics in Sports Media course is not to create effective groups and build team effectiveness but to “apply the principles and laws of freedom of speech and press”; “demonstrate an understanding of the multicultural history and role of professionals and institutions in shaping communities”; “apply critical thinking skills in conducting research and evaluating information”; “demonstrate an understanding of professional ethical principles and work ethically in pursuit of truth, accuracy, fairness and service to all people and communities”; and to build “skilled and culturally proficient communicators capable of collaborating through multicultural teams and understanding global perspectives” (“Principles of Accreditation,” 2025, para. 13). Though the final competency may be applicable, the instructor/author was more focused on piloting three measures from three disparate fields and deter-

mining their success while creating this study.

Future research could also incorporate audience participation as a layer of peer assessment. Though outside of the purposes of the current study, the instructor/author also has students watching case study presentations complete an audience feedback questionnaire after each presentation. In this questionnaire, students are asked to rate, on a scale of 1 (strongly disagree) to 5 (strongly agree) their level of agreement with five statements: “The presenters are knowledgeable about the materials;” “The presenters clearly presented the subject material;” “The presenters conveyed enthusiasm for the material;” “The presenter(s) encouraged audience input and discussion;” “Because of this presentation, I know something about the subject I did not know before coming to class today.” On the last day of class, the instructor/author honors groups whose case studies garnered the three highest feedback scores, called “The People’s Choice Awards.” The instructor/author also keeps an all-time list of the course’s most highly rated presentations and shows the class the list when introducing students to the assignment. This is done so students can see examples of quality case studies and for students to understand they are part of a larger school tradition. However, these measures and/or their validity need to be modified. Multiple students have been caught completing these forms before groups even present, giving groups perfect scores of 25. One student in the Fall 2023 class actually completed audience feedback for the presentations that never happened—class had been canceled that day. This was a violation of the school’s Academic Integrity Policy and defeated the purpose of the audience feedback. Creating more valid peer-review audience feedback measures is the next step in this research.

References

Abernethy, A., & Lett, W. (2005). “You are fired!” A method to control and sanction free riding in group assignments. *Marketing Education Review*, 15(1), 47–54. <https://doi.org/10.1080/10528008.2005.11488891>

Ancona, D. G. (1990). Outward bound: Strategies for team survival in an organization. *Academy of Management Journal*, 33(2), 334–365. <https://psycnet.apa.org/record/1990-27049-001>

Ancona, D. G., & Caldwell, D. F. (1992a). Bridging the boundary: External activity and performance in organizational teams. *Administrative*

Science Quarterly, 37(4), 634–665. <https://doi.org/10.2307/2393475>

Ancona, D. G., & Caldwell, D. F. (1992b). Demography and design: Predictors of new product team performance. *Organization Science*, 3(3), 321–341. <https://doi.org/10.1287/orsc.3.3.321>

Brooks, C. M., & Ammons, J. L. (2003). Free Riding in Group Projects and the Effects of Timing, Frequency, and Specificity of Criteria in Peer Assessments. *Journal of Education for Business*, 78(5), 268–272. <https://doi.org/10.1080/08832320309598613>

Brown, R. W. (1995, November). Autorating: Getting individual marks from team marks and enhancing teamwork. Proceedings Frontiers in Education 1995 25th Annual Conference. *Engineering Education for the 21st Century*, 2, 3c2–15. <https://doi.org/10.1109/FIE.1995.483140>

Chicago Psychoanalytic Institute. (2024, October 24). Exploring cathexis: Understanding emotional energy investment for personal growth. <https://chicagoanalysis.org/blog/conditions-and-diagnoses/cathexis-emotional-energy/>

Cohen, S. G., & Bailey, D. E. (1997). Trends in team effectiveness research: From the shop floor to the executive suite. *Journal of Management*, 23(3), 239–290. <https://doi.org/10.1177/014920639702300303>

Dix, S. (2024). A streamlined survey instrument to moderate university students’ grades for group projects. *Asian Pacific Journal of Marketing and Logistics*, 36(7), 1767–1781. <https://doi.org/10.1108/APJML-09-2023-0858>

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations*, 2(2), 129–152. <https://doi.org/10.1177/001872674900200204>

Deutsch, M. (1962). Cooperation and trust: Some theoretical notes. In M. R. Jones (Ed.), *Nebraska symposium on motivation* (Vol. 10, pp. 275–319). University of Nebraska Press.

