



OPEN ACCESS

EDITED BY

Danijel Dankovic,
University of Niš, Serbia

REVIEWED BY

Mark Vicars,
Victoria University, Australia
Liat Eyal,
Levinsky College of Education, Israel
Topias Ikävalko,
University of Helsinki, Finland

*CORRESPONDENCE

Verena Schürmann
✉ verena.schuermann@hsrw.eu

RECEIVED 11 November 2024

ACCEPTED 10 March 2025

PUBLISHED 21 March 2025

CITATION

Schürmann V, Bodemer D and
Marquardt N (2025) Exploring the use of
regular reflections in student collaboration: a
case study in higher education.
Front. Educ. 10:1526487.
10.3389/feduc.2025.1526487

COPYRIGHT

© 2025 Schürmann, Bodemer and Marquardt.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Exploring the use of regular reflections in student collaboration: a case study in higher education

Verena Schürmann^{1,2*}, Daniel Bodemer² and Nicki Marquardt¹

¹Faculty of Communication and Environment, Rhine-Waal University of Applied Sciences, Kamp-Lintfort, Germany, ²Research Methods in Psychology – Media Based Knowledge Construction, University of Duisburg-Essen, Duisburg, Germany

Introduction: With the growing emphasis in higher education on fostering collaboration and reflection, this study examines the intersection of these two concepts by exploring the use of regular reflections in student collaboration.

Methods: An embedded case study approach was employed, investigating four student teams over a 15-week interdisciplinary project course at a higher education institution. Each team participated in four joint reflections, supported by Behaviorally Anchored Rating Scales (BARS). Multiple data sources, including questionnaires, interviews, and documents, were collected at both the team and individual levels.

Results: The findings reveal a positive improvement in students' self-assessed collaboration in three out of the four teams over the semester. These teams also experienced an increase in psychological safety. Triangulation and the comparison of two contrasting cases provided deeper insights into these patterns. While the data indicated general satisfaction with the reflection sessions and the BARS, several challenges, influencing factors, and areas for improvement were identified.

Discussion: This study offers valuable insights into the dynamics and quality of joint reflections within student teams. It provides practical recommendations for integrating reflective practices into higher education courses and highlights avenues for future research.

KEYWORDS

collaboration, student teamwork, reflective practice, debriefings, case study, higher education

1 Introduction

Given the complex, fast-changing and ambiguous circumstances in the contemporary global environment, education institutes aim to support students in becoming lifelong learners. To ensure graduates' employability and success, curriculum initiatives related to soft skills are required (Ritter et al., 2018). In particular, collaboration and reflection competencies are highly valued and demanded on the job market. Collaborative learning approaches should be widely integrated into higher education curricula to prepare students for working in teams (Robbins and Hoggan, 2019). The ability to reflect is considered as being a key in becoming employable (Heymann et al., 2022) and as being essential for lifelong learning (Bharuthram, 2018; Tan, 2021).

Considering the growing interest in higher education to foster practices of collaboration and reflection, the current study explores these two emerging concepts in combination by investigating the application of regular reflections in student collaboration. One key rationale for integrating reflective practices into collaborative learning is its established role in enhancing professional practice, particularly in fields such as medical and teacher education (e.g., [Chan and Lee, 2021](#)). Despite this, structured reflection has received limited attention in other academic disciplines (*ibid.*). By implementing structured and regular reflections, this study seeks to bridge this gap, examining whether and how reflective practices contribute to students' overall learning experiences, including their ability to collaborate effectively and develop a deeper understanding of group dynamics.

As many students find it difficult to navigate the social dynamics inherent in group work ([Näykki et al., 2014](#)), [Jones et al. \(2022\)](#) highlight the need for a better understanding of these phenomena through case studies of authentic group situations. Aimed at exploring whether and how joint reflective practices, which are relatively unexplored in the higher education context, can support students' collaborations, this study employs a case study approach combining multiple sources of evidence. We analyze student teams conducting four group reflections during a 15-week interdisciplinary project course at a higher education institution. To understand whether and how regular reflection helps students working in collaborative learning environments, we present insights from quantitative and qualitative data on both team level (e.g., collaboration) and individual level (e.g., perceived usefulness of the reflections). Examining and evaluating collaboration and group reflection in everyday education, this study seeks to present explicit recommendations for researchers and practitioners working with student teams. For instance, how students can be encouraged to reflect about their collaboration, what is important to them in the process, and which potential influencing factors need to be considered.

1.1 Collaboration

During collaboration, two or more learners are involved in interaction with each other as well as with tools and resources, working toward a shared goal ([Bedwell et al., 2012](#); [Dillenbourg, 1999](#); [Patel et al., 2012](#)). Within this process, students actively engage in constructing and maintaining a joint problem space through sharing, expanding on and negotiating ideas ([Borge and Mercier, 2019](#); [Roschelle and Teasley, 1995](#)). Ideally, the collaborative group achieves both a common solution to solve the problem as well as skill and knowledge gain of each group member ([Schürmann et al., 2024](#)).

The importance of collaborative working and learning approaches has already been known for a long time and research on collaboration has a long tradition. Collaborative settings encourage students to participate in discussions that not only deepen their understanding of the material but also help them develop shared mental models of complex phenomena, promoting positive interdependence that nurtures collaboration and mutual support in achieving common goals ([Johnson et al., 2014](#)). As emphasized by [Scager et al. \(2016\)](#), cognitive benefits (e.g., conceptual understanding) as well as an improvement of social skills are provided through collaborative learning approaches. Therefore, using collaborative learning approaches might function as both a *mean* (collaborating to learn, i.e.,

knowledge construction, solving problems) and an *objective* (learning to collaborate, i.e., development of social skills; see [Salmons, 2019](#)). These dimensions represent two traditions in the collaboration literature. One perspective is grounded in the constructivist theories of Piaget and Vygotsky, which argue that knowledge is actively built through social interaction and dialogue, and the other perspective represents a shift toward recognizing collaboration as an essential human skill (see [Child and Shaw, 2019](#); [Rowe, 2020](#)).

While several studies highlight the effectiveness of learning and working together in teams (e.g., [Lou and Macgregor, 2004](#); [Nokes-Malach et al., 2015](#); [Tao and Gunstone, 1999](#)), collaborative learning is not always superior to individual learning ([Retnowati et al., 2018](#)). Moreover, research and practice show that students are indecisive, predominantly skeptical about group work because of interpersonal conflicts, logistical challenges such as scheduling meetings and time management issues as well as unequal contributions of group members ([Donelan and Kear, 2024](#); [Scager et al., 2016](#); [Wilson et al., 2018](#)). This is particularly true for students at higher educational levels as they often work on challenging, ill-defined tasks characterized by ambiguity while receiving less support compared to earlier stages of education. Their collaboration is mostly self-organized outside class hours and in absence of instructors ([Scager et al., 2016](#)). Additionally, higher education students already bring with them a range of prior experiences and attitudes toward group work which, in turn, can influence team orientation, conflict resolution and outcomes ([Fransen et al., 2013](#); [Pauli et al., 2008](#)).

Although collaborative learning approaches are ubiquitous, students at various educational levels are often not properly instructed on how to work in groups ([Le et al., 2018](#); [Leopold and Smith, 2020](#)). Research suggests that instructors often engage primarily in logistical aspects of group work, such as assigning tasks and group size, rather than actively facilitating collaboration ([Xu, 2024](#)). Consequently, higher education students often report a lack of collaborative skills leaving them feeling ill-prepared for collaborative work—a phenomenon observed globally, including in both Western (e.g., [Wilson et al., 2018](#)) and Eastern contexts (e.g., [Xu, 2024](#)). As summarized by [Borge and Mercier \(2019\)](#), “collaboration is a collective cognitive endeavor that is hard, requires sustained effort from participants to work well, and is prone to breakdowns caused by different socio-emotional, cognitive, metacognitive, and socio-metacognitive problems” (p. 220). Typical problems encountered in student group work include different task understandings, altering levels of interests and commitment, or incompatible working and interaction styles, among others ([Näykki et al., 2014](#)). Especially in interdisciplinary student groups, collaboration issues often arise due to conflicting schedules, competing commitments (e.g., extracurricular activities), and differing priorities (see [Hussein, 2021](#)). Furthermore, the rise of virtual and hybrid learning settings since the COVID-19 pandemic has introduced additional challenges for students, particularly in managing their socially shared regulation of learning (e.g., [Donelan and Kear, 2024](#); [Oshima et al., 2024](#); [Wildman et al., 2021](#)).

But not only students face various obstacles with regard to collaborative learning, so do higher education instructors. For instance, they are struggling with crowded curricula and limited resources. Besides, although many higher education instructors are strong subject matter experts, they have received little pedagogical training. This impedes the selection of appropriate collaborative

settings and tasks as well as giving instructions on how to work together successfully. Therefore, the focus of their teaching often is on disciplinary knowledge while collaborative learning goals take a back seat, with collaboration skills typically treated as a mere byproduct of group work (Boud and Bearman, 2024). Yet, even if instructors aim to develop both cognitive and collaborative skills through applying collaborative learning practices, they face challenges in assessing the collaborative processes and skills (Boud and Bearman, 2024; Le et al., 2018).

1.2 Reflection on collaboration

Based on John Dewey's criteria for reflection, Rodgers (2002) describes reflection as "a meaning-making process that moves a learner from one experience into the next with deeper understanding of its relationships with and connections to other experiences and ideas" (p. 845). Hence, reflection is a purposeful thinking process involving affect and cognition, which is controlled by a learner (Nguyen et al., 2014; Yukawa, 2006). This process engages critical thought about an experience and generates insights that can inform future learning and development (Quinton and Smallbone, 2010). Besides reflecting on one's own thoughts and actions, it is promising to reflect in interaction with others about collaborative processes. Collaborative environments provide greater opportunities for reflection compared to individual settings, as they facilitate the exchange of diverse perspectives and promote deeper critical engagement with experiences (Yang and Choi, 2023). Reflecting on peer feedback is expected to enhance group members' awareness of their own behavior, its impact on others, and the potential need for adjustments (Eshuis et al., 2019; Phielix et al., 2011). By doing so, not only self-regulation but also mutual understanding and coordination within the group can be fostered. For instance, in medical education, reflective practice often takes place in form of orally conducted collaborative discussions giving students the opportunity to prepare themselves for their career including learning to collaborate (Johansson et al., 2017; Paige et al., 2021). This technique is referred to as *debriefing*.

Originally stemming from high reliability industries such as military and medicine, debriefings are a well-established and systematically applied concept for learning in teams in these contexts (Allen et al., 2018; Keiser and Arthur, 2021). Debriefings are structured group reflexivity interventions, in which teams are asked to reflect about recent events during which they collaborated (Allen et al., 2018). This process is often led by a facilitator but is associated with comparable effectiveness if self-led (Boet et al., 2011; Keiser and Arthur, 2021). Particularly with the advent of agile working methods, similar concepts can be found under different terms in many other working contexts (e.g., retrospective meetings conducted by scrum project teams). What they have in common is that they are based on the concept of experiential learning (Kolb, 1984), and combine feedback, reflection, and discussion (Keiser and Arthur, 2021). The purposes of group reflexivity interventions are manifold inclusive of improving teams' problem-solving and decision-making processes, establishing psychological safety, enhancing group identity, and reducing failures (e.g., Allen et al., 2018; Schippers et al., 2014; Schippers and Rus, 2021). Their effectiveness in terms of several criteria such as individual and team performance has been confirmed

in a number of studies including meta-analyses (e.g., Keiser and Arthur, 2021; Tannenbaum and Cerasoli, 2013).

Considering the importance of reflective practices for lifelong learning, reflection in higher education is receiving increased attention in recent years (e.g., Tan, 2021; Veine et al., 2020). As demonstrated by Merkebu et al. (2023), factors enhancing reflection are metacognition and emotional regulation. In general, the concepts of metacognition and regulation (e.g., self-regulation, socially shared or emotional regulation) have been advancing in significance within educational theory, research and practice (e.g., Hadwin and Oshige, 2011; Järvelä et al., 2015; Järvenoja et al., 2013; Kaplan, 2008). However, as of yet, reflective practices are predominantly common and studied in medical (e.g., Dornan et al., 2019; van Braak et al., 2021) and teacher education (e.g., Mortari, 2012; Slade et al., 2019) but rather scarce in other disciplines (Chan and Lee, 2021).

Moreover, research and practice revealed that students tend to struggle with reflection tasks (e.g., Moreland and McMinn, 2010; Tan, 2021). In their literature review, Chan and Lee (2021) refer to misconceptions by both students and teachers. Challenges at the student level include little motivation, lacking understanding of and ability to engage in reflection, as well as emotional impact through feelings of stress, vulnerability and anxiety (Chan and Lee, 2021; Tan, 2021). Examples for challenges at the instructor level are difficulties with pedagogical logistics like large classes and tight teaching schedules, choosing appropriate reflection approaches as well as assessment and feedback issues due to subjectivity and lacking criteria (Chan and Lee, 2021). Typically, students are asked to reflect individually and in written form such as through learning journals, diaries and portfolios (Chan and Lee, 2021; Heymann et al., 2022). Nevertheless, some research findings indicate that writing is not students' preferred mode of expression (see Chan and Lee, 2021).

1.3 Fostering students' collaboration and reflection using behaviorally anchored rating scales: the current study

Although some students might have a natural tendency to reflect on a collaborative learning experience, it is rather unlikely that students reflect together as a group and in a systematic way about what happened and what could have been improved (see Fanning and Gaba, 2007; Harrison et al., 2003). Considering the positive findings of group reflexivity interventions (i.e., debriefings) in medical education and workplace settings, the current study transfers their systematic approach to higher education context.

Asking and supporting students to reflect regularly on their collaboration seems to be promising in various ways: Engaging in a joint meaning-making process of experienced collaborative work might help students to overcome challenges like task understanding and altering levels of interests or working styles. Additionally, students' (re-)configuration of internal collaboration scripts (i.e., learner's knowledge that guides understanding and action in collaborative practices, Fischer et al., 2013; Kollar et al., 2007) could be fostered. As students often find it difficult to raise sensitive issues (e.g., incompatible interaction styles or sharing the feeling of frustration due to low commitment of some collaborative partners) in collaborative work, debriefings provide a good opportunity for students to address these issues as they are explicitly asked to.

Furthermore, the systematic and regular implementation of reflexivity tasks might help to raise students' understanding of and ability to engage in reflection.

Debriefings should be ideally designed in accordance with the participants' needs (Allen et al., 2018), such being provided with a clear structure and guidance during joint reflections (Schürmann et al., 2025). Behaviorally Anchored Rating Scales (BARS) might be promising to this end. In this rating format, performance dimensions and scale values are defined in behavioral terms (Schwab et al., 1975). Typically, rating scales ranging from five to 10 points are employed, incorporating behavioral anchors to illustrate varying levels of quality within a given construct, with higher scale values indicating more effective behaviors (Debnath et al., 2015). Research indicates promising findings in terms of their use for self- and peer-evaluation purposes (Ohland et al., 2012) and their feedback potential (Horn et al., 1982). Moreover, BARS have been successfully applied in team-based contexts, particularly for assessing team member effectiveness and evaluating team adaptation processes (Georganta and Brodbeck, 2020; Ohland et al., 2012). By providing information on what constitutes a poor, a satisfying and a good performance, using BARS might (re)configure students' knowledge about collaboration (see Ohland et al., 2012) and stimulate reflection processes. Accordingly, BARS might be a beneficial tool to support and structure students' self-led debriefings.

Taken together, the current study explores the application of structured and regularly conducted group reflexivity interventions (i.e., debriefings) in higher education context. We attempt to understand whether and how this approach helps students working in collaborative learning environments. In addition, this study investigates the potential of BARS as a tool to support students' debriefings. We aim to answer the following questions, thus yielding valuable implications for higher education practitioners:

- RQ 1: How does collaboration of student teams develop over time when integrating regular debriefings in course design?
 RQ 2: (How) do debriefings help to foster collaboration in an educational setting?
 RQ 3: Which features of debriefings are valuable to students in higher education context?

Our assumption is that students' collaboration improves through regular reflection. We further postulate that the debriefings are perceived as useful for learning about how to collaborate. While RQ 2 explores the general effectiveness of regular debriefings in fostering collaboration within the context of higher education, RQ 3 delves more specifically into the role of BARS and examines their value in enhancing the quality and impact of debriefings for students.

2 Methods

Given that debriefings using BARS have not yet been investigated as well as the descriptive and explanatory nature of our research questions, we conducted a case study in higher education context. We investigated student teams conducting four group reflections in a naturalistic setting, namely during a 15-week interdisciplinary project course. Case studies come with the strength of providing rich descriptions and insightful explanations on social

phenomena (Yin, 2018). More precisely, we used an embedded single case study design including multiple sources of evidence such as questionnaires, interviews and documents (Yin, 2018). This way, we were able to analyze student project teams and their perception regarding team-level data (e.g., collaboration) as well as learners' individual opinions and attitudes in real-life context. We pre-registered our approach on [aspredicted.org](https://aspredicted.org/#112332) (#112332). Furthermore, an ethics committee reviewed and approved the study.

2.1 Research context

The study was conducted in a one semester-long, interdisciplinary project course offered annually at a university in Germany. In this course, all faculty students being in their fifth or higher semester can choose from a range of topics presented by instructors of various disciplines. The learning objectives include encouraging students to discover new topics and enabling them to work together in international and interdisciplinary teams in authentic learning settings. Each summer term, students receive basic information about the project options; afterwards, they can enroll in their favored project taking place in winter term. The interdisciplinary projects take places throughout the entire winter term, i.e., 15 weeks.

The project investigated in the current study is one of 16 interdisciplinary projects ($N = 199$ participants) that took place in winter term 2022/2023. It dealt with creating a podcast for supporting students during thesis writing. Prior to enrollment, students were informed that this project is to be accompanied by a case study. The course limit was set to 16 participants. After the first week, one student withdrew her enrollment due to personal reasons.

2.1.1 Participants

In total, 15 students (12 female, three male) nested in four teams were involved in the study. The participants aged 20–26 ($M = 22.6$, $SD = 1.8$) were bachelor and master students of the disciplines *psychology*, *business* and *computer sciences*. Table 1 provides further information on team's composition. In three of the teams, students reported having worked with some, but not all, of their teammates prior to the course. In contrast, members of team 4 indicated that they had no prior experience working together before the course. In addition to fulfilling regular course requirements, participants received either course credits or monetary compensation for

TABLE 1 Overview of student team's composition.

Team	Composition
Team 1	Four female undergraduate students: Two psychology students, two business students ($M_{age} = 22.5$, $SD_{age} = 1.7$)
Team 2	Four female undergraduate students: Two psychology students, two business students ($M_{age} = 22$, $SD_{age} = 2.2$)
Team 3	Four female students (three undergraduate, one graduate): One psychology student, two business students, one psychology and business student ($M_{age} = 23.3$, $SD_{age} = 1$)
Team 4	Three male undergraduate students: Two psychology students, one computer sciences student ($M_{age} = 23$, $SD_{age} = 2.6$)

The students formed the teams themselves under the condition of having at least two different disciplines represented in each team.

contributions beyond standard tasks, such as participation in interviews or completing additional assignments.

2.1.2 Setting: interdisciplinary project course

2.1.2.1 Project aim and topic

The aim of the podcast to be created was to provide students writing their thesis with helpful organizational information (e.g., what to consider for thesis registration) and motivational tips (e.g., how to deal with writer's block). The podcast was to be published in a thesis writing course on the learn management system of the university.

2.1.2.2 Project instructors

The project was offered and taught by the first author of the study who has a background in psychology (VS). Moreover, two colleagues (one educator in the field of media and computer sciences, one educator in the field of psychology) and a student assistant (familiar with podcast equipment and audio editing) supported course instruction and podcast production but were not involved in the research project.