Durdubas, D., & Koruc, Z. (2023). Effects of a multifaceted team goal-setting intervention for youth volleyball teams. *Journal of Applied Sport Psychology* 35(2), 224–243. <https://doi.org/10.1080/10413200.2021.2021564>

Feichtner, S. B., & Davis, E. A. (2015). Why some groups fail: A survey of students’ experiences with learning groups. *Journal of Management Education*, 40(1), 12–29. <https://journals.sagepub.com/doi/10.1177/1052562915619639>

Felder, R. M., & Brent, R. (2007). Cooperative learning. In P.A. Mabrouk (Ed.), *Active learning: Models From the analytical sciences* (pp. 34–53). American Chemical Society. <https://doi.org/10.1021/bk-2007-0970.ch004>

Felder, R. M., & Brent, R. (2009). Effective teaching: A workshop. *Chemical Engineering Education*, 43(1), 15–16. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=9d80bdc46c1205e5907dfdf34b98296d130386eb>

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education (Online)*, 41(3), 39–54. <https://www.tandfonline.com/doi/abs/10.1080/0260293940190103>

Goldfinch, J. (1994). Further developments in peer assessment of group projects. *Assessment & Evaluation in Higher Education*, 19(1), 29–35.

Halfhill, T., Sundstrom, E., Lahner, J., Calderone, W., & Nielsen, T. M. (2005). Group personality composition and group effectiveness: An integrative review of empirical research. *Small Group Research*, 36(1), 83–105. <https://doi.org/10.1177/1046496404268538>

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A metanalysis. *Psychological bulletin*, 89(1), 47.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (1998). Cooperative learning and social interdependence theory. In R. S. Tindale (Ed.), *Theory and research on small groups* (pp. 9–35). Springer US. https://link.springer.com/chapter/10.1007/0-306-47144-2_2

Johnson, D. W., & Johnson, R. T. (2005). Cooperative learning, values, and culturally plural classrooms. In *Classroom Issues* (pp. 29–47). Routledge.

Johnson, D. W., & Johnson, R. T. (2015). Theoretical approaches to cooperative learning. In R. M. Gillies (Ed.), *Collaborative learning: Developments in research and practice* (pp. 17–46). Nova Science Publishers.

Khuzwayo, M. (2018). Assessment of group work in initial teacher education and training. *South African Journal of Education*, 28(2), 271–271. <https://journals.co.za/doi/abs/10.15700/saje.v38n2a1386>

Kim, Y. S., & Kim, B. J. (2012). The change of team cohesion in a university baseball team by applying team-building intervention. *Korean Journal of Sport Science*, 23(2), 334–342. https://www.kci.go.kr/kciportal/landing/article.kci?article_id=ART001670727

Koffka, K. (1935). Points and lines as stimuli. In *Principles of Gestalt Psychology* (pp. 148–160). Harvacourt Brace.

Lax, J., & Fentiman, A. W. (2016, June), Improving graduate student oral presentations through peer review [Paper presentation]. American Society for Engineering Education Conference 2016, New Orleans, Louisiana. <https://peer.asee.org/improving-graduatestudent-oral-presentations-through-peer-review>

Levine, J. M., & Moreland, R. L. (2012). A history of small group research. In A. W. Kruglanski & W. Stroebe (Eds.), *Handbook of the history of social psychology* (pp. 383–405). Psychology Press. <https://psycnet.apa.org/record/2011-25229-018>

Lewin, K. (1948). Resolving social conflicts; selected papers on group dynamics.

Neto, J., Neto, F., & Furnham, A. (2023). Predictors of students’ preferences for assessment methods. *Assessment & Evaluation in Higher Education*, 48(4), 556–565.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://www.tandfonline.com/doi/abs/10.1080/03075070600572090>

Oakley, B., Felder, R. M., Brent, R., & Elhadj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, 2(1), 9–34. [https://engr.ncsu.edu/wpcontent/uploads/drive/1ofGhdOciEwloA2zofffqkr7jG3SeKRq3/2004-Oakleypaper\(JSCL\).pdf](https://engr.ncsu.edu/wpcontent/uploads/drive/1ofGhdOciEwloA2zofffqkr7jG3SeKRq3/2004-Oakleypaper(JSCL).pdf)

Obaya, A. (1999). Getting cooperative learning. *Science Education International*, 10(2), 25–27. <https://www.scirp.org/journal/paperinformation?paper-id=115369>