2.1.2.3 Project procedure

The course took place in a hybrid setting with nine of 15 weeks having fixed meetings at university. The remaining time was to be self-organized by the student teams. The project started with a kick-off meeting in which organizational aspects like project background and aim, timetable and requirements to complete the course were presented. Additionally, the students were given time to get to know each other. They were explicitly asked to get in touch with students from other disciplines. In the second session, the 15 students formed four interdisciplinary teams (see Table 1), each of which was to create two podcast episodes related to thesis writing. Hereby, the students could either choose a topic from a list (provided by the instructors) or suggest an own topic while keeping the aim of the podcast in mind. For each episode, the teams should prepare and present their idea (concept) to the course before recording it.

2.2 Data collection procedure

At the beginning of the course, all students received an instruction on collaboration. The aim of the instruction was to inform students about the main characteristics of student collaboration while highlighting behaviors representing an effective collaboration (see Eshuis et al., 2019) based on the framework on collaboration in higher education by Schürmann et al. (2024). It was delivered by the course instructor (VS). Moreover, one team member of each team volunteered as so called multiplier. The four multipliers were psychology students who were instructed on how to implement, conduct and document a debriefing (focus: collaboration) within their teams. They received a document with explanations and guiding questions for the debriefing. Furthermore, they were compensated for their additional efforts with research participation credit. As depicted in Figure 1, we combined several research methods to collect quantitative and qualitative data on both individual and team level. Data were collected *before* (t_1), *during* (t_2 – t_5) and *after* (t_6 , t_7) student teams' main working phase. The first author took field notes throughout the project course to document her observations.

2.2.1 Before teams' main working phase

Prior to teams' main working phases, students filled out a questionnaire which comprised the informed consent and questions on demographics and psychological safety. The students were asked whether they have ever worked with their team members *before* the project course (1 = *no*; 2 = *yes, partly*; 3 = *yes, with all of them*). Psychological safety being strongly related to collaboration, is often considered an important correlate in debriefing and team literature (e.g., Edmondson and Lei, 2014; Newman et al., 2017). Psychological safety describes a "a sense of confidence that the team will not embarrass, reject, or punish someone for speaking up" (Edmondson, 1999, p. 354). A low level of psychological safety is likely to undermine reflective learning conversations (Kolbe et al., 2020). In this study, we aimed to consider teams' perceived psychological safety prior to the main working phase serving as a baseline. It was measured with the German PsySafety-Check scale by Fischer and Hüttermann (2020) containing seven items answered on a 7-point Likert scale (1 = *completely disagree*, 7 = *completely agree*). One illustrative item is "No one in this team would deliberately act in a way that would undermine my efforts" ($\alpha_{11} = 0.88$).

2.2.2 During teams' main working phase

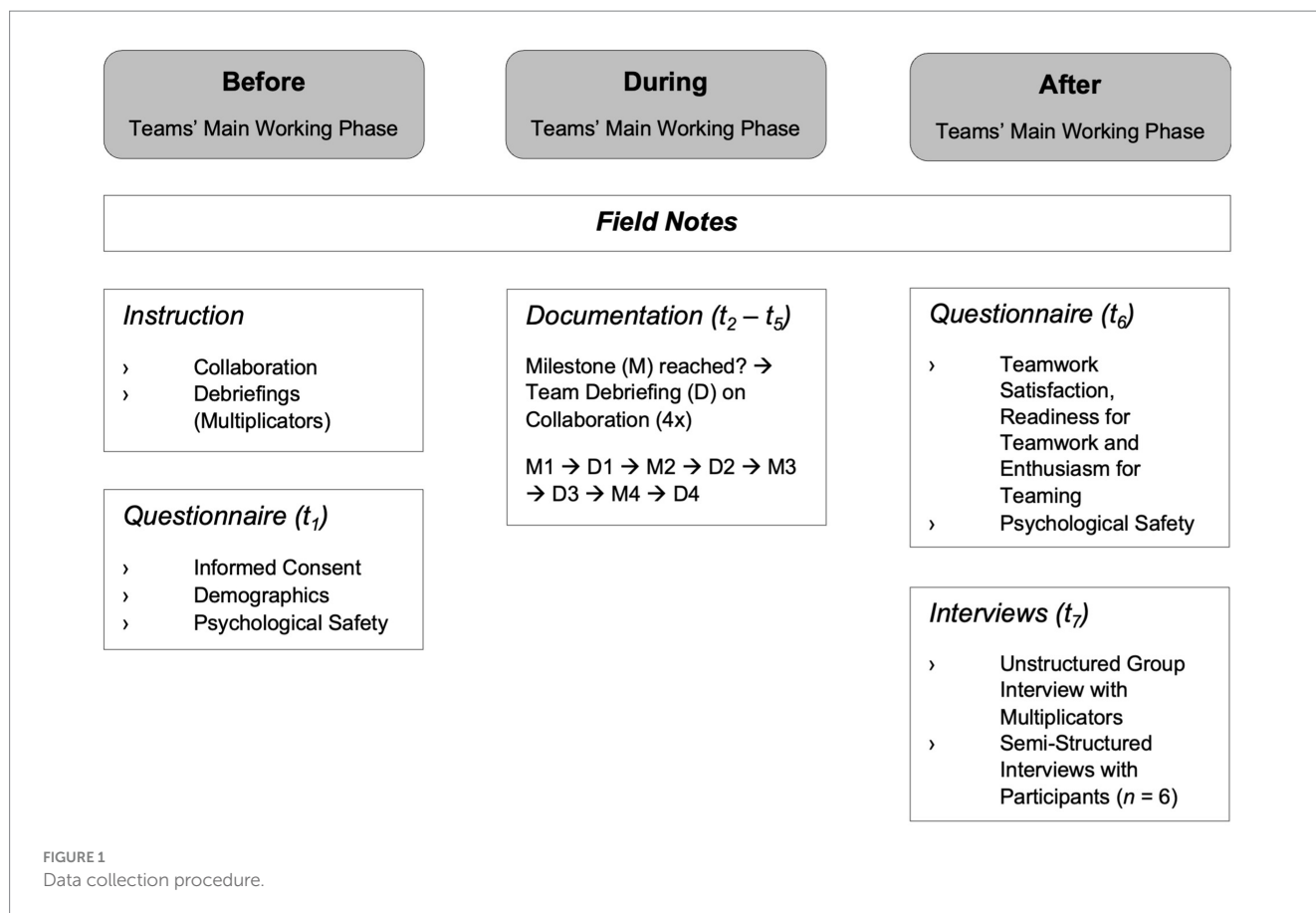
The student teams were asked to conduct a debriefing about their collaborative working phase when they reached a milestone (representing a performance episode, see Eddy et al., 2013). A total of four milestones were set (M1–M4, see Figure 1), so each student team had four debriefing sessions (D1–D4, see Figure 1). The multipliers documented these sessions with the help of a template containing information on setting (e.g., online or in person, all members attending, duration), a Behaviorally Anchored Rating Scale for student collaboration as well as an open field for notes and appendices. The debriefing was structured in accordance with the typical three debriefing phases (description, analysis and application, e.g., Johansson et al., 2017; Steinwachs, 1992). The BARS was particularly developed for student collaboration (Schürmann et al., 2023). It is based on a systematic literature review (Schürmann et al., 2024) and observation (critical incident analysis) of student teams performing a collaborative problem-solving task (Schürmann et al., 2023). It covers 10 dimensions of collaboration (e.g., *planning activities, gathering and sharing information*) with 5-point scale with anchors for high, medium, and low level of the respective sub-facet. The "5," "3" and "1" ratings each comprise four behavioral anchors, reflecting excellent, satisfactory and poor behaviors, respectively.

In each debriefing, the teams

- 1 talked about their prior collaborative working phase and rated their collaboration quality as a team on the BARS dimensions (description),
- 2 used this rating to discuss reasons for poor, satisfactory or excellent performance (analysis),
- 3 summarized and generalized their experiences for future situations (application).

2.2.3 After teams' main working phase

To evaluate the course concept, we collected students' perceptions, for instance, how helpful they perceived the regular reflections. On 10-point Likert scales ranging from 1 = *not useful* to 10 = *very useful*,



students were asked to rate the perceived usefulness with respect to (a) their collaboration in the team and (b) for them personally. In addition, a global rating regarding their satisfaction with the regular reflections was considered (1 = *not satisfied*, 10 = *very satisfied*). We used stimulating questions and open-text fields enabling the students to share further insights with us (e.g., *Did anything change after the reflections? Did the collaboration go better or worse than before?*). Moreover, the second questionnaire comprised questions on teamwork satisfaction, readiness for teamwork, enthusiasm for teaming, and again psychological safety ($\alpha_{t6} = 0.87$). While the former variables were collected for exploratory purposes, the latter (psychological safety) enables pre-post comparison. Teamwork satisfaction was measured based on Tseng et al. (2009). The 10 items (5-point likert scale, 1 = *completely disagree* to 5 = *completely agree*) were translated to German and adapted to the current setting. An example item is “I have benefited from interacting with my teammates” ($\alpha = 0.96$). Readiness for teamwork and enthusiasm for teaming were measured with German translations of the scales developed by Eddy et al. (2013). Each of the scales includes three items answered on a 5-point likert scale ranging from 1 = *completely disagree* to 5 = *completely agree*. Example items are “Being a part of this team will help me be a more effective member of teams in the future” (Readiness for Teamwork, $\alpha = 0.86$) and “Being on this team has decreased my enthusiasm for working in team settings in the future” (Enthusiasm for Teaming, reverse coded, $\alpha = 0.94$). The reported alpha values indicate strong reliability of the scales used in the study.

To gain deeper insights into students’ learnings and perceptions, we further conducted interviews. On the one hand, we met with the four students who led the debriefings in their groups to discuss their

experiences as multiplier. This unstructured group interview took place at the end of the semester and endured around 20 min. On the other hand, all course participants were asked to participate in semi-structured interviews a few weeks after course completion. Six of the 15 students participated. Interview participants were offered the option to receive either research participation credit or monetary reimbursement for their involvement in the interviews. All interviews were conducted by the same interviewer and lasted approximately 20 min to 1 h. The interview guideline comprised questions on prior experiences with group work (serving as opening questions) and reflection and questions regarding the conducted debriefings using the BARS. For instance, we sought to delve deeper into whether the students found the reflections beneficial, examining the specific ways in which they were helpful or not, and the reasons behind their perceptions (e.g., *Did you find the reflections helpful with regard to your collaboration? If so, in what way? Or why not?*). Given that students are unfamiliar with the BARS format, we were interested in their handling with this scale. Sample questions included: *How did you find the process of using the scale? Please elaborate on aspects you appreciated, those you found beneficial, and any challenges you encountered.* Finally, participants were given the possibility to share their suggestions and recommendations concerning the course concept including regular reflections.