Ocampo, J. C., Panadero, E., & Diez, F. (2023). Are men and women really different? The effects of gender and training on peer scoring and perceptions of peer assessment. *Assessment and Evaluation in Higher Education*, 48(6), 760–776. <https://doi.org/10.1080/02602938.2022.2130167>

Principles of Accreditation (2025). Accrediting Council on Education in Journalism and Mass Communications. <https://www.acejmc.org/policies-process/principles>

Scott, J. (2023, April 10). Paul Pierce decries firing by ESPN: “What did I do wrong?” *Sports Illustrated*. <https://www.si.com/nba/2023/04/10/paul-pierce-comments-firing-espninstagram-live>

Sharp, J., & Dyrud, M. (2009, June). Two perspectives on peer review [Paper presentation]. Annual Conference & Exposition, Austin, TX, United States. <https://peer.asee.org/twoperspectives-on-peer-review>

Smith, K. A., Sheppard, S. D., Johnson, D. W., & Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education*, 94(1), 87–101. <https://doi.org/10.1002/j.2168-9830.2005.tb00831.x>

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21–51. <https://doi.org/10.3102/00346543069001021>

Sridharan, B., Tai, J., & Boud, D. (2018). Does the use of summative peer assessment in collaborative group work inhibit good judgement? *Higher Education*, 77(5), 853–970. <https://link.springer.com/article/10.1007/s10734-018-0305-7>

Tassi, J. M., López-Gajardo, M. A., Leo, F. M., Díaz-García, J., & García-Calvo, T. (2023). An intervention program based on team building during tactical training tasks to improve team functioning. *Frontiers in Psychology*, 14, Article 1065323. <https://doi.org/10.3389/fpsyg.2023.1065323>

Terenzini, P. T., Cabrera, A. F., Colbeck, C. L., Parante, J. M., & Bjorklund, S. A. (2001). Collaborative learning vs. lecture/discussion: Students’ reported learning gains. *Journal of Engineering Education*, 90(1), 123–130. <https://doi.org/10.1002/j.2168-9830.2001.tb00579.x>

Willmot, P., & Crawford, A. (2007). Peer review of team marks using a web-based tool: An evaluation. *Engineering Education: The Journal of Higher Education Academy Engineering Subject Centre*, 2(1), 59–66. <https://doi.org/10.1112/ened.2007.02010059>

Wong, F.M.F., Kan, C.W.Y., & Chow, S.K.Y. (2022). From resistance to acceptance in small group work: Students’ narratives. *Nurse Education Today*,

111, 105317. <https://www.sciencedirect.com/science/article/abs/pii/S0260691722000533?via%3Dihub>

Zhang, B., & Ohland, W. M. (2009). How to assign individualized scores on a group project: An empirical evaluation. *Applied Measurement in Education*, 22(3), 290–308. <https://www.tandfonline.com/doi/abs/10.1080/08957340902984075>

**Appendix A:
Case Study Rubric**

The instructor/author grades the case study based on the following elements:

Facts of the case/story. This includes the who, what, where, when, and why of the story. What is the news organization or sports journalist involved? When and where did this incident occur? Why did this happen?

Ethical issue(s) involved. Which SPJ value or values were violated (i.e., Seek truth and report it, minimize harm, act independently, and/or be accountable and transparent)?

Stakeholders. Who was affected by this incident?

Arguments PRO. Why might the sports journalist or news organization have done this? What might they have been trying to achieve?

Arguments CON. Why should the sports media organization or sports journalist have not done this?

Your decision. What would you do to address this situation?

What happened: What did the news organization or sports journalists do in response to this situation, if anything?

Reaction. What was the public reaction, if any?

Questions for the audience. How will you engage the audience? What do you want the audience to consider when wrapping their heads around this case? Plan in-class activities that will engage your classmates.

Time. Presentations, including audience participation, must be at least 20 minutes.

Sada J. Reed’s research explores how sports journalists do their job, particularly how their routines and practices shape the content they create. Examples include challenges associated with sports journalists’ use of social media, hero mythology, and source selection in content creation, as well as role conflict and differences between how American and Russian sports journalists cover athlete doping scandals.

Correspondence concerning this article should be addressed to Sada Reed, Walter Cronkite School of Journalism and Mass Communication at Arizona State University, 555 N. Central Avenue, Phoenix, AZ 85004