2.3 Data analyses

2.3.1 Quantitative data

One participant (member of team 1) did not complete the second questionnaire, reducing the sample size to $N = 14$

participants at t_6 . Prior to data analyses, some data preparation was done. Collaboration and psychological safety are team level variables. For collaboration, we added up the teams' ratings for each of the 10 dimensions resulting in a sum score for each debriefing (with 50 being the highest possible value; see Table 2). Using the Excel tool provided by Biemann et al. (2012) to calculate within-group agreement indices [r_{WG} and $r_{WG(j)}$], (James et al., 1984) and intraclass correlation coefficients [ICC(1) and ICC(2), Bliese, 2000], individual-level data for psychological safety was evaluated to justify aggregation to the team level. The $r_{WG(j)}$ values of 0.92 for t_1 and 0.89 for t_6 represent strong to very strong agreement within the teams (LeBreton and Senter, 2008). The calculated ICC(1) values being >0.50 for both measurement times can be considered to be large effects (Bliese, 2000; LeBreton and Senter, 2008). The ICC(2) values being >0.80 for both measurement times represent the reliability of the group means (Bliese, 2000) and can be interpreted as good. Hence, we concluded that data aggregation was justified. IBM SPSS Statistics 28 was used for the analysis of quantitative data, primarily for conducting descriptive statistical analyses to summarize and present key patterns.

2.3.2 Qualitative data

All interview recordings were transcribed automatically using the transcription software f4x (Audiotranskription, 2025). After checking and revising the transcripts manually, they were uploaded to MAXQDA (VERBI Software, 2019) together with students' answers to open-ended questions, debriefings' documentations as well as the field notes. We used qualitative content analysis (Schreier, 2012) to code and structure the data using the functionalities of MAXQDA. The coding system was developed deductively based roughly on the Kirkpatrick model of training evaluation (e.g., Kirkpatrick, 1996; Paull

et al., 2016; Praslova, 2010) and adapted inductively. This model has also been used in higher education settings and provides a useful starting point for evaluating educational interventions and programs (Paull et al., 2016; Praslova, 2010).

The combined deductive-inductive approach facilitated the identification of recurring concepts and categories related to the participants' experiences and perceptions. The coding system was iteratively adapted and refined through several discussions. Afterwards, the first author coded the data accordingly. In addition, a second independent coder (psychologist) was introduced to the coding scheme and coded parts of the material (~25%). For that we selected the group interview, three interviews and document including the open-text field answers of one group so that responses of all four teams were covered. The intercoder agreement of approximately 60% displayed fair to moderate agreement. While discussing sections of non-agreement, it became evident that the second coder would have needed more context knowledge and longer training. To ensure validity and trustworthiness of data interpretation, we conducted member checking by soliciting feedback from the interview participants (Motulsky, 2021). Three participants voluntarily engaged in this process. They were provided with a summary of the coding scheme in order to give a short feedback with having the following questions in mind: (a) *Do the described experiences match your own?*, (b) *Is there something you see differently or would like to add?*, (c) *Are there any misunderstandings or errors in the representation?*. Moreover, the three participants were asked to highlight the codes matching most to their experiences. In summary, they expressed overall agreement with the findings. Their feedback, along with discussions from the intercoder agreement process, contributed to the refinement of code

TABLE 2 Descriptive statistics of study variables.

Variable	Team 1		Team 2		Team 3		Team 4		Overall	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Team level										
Collaboration (D1–D4)	37.75		36.75		45.5		36.75		39.19	
D1	42		33		46		32		38.25	(6.85)
D2	45		40		45		34		41.00	(5.23)
D3	41		37		42		34		38.50	(3.70)
D4	23		37		49		47		39.00	(11.89)
Psychological safety										
t_1	6.43		4.47		6.36		5.66		5.73	(0.91)
t_6	4.24		4.82		6.71		6.10		5.47	(1.14)
Individual level										
Perceived usefulness (teamwork)	6.33	(3.22)	4.00	(2.58)	5.75	(1.89)	7.00	(0.00)	5.64	(2.30)
Perceived usefulness (personally)	8.33	(2.89)	3.25	(2.06)	4.50	(1.73)	7.00	(0.00)	5.50	(2.68)
Satisfaction with debriefings	6.00	(3.46)	6.25	(2.06)	7.00	(1.41)	7.67	(1.16)	6.71	(1.98)
Teamwork satisfaction	2.90	(1.13)	2.70	(0.70)	4.83	(0.13)	4.17	(0.51)	3.66	(1.12)
Readiness for teamwork	3.89	(1.07)	3.17	(0.33)	4.00	(0.86)	4.00	(1.00)	3.74	(0.82)
Enthusiasm for teaming	3.00	(1.00)	1.33	(0.67)	5.00	(0.00)	4.33	(0.88)	3.38	(1.65)

The table depicts mean values including standard deviations in brackets. The maximum scale values are: Collaboration = 50; Psychological safety = 7; Perceived usefulness, Satisfaction with debriefings = 10; Teamwork satisfaction, Readiness for teamwork, Enthusiasm for teaming = 5. The bold values represent the total values of the teams' collaboration on the BARS scales.

descriptions and the coding process. For example, the code definitions of “disagreements” and “uncertainty” have been refined and some codes were split into more specific categories. The revisions helped to increase precision and clarity in the data analysis. Finally, the first coder checked the coding of the whole material and adapted it based on the refinements made. The final coding scheme including code descriptions and examples is provided as [Supplementary material](#).

Taken together, the mixed methods approach allowed us to integrate both quantitative and qualitative data sources, enhancing the depth and validity of our findings. Data triangulation was employed by combining quantitative data with qualitative insights from field notes, open-text responses, and interview data, providing a more comprehensive and reliable interpretation of the collaborative processes.

3 Results

The descriptive statistics of study variables are presented in [Table 2](#). A positive development in students’ self-assessed collaboration over the semester is observed in teams 2, 3, and 4. Additionally, an increase in psychological safety can be observed in these three teams. Taking the individual level data into account, the usefulness of the debriefings (for both teamwork and personal development) was rated quite differently by the students, as indicated by the range and variance in responses. Despite this variability, the data suggests a general tendency toward satisfaction with the debriefings. Furthermore, the mean values for teamwork satisfaction and enthusiasm for teaming are rather high in teams 3 and 4 but rather low in team 1 and 2. Whereas mean values for readiness for teamwork are comparable in teams 1, 3 and 4, perceived readiness for teamwork is notably lower in team 2.

3.1 Development of collaboration within the groups (RQ 1)

[Table 2](#) and [Figure 2](#) depict how student teams’ collaboration as rated during the debriefings developed over the semester. Findings indicate a positive development of students’ self-assessed collaboration quality for three of the four teams. Looking at the sub-facets of collaboration, the teams felt most confident in *Planning Activities*, *Handling Tools and Resources* and *Establishing a Positive Atmosphere and Cohesion*. On the other hand, *Monitoring and Reflecting*, *Managing Emotions* and *Individual and Joint Participation* were rated lowest (see [Table A1](#)).

Team 1 initially showed an increase in collaboration scores, which then fell intensively during the final debriefing. Taking the qualitative data into account, it becomes evident that the interviewees underestimated potential problems in the beginning and team members did not raise their issues in the debriefings until the last one. Two sub-groups (psychology, business) were formed with one sub-group feeling more actively involved in the group work than the other. Although this was thematized in the beginning, the students found it challenging to repeat on addressing this issue. For instance, one team member highlights that one hindering factor was that they got along very well personally. It seems like they did not want to risk the positive atmosphere with addressing negative issues. However, during the last debriefing they discussed their issues very honestly. One member of team 1 believes “*that if this last reflection meeting had not taken place, it would have remained unspoken in the end*.”

Team 2 also experienced the formation of two sub-groups (psychology, business). Prior to the first debriefing, the psychology students sought dialogue with the course instructor to complain about the unbalanced division of work and the passive behavior of their teammates. They raised their issues within the first debriefing leading to a slight improvement afterwards (see D2). However, coordination and joint participation remained a challenge until the course end (see

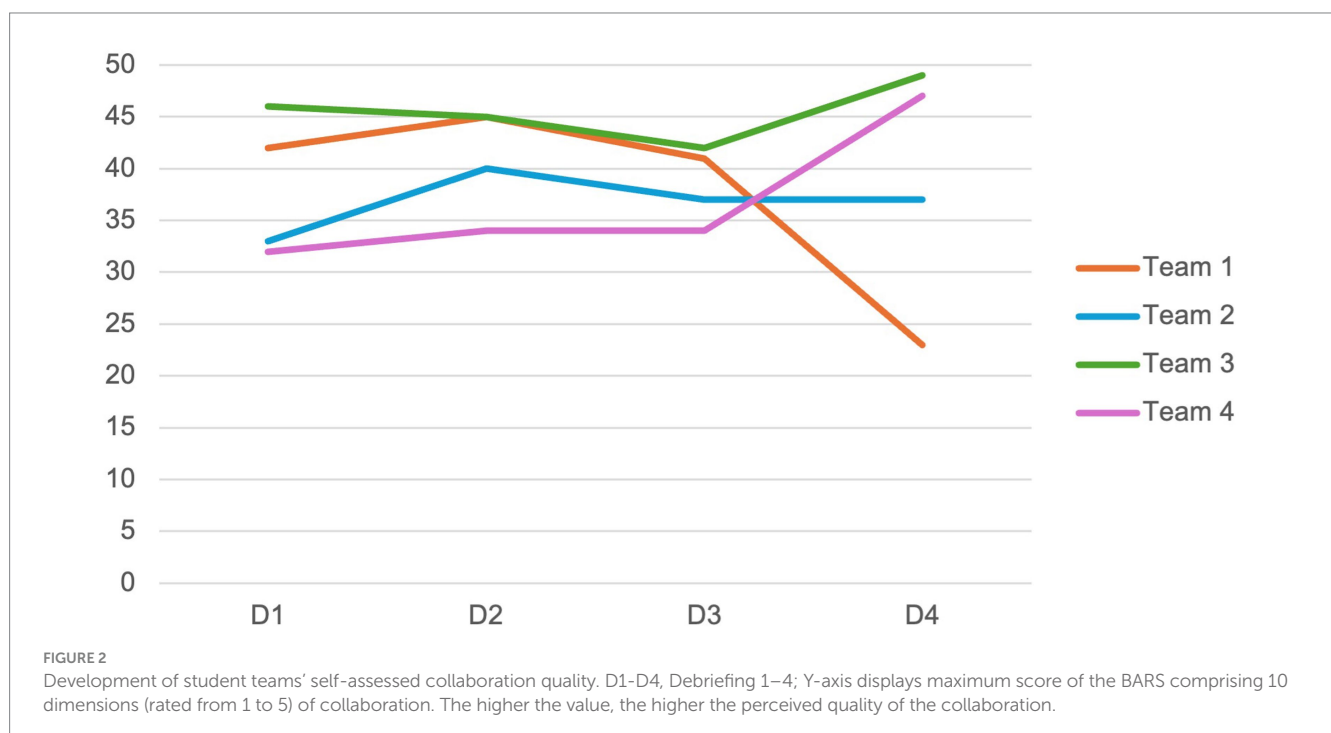


Table A1). For instance, during the group interview, interdisciplinarity came up as a potential reason.

Team 3 reported the highest self-assessed collaboration values. This is supported by the open text-field answers and interview data in which team members report that they functioned very well as a group.

Team 4 behaved rather unobtrusively during the semester. Despite having the lowest collaboration values for D1 to D3, they did not report on any challenges or troubles during group work during the semester. After initially increasing slightly, their collaboration values rose sharply toward the fourth debriefing.

In addition to the quantitative data gained in the debriefings, the analyses of the field notes, open-text fields and interview data further supports these findings. Specifically, two teams (teams 1 and 2) reported challenges and difficulties during group work. During coding, we also collected information on whether the debriefings lead to any modifications during student's group work. Whereas member(s) of team 1 experienced hardly any recognizable change other than poorer interaction and atmosphere, statements of team 3 and 4 suggest behavioral adoptions after the debriefings (*"In the first reflection it turned out that I contributed a little less to the group, which could be discussed immediately and improved for the next reflection,"* member of team 4). Furthermore, they reported an improved communication (e.g., member of team 3) and highlighted that the debriefings helped to immediately clarify disagreements (e.g., member of team 4). Taken together, the qualitative findings provide a nuanced perspective that complements the teams' self-assessments gathered during the debriefings.

3.2 Student views on the benefits of reflective practices for collaboration (RQ 2 + RQ 3)

During data analysis, it became evident that the responses to RQ 2 and RQ 3 were intertwined, making it challenging to draw clear boundaries between them. Consequently, we will combine our findings related to both research questions. In the upcoming sections, we provide a detailed analysis of the results, offering insights that not only address the research questions but also reveal additional challenges and limitations that extend beyond the scope of the original inquiries. While the responses to RQ 2 and RQ 3 are discussed in an intertwined manner throughout the text, Table 3 presents the core findings specifically related to each research question, allowing for a focused overview of the key answers. The selection of the most relevant codes was based on their frequency and the information gained during the member checking procedure.

In the following, particular attention is given to the insights from team 1 and team 3 allowing for a deeper understanding of how joint reflections function in varied collaborative contexts. Whereas team 1 encountered significant challenges during their collaboration, team 3 experienced a smooth and successful working process with no major problems. By comparing these two contrasting cases, we provide insights into how teams with different experiences benefit from the debriefing process. Due to the breadth of the coding scheme, it is not possible to present all categories in detail (please see the [Supplementary material](#) for a complete overview of the coding scheme). Hence, we will focus on the most frequently occurring categories, as well as those identified as particularly important during

the member checking process as these categories were highlighted by participants as central to their experiences.

3.2.1 Reactions and evaluations

On the one hand, students perceived the regular reflections as being helpful (*"I thought the idea of regularly sitting down together to reflect after the small group work was very good,"* member of team 3). Particularly, the debriefings helped them to not only focus on taskwork (*"... not just to focus on the organizational aspects, but also to really look at how we function as a team,"* member of team 1) and enabled open communication. Moreover, students emphasized that the debriefings helped to talk *"about things that otherwise might not have come up at all and would have been swept under the rug"* (member of team 2). Hence, students perceived the debriefings as awareness raising and as a good possibility to raise issues and clarify problems. This way, they were able to experience the perspective of others and could give each other feedback. For example, a team member of team 4 highlighted that he liked reflecting *as a group* as he was not experienced with reflective practices. He emphasized that that getting direct feedback from his teammates was extremely valuable for him.

On the other hand, two participants (members of team 3 and 4) emphasized that the debriefings were not necessary (*"But I do not think it would have been necessary for my group because, as I said, it went well in my group,"* member of team 3). Hence, it was mentioned that conducting the debriefing was rather a sense of duty than a necessity (*"Well, at least I felt it was more of a duty,"* member of team 4).

In team 1, which experienced problems during teamwork, participants underlined that during the debriefings miscommunications occurred. Additionally, they questioned whether the communication in the debriefings were open and honest (*"Negative about reflection: Question how honest the reflections were,"* member of team 1). In general, the students perceived the scope of the debriefings (i. e., frequency, length, interval) as too large (e.g., *"the number of reflections was too much,"* member of team 2).

This is also associated with participants feedback regarding the BARS which they perceived as too long leading to less attention and motivation during the debriefings (*"Because it was such a long questionnaire with lengthy questions and answer options, the motivation might have been a bit lower,"* member of team 3). The level of detail within the BARS is associated with advantages and disadvantages as emphasized by another member of team 3:

"This probably sounds a bit paradoxical, because I think it can be seen as both positive and negative that it is so detailed. Because it is so detailed and we have many different bullet points for each evaluation point, it makes the whole thing more tangible. You understand what the reflection form wants to convey and what it expects from you. But this does somewhat conflict with what I said earlier that it is indeed a lot of input."

Students claimed to like the type of scale and answered that the BARS was easy to understand. Moreover, the BARS provided a helpful orientation within the reflections (*"Positive about the reflection: Questionnaire with scales for orientation. This way, the team has a common reference point,"* member of team 1). Yet, regarding the BARS, some items were perceived as inappropriate for the context or redundant. Team 1, which experienced challenges,

TABLE 3 Essential Results in Response to the Research Questions.

RQ 1: How does collaboration of student teams develop over time when integrating regular debriefings in course design?	
Quantitative data	See Figure 2, Table 2, Table A1
Qualitative data (codes)	Debriefings led to...
	• improved communication
	• poorer interaction and atmosphere
	• no/ hardly any recognizable change
	• adapting and optimizing the way of working
	• immediate clarification of disagreements
RQ 2: How do debriefings help to foster collaboration in an educational setting?	
Qualitative data (codes)	During Teamwork, debriefings helped to...
	• not only focus on taskwork
	• raise awareness
	• enable open communication
	• raise issues and clarify problems
	• experience the perspective of others
	Personally, debriefings helped to...
	• understand the importance of open communication/ raising issues
	• get ready for groupwork
	• identify own strength and weaknesses and work on them
	• analyze own behavior in groups
RQ 3: Which features of debriefings are valuable to students in higher education context?	
Qualitative data (codes)	The Behaviorally Anchored Rating Scales (BARS) were perceived as valuable because they...
	• provided helpful orientation (common ground)
	• were easy to understand
	• are very detailed allowing to learn more about the dimensions of collaboration

The table provides a summary of key answers to the research questions (RQs). For a detailed discussion, the reader is referred to the main text. A comprehensive list of all codes generated from the qualitative data is available in the [Supplementary material](#).

further indicated some disagreements while choosing the level of scale. The team member interviewed (a multiplier) attempted to resolve these issues by explaining the reasoning behind her suggestions, but received no response, leaving her uncertain about how to proceed.

3.2.2 Personal take aways

We asked students on whether they learned something from the debriefings. Whereas some students explicitly stated that they do not think the reflections led to a personal learning gain (e.g., members of team 3), several personal take-aways from the debriefings (e.g., statements on what they plan to apply to future work) became evident. Also, in interviews with participants who did not directly see a personal take away (members of team 3). As can also be observed with the quantitative data (see Table 1, readiness for teamwork), some

students claimed to feel better prepared for future collaborations and take insights for future collaborations with them. In particular, this code was found in interviews with members of team 1. Furthermore, the debriefings supported in analyzing one’s own behavior during groupwork and helped to identify own strength and weaknesses (“Personally, I was able to work on my strengths and weaknesses even better through reflections and continuously improve them,” member of team 4). Especially with the help of the BARS, students gained a greater awareness of collaboration. For instance, how collaboration works and what dimensions/facets are associated with it (“To be honest, I do not think I even thought about it before. If someone had asked me, ‘In your opinion, what dimensions of collaboration are there?’ I would never have been able to name so many“, member of team 3).

3.2.3 Determining factors and suggestions

Several statements of the students illustrate what can influence the quality and utility of debriefings. Most frequently, team dynamics like psychological safety were mentioned. For example, a member of team 1 summarized:

But I think it’s really important as a foundation that there is a certain level of trust within the group, that the group members are well-attuned, and that we can truly say what we think without trying not to step on anyone’s toes. Because that was the problem I had, feeling like I had to handle things with kid gloves, so to speak.

This impression is also confirmed by another member of team 1 saying “because even if you were honest, it wasn’t really accepted and denied.” Whereas team 1 experienced these issues, members of team 3 were very happy about their way of communicating openly.

The answers of several participants further suggest that the attitude of those involved are significant (“Half of the team did not want to allow much time for extensive reflection, which is why there was no intensive discussion. It was agreed rather quickly.”, member of team 1; “And I think it depends on the individual or perhaps also a bit on the team, how the reflection is carried out, how seriously it is taken and what extent it brings with it.”, member of team 3). This is closely linked to the code “extrinsic or intrinsic motivation,” suggesting the importance of having students who are intrinsically motivated to engage in the debriefing process.

Furthermore, students provided several suggestions for integrating regular debriefings and using BARS in higher education courses. They expressed a desire for early integration into study programs to help make sense of their experiences and better prepare for future group work. Ideally, reflections should occur during class time to prevent a negative attitude, as some students found it “annoying to take extra time for it” (member of team 2). Regarding the use of BARS, interview participants engaged in constructive discussions on how to optimize its application during debriefings. For instance, a member of team 3 proposed adjusting the sequence of the 10 scales in each debriefing or limiting the use to fewer scales per session.

3.2.4 Further insights

While several students viewed the role of the multiplier positively, and some multipliers felt it did not involve much additional work, others emphasized the responsibility it entailed and suggested that incentives would be necessary to sustain the role. One participant mentioned feeling somewhat dependent on the

multiplicator. As previously noted, sub-groups formed within two of the teams, which may explain why several participants perceived interdisciplinarity as a challenge during the project work. Further, it is worth noting that one interviewee (member of team 4) implied that knowledge of being part of a case study may have influenced the way they answered the BARS.

4 Discussion

The current study explored the application of structured and regularly conducted group reflexivity interventions (i.e., debriefings) in the context of higher education. Our findings indicate a positive development in students' self-assessed collaboration (quantitative data) over the semester in three of the four teams. Additionally, these teams also showed an increase in psychological safety. Triangulation with qualitative data (e.g., interviews, field notes) and comparing two contrasting cases help to understand this pattern. While one team initially demonstrated high collaboration, this progress was superficial, with genuine discussions only occurring during the final debriefing. Interviews and open-text responses revealed negative reactions toward the debriefings (e.g., unpleasant feelings, openness and honesty questionable). One interviewee described how reflecting on group work became progressively more uncomfortable over time indicating participants' discomfort with addressing problems and offering criticism. Furthermore, the interviewed members implied that the discussions during the debriefings were not transferred to group work (see code "no/hardly any recognizable change"). On the other hand, there was a noticeable prevalence of codes highlighting personal takeaways (e.g., the importance of open communication and addressing issues). This suggests that while the debriefings may not have been perceived as beneficial for teams' collaboration, they did appear to offer personal value to the interviewees. The individual-level data in [Table 1](#) also supports this observation [see mean values for Perceived Usefulness (personally) and Readiness for Teamwork, team 1].

In contrast, another team consistently reported the highest self-assessed collaboration scores throughout. This trend is corroborated by qualitative interview data, which indicate that this team did not perceive much need for the debriefings, as their collaboration was functioning very well. Nevertheless, interviews and open-text responses from this team revealed several codes related to positive reactions toward the debriefings (e.g., awareness raising, not only focusing on taskwork). Yet, data of this team implied little learning gain on the individual level [see code little learning, [Table 1](#) mean values for Perceived Usefulness (personally) rather low]. In summary, the findings suggest that teams with well-functioning collaboration tend to question the necessity of debriefings and struggle to perceive a personal benefit from them.

Taking into account the data from the remaining two teams, one team started with low collaboration ratings, facing issues due to unequal work distribution. Data suggests these issues improved after being addressed in the first debriefing but then declined again over time. In addition, the ratings for perceived usefulness, teamwork satisfaction, readiness for teamwork and enthusiasm for teaming are lowest in this team. Notably, no member of this team participated in the interviews, which makes it challenging to fully understand or interpret the data and the underlying reasons for these low ratings.

While this group's members sought support from the course instructor due to early challenges, another team reported no issues despite low initial ratings. Their scores increased slightly at first and then sharply. However, this development should be interpreted with caution, as one interviewee hinted that team members may have deliberately tried to present "new material" during self-assessments.

The common issues found in group work across various settings, such as unequal participation and incompatible working styles (e.g., [Donelan and Kear, 2024](#); [Näykki et al., 2014](#)), were also evident in this study, confirming previous research on the persistence of these challenges. Although our findings suggest that incorporating debriefings offers students a valuable opportunity to raise issues during teamwork, they seem to encounter difficulties when it comes to engaging in critical discussions, as similarly observed by [Vizcarrondo \(2021\)](#). In line with the review by [Chan and Lee \(2021\)](#), the integration of regular reflection led to some negative attitudes among students (e.g., high workload) which in turn can lead to a loss of engagement with the reflective practice. These challenges may be even more pronounced in groups composed of students from different study programs, where conflicting schedules, competing commitments, and varying priorities (see [Hussein, 2021](#)) further complicate collaborative learning and reflective engagement. Unequal contributions during teamwork are a significant factor contributing to poor teamwork experiences ([Wilson et al., 2018](#)), particularly in online or hybrid settings, where communication technologies can make it easier for individuals to disengage ([Donelan and Kear, 2024](#)). In this study, this was also observed for two teams in which sub-groups were formed and team satisfaction was low. Keeping in mind that the reflection itself is also a collaborative task, a lack of engagement from some team members can be perceived as unequal contribution potentially increasing conflict potential within the team. Conflicts (process and relationship conflicts in particular) decrease positive team emergent states like trust or respect as well as hinder psychological safety ([Huerta et al., 2024](#); [Jehn et al., 2008](#)). This, along with other factors identified in this study, can influence the quality and effectiveness of debriefings. Especially, team dynamics such as psychological safety seem to play a significant role. Quantitative and qualitative data revealed that psychological safety remained high in the team which consistently reported the highest collaboration values (respectively even increased – which fits to the aspired goal of debriefings, see [Allen et al., 2018](#)), whereas it significantly declined over time in another team. The interview data suggest that members of the latter team felt not heard by their teammates when voicing concerns. Feeling heard describes "the feeling that one's communication is received with attention, empathy, respect and in a spirit of mutual understanding" ([Roos et al., 2023](#), p. 5). Although not much is known yet about how psychological safety gets destroyed ([Edmondson and Bransby, 2023](#)), it can be assumed that feelings of being unheard are likely to impact psychological safety and induce socio-emotional challenges in student collaboration. In line with [Kolbe et al. \(2020\)](#), we also observed that a low level of psychological safety (like observed in this team, i.e., team 1) is likely to undermine reflective learning conversations.

We further investigated the potential of BARS as a tool to support students' debriefings. In general, students claimed to like the type of scale as it provided a helpful orientation within the reflections. The level of detail within the BARS is associated with advantages and disadvantages. On the one hand, the detailed

descriptions helped students understand how collaboration works and the dimensions or facets associated with it, serving as a helpful orientation during debriefings. On the other hand, the BARS was perceived as too long, potentially leading to negative feelings such as boredom and frustration. Although we are not yet aware of any studies that have integrated the use of BARS into debriefings, our findings align with research on the use of BARS in various contexts. For instance, research by MacDonald and Sulsky (2009) suggests that the specific behavioral anchors in BARS are valued by both feedback providers and recipients. In higher education context, Ohland et al. (2012) created a one-page BARS, noting that students are less likely to respond conscientiously as the length of an instrument increases. Taken together, we believe that BARS have the potential to support students' reflective processes. Based on the findings depicted above, however, adjustments need to be made regarding the length and presentation of the BARS.

4.1 Strengths and limitations

This study transferred the well-established intervention of debriefings to the higher education context, while innovatively integrating BARS into the debriefing process. Using an embedded case study design with mixed methods, this study has a small sample size, which limits generalizability. However, it provides rich and nuanced insights into student collaboration. The analysis of contrasting cases has deepened our understanding of how teams with varying experiences benefit from the integration of regular reflections. Further strengths are data triangulation and methodological rigor (e.g., member checking) which enhanced reliability of the results by incorporating multiple perspectives and methods. Additionally, we combine insights of team research and educational research allowing for a nuanced understanding of the findings. In response to the call for incorporating qualitative data on team reflection (e.g., Gabelica et al., 2014), our article offers deep insights into the characteristics and quality of joint reflections in student teams. Moreover, our study complements previous research by employing a time-series design, which sheds light on the repeated support of student collaboration (cf. Eshuis et al., 2019), and by highlighting challenges when students are asked to engage with multiple perspectives during reflection (cf. Tan, 2021).

Despite these strengths, several limitations must be acknowledged. A primary limitation is the potential for self-selection bias, as participants who chose to engage in the interviews were mainly psychology students. This may have skewed the results and limit the transferability of the findings. Considering the formation of sub-groups, it would have been valuable to gather perspectives from both "sides" for a more balanced view. Another limitation is the heterogeneity of the teams, as differences in composition and experience, for example, may have influenced the findings. To ensure fairness, we decided that all teams should undergo the debriefing intervention. However, without a control group, the study lacks a comparative baseline. Due to the small sample size and the focus on a few specific cases, the results cannot be easily extrapolated to a wider population. At the same time, based on our experience working with a wide range of student teams over many years, we have no reason to believe that the insights gained here are not relevant to other similar contexts.

4.2 Implications

The outcomes of the current study offer valuable guidance for researchers and higher education practitioners. In the following, we outline and discuss the implications derived directly from our findings.

In summary, instructors are encouraged to integrate regular reflections as this approach shows potential for improving students' collaborative experiences by prompting them to focus not only on task-related work but also on critically evaluating their own behavior within groups. Importantly, joint reflections should be integrated during course time. While holding four reflective sessions over the course of one semester was considered as *too much* by the students, we propose limiting this to two or three sessions of 30 min each. However, it is important to note that reducing the frequency of reflections may inadvertently shift the focus toward taskwork at the expense of teamwork. Without regular opportunities for students to reflect on team dynamics and improve their collaboration skills, the emphasis may fall primarily on completing tasks, sidelining the critical aspect of teamwork development. The optimal frequency and duration can vary considerably between teams, making a flexible strategy more effective than a uniform approach.

Previous research indicates that students require structure and guidance during joint reflections (Schürmann et al., 2025). To prevent superficial discussions and improve the depth of reflections, it is essential to provide clear and precise instructions. In this study, BARS emerged as a promising tool for enhancing student reflections. The students appreciated the clarity provided by the detailed descriptors of BARS, as it helped them better understand the components of effective collaboration. This suggests that BARS can serve as an effective framework for guiding students during self-led debriefings. Yet, students also perceived the BARS as too lengthy, which could discourage engagement during debriefings. A possible solution to this issue would be to shorten the BARS by not requiring students to answer all 10 dimensions in every debriefing. Instead, specific dimensions could be selected based on students' interests, needs, or temporal aspects. For example, early in the collaboration process, the *planning activities* dimension may be more relevant, whereas other dimensions might take precedence later. This flexible approach could make BARS more manageable while still maintaining its effectiveness as a tool supporting student reflection. Further research is needed to evaluate the impact of using BARS in this modified format, particularly in comparison to reflections conducted without such structured scales.

Despite the usefulness of BARS and instructing one member of each team (i.e., multiplier), the findings also suggest that students wished for additional support during the reflections. When teams encounter difficulties, they often struggle to articulate criticism constructively and manage differing viewpoints, indicating a need for stronger facilitation. On the other hand, when group work proceeds smoothly, students may find it difficult to appreciate the value of debriefings, perceiving them as unnecessary. These challenges suggest that educators should offer more guidance during debriefings, helping students navigate difficult conversations and reinforcing the benefits of reflection, even when collaboration appears to be successful. Yet, with regard to this, the question arises of how to do so in large classes with limited resources. Additionally, the student-teacher relationship itself may hinder students' engagement with reflective practices (Chan

and Lee, 2021). To address these challenges, it may be beneficial to have a more neutral facilitator support student teams during reflections. This could help mitigate any perceived power dynamics between students and instructors, fostering a more open and productive environment for reflection. A potential solution is to train student assistants in facilitating reflections. Given their close connection to the target group, they may be better equipped to understand students' experiences and thereby foster a more open, non-hierarchical dialogue. With appropriate institutional support, another potential solution could involve interdisciplinary collaboration. For instance, students from disciplines that focus on team dynamics, such as psychology, could be trained as facilitators for student teams from other disciplines. This approach would support not only reflective practice but also provide valuable real-world experience for facilitators.

Taken together, this study provides deep insights into the role of reflection in fostering students' collaborative experiences; however, challenging these observations in larger samples would provide an enhancing empirical basis for the findings. Future studies may also explore potential influencing factors, such as psychological safety or feeling heard, to better understand how these elements impact the quality and effectiveness of reflective practices. Gaining a deeper understanding of how reflection fosters collaboration and personal growth is crucial, particularly in preparing students for the demands of the 21st-century workplace. Expanding this research will help refine strategies that can better equip students for collaborative environments, addressing the pervasive gap in students' preparation for group work across various global contexts (e.g., Le et al., 2018; Wilson et al., 2018; Xu, 2024).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee for Non-Invasive Research on Humans in the Faculty of Society and Economics of the Rhine-Waal University of Applied Sciences. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

References

- Allen, J. A., Reiter-Palmon, R., Crowe, J., and Scott, C. (2018). Debriefs: teams learning from doing in context. *Am. Psychol.* 73, 504–516. doi: 10.1037/amp0000246
- Audiotranskription (2025). f4x Speech Recognition [Computer Software]. Available online at: <https://www.audiotranskription.de/en/f4x/> (Accessed November 24, 2023).
- Bedwell, W. L., Wildman, J. L., DiazGranados, D., Salazar, M., Kramer, W. S., and Salas, E. (2012). Collaboration at work: an integrative multilevel conceptualization. *Hum. Resour. Manag. Rev.* 22, 128–145. doi: 10.1016/j.hrmr.2011.11.007
- Bharuthram, S. (2018). Reflecting on the process of teaching reflection in higher education. *Reflective Pract.* 19, 806–817. doi: 10.1080/14623943.2018.1539655
- Biemann, T., Cole, M. S., and Voelpel, S. (2012). Within-group agreement: on the use (and misuse) of r WG and r WG(j) in leadership research and some best practice guidelines. *Leadership Q.* 23, 66–80. doi: 10.1016/j.leaqua.2011.11.006
- Bliese, P. D. (2000). "Within-group agreement, non-independence, and reliability: implications for data aggregation and analysis" in *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. eds. K. J. Klein and S. W. J. Kozlowski (Hoboken, NJ: Jossey-Bass), 349–381.
- Boet, S., Bould, M. D., Bruppacher, H. R., Desjardins, F., Chandra, D. B., and Naik, V. N. (2011). Looking in the mirror: self-debriefing versus instructor debriefing

Author contributions

VS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. DB: Supervision, Writing – review & editing. NM: Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. They acknowledge support by the Open Access Publication Fund of Rhine-Waal University of Applied Sciences.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. Portions of the text were edited with assistance from ChatGPT (model GPT-4, provided by OpenAI). The AI was used solely to enhance clarity and readability; factual content was verified independently by the authors.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1526487/full#supplementary-material>

- for simulated crises. *Crit. Care Med.* 39, 1377–1381. doi: 10.1097/CCM.0b013e31820eb8be
- Borge, M., and Mercier, E. (2019). Towards a micro-ecological approach to CSCL. *Int. J. Comput.-Support. Collab. Learn.* 14, 219–235. doi: 10.1007/s11412-019-09301-6
- Boud, D., and Bearman, M. (2024). The assessment challenge of social and collaborative learning in higher education. *Educ. Philos. Theory* 56, 459–468. doi: 10.1080/00131857.2022.2114346
- Chan, C. K. Y., and Lee, K. K. W. (2021). Reflection literacy: a multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review. *Educ. Res. Rev.* 32:100376. doi: 10.1016/j.edurev.2020.100376
- Child, S. F. J., and Shaw, S. (2019). Towards and operational framework for establishing and assessing collaborative interactions. *Res. Pap. Educ.* 34, 276–297. doi: 10.1080/02671522.2018.1424928
- Debnath, S. C., Lee, B. B., and Tandon, S. (2015). Fifty years and going strong: what makes behaviorally anchored rating scales so perennial as an appraisal method? *Int. J. Bus. Soc. Sci.* 6, 16–25.
- Dillenbourg, P. (1999). "Introduction: what do you mean by "collaborative learning"?" in *Collaborative learning: Cognitive and computational approaches*. ed. P. Dillenbourg (Amsterdam: Elsevier), 1–19.
- Donelan, H., and Kear, K. (2024). Online group projects in higher education: persistent challenges and implications for practice. *J. Comput. High. Educ.* 36, 435–468. doi: 10.1007/s12528-023-09360-7
- Dornan, T., Conn, R., Monaghan, H., Kearney, G., Gillespie, H., and Bennett, D. (2019). Experience based learning (ExBL): clinical teaching for the twenty-first century. *Med. Teach.* 41, 1098–1105. doi: 10.1080/0142159X.2019.1630730
- Eddy, E. R., Tannenbaum, S. I., and Mathieu, J. E. (2013). Helping teams to help themselves: comparing two team-led debriefing methods. *Pers. Psychol.* 66, 975–1008. doi: 10.1111/ppps.12041
- Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Adm. Sci. Q.* 44, 350–383. doi: 10.2307/2666999
- Edmondson, A. C., and Bransby, D. P. (2023). Psychological safety comes of age: observed themes in an established literature. *Annu. Rev. Organ. Psych. Organ. Behav.* 10, 55–78. doi: 10.1146/annurev-orgpsych-120920-055217
- Edmondson, A. C., and Lei, Z. (2014). Psychological safety: the history, renaissance, and future of an interpersonal construct. *Annu. Rev. Organ. Psych. Organ. Behav.* 1, 23–43. doi: 10.1146/annurev-orgpsych-031413-091305
- Eshuis, E. H., Vrugte, J., Anjewierden, A., Bollen, L., Sikken, J., and de Jong, T. (2019). Improving the quality of vocational students' collaboration and knowledge acquisition through instruction and joint reflection. *Int. J. Comput.-Support. Collab. Learn.* 14, 53–76. doi: 10.1007/s11412-019-09296-0
- Fanning, R. M., and Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simul. Healthc.* 2, 115–125. doi: 10.1097/SIH.0b013e3180315539
- Fischer, J. A., and Hüttermann, H. (2020). PsySafety-Check (PS-C): Fragebogen zur Messung psychologischer Sicherheit in Teams. *ZIS*. doi: 10.6102/zis279
- Fischer, F., Kollar, I., Stegmann, K., and Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educ. Psychol.* 48, 56–66. doi: 10.1080/00461520.2012.748005
- Fransen, J., Weinberger, A., and Kirschner, P. A. (2013). Team effectiveness and team development in CSCL. *Educ. Psychol.* 48, 9–24. doi: 10.1080/00461520.2012.747947
- Gabelica, C., Van den Bossche, P., De Maeyer, S., Segers, M., and Gijssels, W. (2014). The effect of team feedback and guided reflexivity on team performance change. *Learn. Instr.* 34, 86–96. doi: 10.1016/j.learninstruc.2014.09.001
- Georganta, E., and Brodbeck, F. C. (2020). Capturing the four-phase team adaptation process with behaviorally anchored rating scales (BARS). *Eur. J. Psychol. Assess.* 36, 336–347. doi: 10.1027/1015-5759/a000503
- Hadwin, A., and Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: exploring perspectives of social in self-regulated learning theory. *Teach. Coll. Rec.* 113, 240–264. doi: 10.1177/016146811111300204
- Harrison, M., Short, C., and Roberts, C. (2003). Reflecting on reflective learning: the case of geography, earth and environmental sciences. *J. Geogr. High. Educ.* 27, 133–152. doi: 10.1080/03098260305678
- Heymann, P., Bastiaens, E., Jansen, A., van Rosmalen, P., and Beausaert, S. (2022). A conceptual model of students' reflective practice for the development of employability competences, supported by an online learning platform. *Educ. Train.* 64, 380–397. doi: 10.1108/ET-05-2021-0161
- Hom, P. W., DeNisi, A. S., Kinicki, A. J., and Bannister, B. D. (1982). Effectiveness of performance feedback from behaviorally anchored rating scales. *J. Appl. Psychol.* 67, 568–576. doi: 10.1037/0021-9010.67.5.568
- Huerta, M. V., Sajadi, S., Schibeli, L., Ryan, O. J., and Fisher, M. (2024). An exploration of psychological safety and conflict in first-year engineering student teams. *J. Eng. Educ.* 113, 635–666. doi: 10.1002/jee.20608
- Hussein, B. (2021). Addressing collaboration challenges in project-based learning: the student's perspective. *Educ. Sci.* 11:434. doi: 10.3390/educsci11080434
- James, L. R., Demaree, R. G., and Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *J. Appl. Psychol.* 69, 85–98. doi: 10.1037/0021-9010.69.1.85
- Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., et al. (2015). Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educ. Technol. Res. Dev.* 63, 125–142. doi: 10.1007/s11423-014-9358-1
- Järvenoja, H., Volet, S., and Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: an instrument to measure the adaptive and social nature of the regulation process. *Educ. Psychol.* 33, 31–58. doi: 10.1080/01443410.2012.742334
- Jehn, K. A., Greer, L., Levine, S., and Szulanski, G. (2008). The effects of conflict types, dimensions, and emergent states on group outcomes. *Group Decis. Negot.* 17, 465–495. doi: 10.1007/s10726-008-9107-0
- Johansson, E., Lindwall, O., and Rystedt, H. (2017). Experiences, appearances, and interprofessional training: the instructional use of video in post-simulation debriefings. *Int. J. Comput.-Support. Collab. Learn.* 12, 91–112. doi: 10.1007/s11412-017-9252-z
- Johnson, D. W., Johnson, R. T., and Smith, K. A. (2014). Cooperative learning: improving university instruction by basing practice on validated theory. *J. Exc. Univ. Teach.* 25, 1–26.
- Jones, C., Volet, S., Pino-Pasternak, D., and Heinimäki, O.-P. (2022). Interpersonal affect in groupwork: a comparative case study of two small groups with contrasting group dynamics outcomes. *Frontline Learn. Res.* 10, 46–75. doi: 10.14786/flr.v10i1.851
- Kaplan, A. (2008). Clarifying metacognition, self-regulation, and self-regulated learning: What's the purpose? *Educ. Psychol. Rev.* 20, 477–484. doi: 10.1007/s10648-008-0087-2
- Keiser, N. L., and Arthur, W. (2021). A meta-analysis of the effectiveness of the after-action review (or debrief) and factors that influence its effectiveness. *J. Appl. Psychol.* 106, 1007–1032. doi: 10.1037/apl0000821
- Kirkpatrick, D. (1996). Great ideas revisited. Revisiting Kirkpatrick's four-level model. *Training Dev.* 12, 54–59.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Hoboken, NJ: Prentice-Hall.
- Kolbe, M., Eppich, W., Rudolph, J., Meguerdichian, M., Catena, H., Cripps, A., et al. (2020). Managing psychological safety in debriefings: a dynamic balancing act. *BMJ Simulation Technol. Enhanced Learn.* 6, 164–171. doi: 10.1136/bmjstel-2019-000470
- Kollar, I., Fischer, F., and Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learn. Instr.* 17, 708–721. doi: 10.1016/j.learninstruc.2007.09.021
- Le, H., Janssen, J., and Wubbels, T. (2018). Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration. *Camb. J. Educ.* 48, 103–122. doi: 10.1080/0305764X.2016.1259389
- LeBreton, J. M., and Senter, J. L. (2008). Answers to 20 questions and interrater agreement. *Organ. Res. Methods* 11, 815–852. doi: 10.1177/1094428106296642
- Leopold, H., and Smith, A. (2020). Implementing reflective group work activities in a large chemistry lab to support collaborative learning. *Educ. Sci.* 10, 1–19. doi: 10.3390/educsci10010007
- Lou, Y., and Macgregor, S. K. (2004). Enhancing project-based learning through online between-group collaboration. *Educ. Res. Eval.* 10, 419–440. doi: 10.1080/13803610512331383509
- MacDonald, H. A., and Sulsky, L. M. (2009). Rating formats and rater training redux: a context-specific approach for enhancing the effectiveness of performance management. *Can. J. Behav. Sci.* 41, 227–240. doi: 10.1037/a0015165
- Merkebu, J., Kitsantas, A., Durning, S. J., and Ma, T. (2023). What is metacognitive reflection? The moderating role of metacognition on emotional regulation and reflection. *Front. Educ.* 8:1166195. doi: 10.3389/feduc.2023.1166195
- Moreland, R. L., and McMinn, J. G. (2010). Group reflexivity and performance. *Adv. Group Process.* 27, 63–95. doi: 10.1108/S0882-6145(2010)000027006
- Mortari, L. (2012). Learning thoughtful reflection in teacher education. *Teachers Teach.* 18, 525–545. doi: 10.1080/13540602.2012.709729
- Motulsky, S. L. (2021). Is member checking the gold standard of quality in qualitative research? *Qual. Psychol.* 8, 389–406. doi: 10.1037/qup0000215
- Näykki, P., Järvelä, S., Kirschner, P. A., and Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning: a process-oriented case study in a higher education context. *Int. J. Educ. Res.* 68, 1–14. doi: 10.1016/j.ijer.2014.07.001
- Newman, A., Donohue, R., and Eva, N. (2017). Psychological safety: a systematic review of the literature. *Hum. Resour. Manag. Rev.* 27, 521–535. doi: 10.1016/j.hrmr.2017.01.001
- Nguyen, Q. D., Fernandez, N., Karsenti, T., and Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Med. Educ.* 48, 1176–1189. doi: 10.1111/medu.12583
- Nokes-Malach, T. J., Richey, J. E., and Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educ. Psychol. Rev.* 27, 645–656. doi: 10.1007/s10648-015-9312-8

- Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., et al. (2012). The comprehensive assessment of team member effectiveness: development of a behaviorally anchored rating scale for self- and peer evaluation. *Acad. Manage. Learning Educ.* 11, 609–630. doi: 10.5465/amle.2010.0177
- Oshima, J., Oshima, R., and Kayagi, T. (2024). University students' knowledge creation practices in face-to-face and hybrid blended learning: development of epistemic views and perceptions of the Community of Inquiry. *J. Sci. Educ. Technol.* 17, 1–20. doi: 10.1007/s10956-024-10191-z
- Paige, J. T., Kerdolff, K. E., Rogers, C. L., Garbee, D. D., Yu, Q., Cao, W., et al. (2021). Improvement in student-led debriefing analysis after simulation-based team training using a revised teamwork assessment tool. *Surgery* 170, 1659–1664. doi: 10.1016/j.surg.2021.06.014
- Patel, H., Pettitt, M., and Wilson, J. R. (2012). Factors of collaborative working: a framework for a collaboration model. *Appl. Ergon.* 43, 1–26. doi: 10.1016/j.apergo.2011.04.009
- Pauli, R., Mohiyeddini, C., Bray, D., Michie, F., and Street, B. (2008). Individual differences in negative group work experiences in collaborative student learning. *Educ. Psychol.* 28, 47–58. doi: 10.1080/01443410701413746
- Paull, M., Whitsed, C., and Girardi, A. (2016). Applying the Kirkpatrick model: evaluating an interaction for learning framework curriculum intervention. *Issues Educ. Res.* 26, 490–507. Available at: <https://www.iier.org.au/iier26/paull.pdf>
- Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., and Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: effects of a peer feedback and reflection tool. *Comput. Hum. Behav.* 27, 1087–1102. doi: 10.1016/j.chb.2010.06.024
- Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in higher education. *Educ. Assess. Eval. Account.* 22, 215–225. doi: 10.1007/s11092-010-9098-7
- Quinton, S., and Smallbone, T. (2010). Feeding forward: using feedback to promote student reflection and learning—a teaching model. *Innov. Educ. Teach. Int.* 47, 125–135. doi: 10.1080/14703290903525911
- Retnowati, E., Ayres, P., and Sweller, J. (2018). Collaborative learning effects when students have complete or incomplete knowledge. *Appl. Cogn. Psychol.* 32, 681–692. doi: 10.1002/acp.3444
- Ritter, B. A., Small, E. E., Mortimer, J. W., and Doll, J. L. (2018). Designing management curriculum for workplace readiness: developing students' soft skills. *J. Manag. Educ.* 42, 80–103. doi: 10.1177/1052562917703679
- Robbins, S., and Hoggan, C. (2019). Collaborative learning in higher education to improve employability: opportunities and challenges. *New Directions Adult Cont. Educ.* 2019, 95–108. doi: 10.1002/ace.20344
- Rodgers, C. (2002). Defining reflection: another look at John Dewey and reflective thinking. *Teach. Coll. Rec.* 104, 842–866. doi: 10.1111/1467-9620.00181
- Roos, C. A., Postmes, T., and Koudenburg, N. (2023). Feeling heard: operationalizing a key concept for social relations. *PLoS One* 18:e0292865. doi: 10.1371/journal.pone.0292865
- Roschelle, J., and Teasley, S. D. (1995). “The construction of shared knowledge in collaborative problem solving” in Computer supported collaborative learning. ed. C. E. O'Malley (Cham: Springer), 69–97.
- Rowe, N. (2020). The great neoliberal hijack of collaboration: a critical history of group-based learning in tertiary education. *High. Educ. Res. Dev.* 39, 792–805. doi: 10.1080/07294360.2019.1693518
- Salmons, J. (2019). Learning to collaborate, collaborating to learn. Engaging students in the classroom and online (stylus). London: Routledge.
- Scager, K., Boonstra, J., Peeters, T., Vulperhorst, J., and Wiegant, F. (2016). Collaborative learning in higher education: evoking positive interdependence. *CBE Life Sci. Educ.* 15, 1–9. doi: 10.1187/cbe.16-07-0219
- Schippers, M. C., Edmondson, A. C., and West, M. A. (2014). Team reflexivity as an antidote to team information-processing failures. *Small Group Res.* 45, 731–769. doi: 10.1177/1046496414553473
- Schippers, M. C., and Rus, D. C. (2021). Optimizing decision-making processes in times of COVID-19: using reflexivity to counteract information-processing failures. *Front. Psychol.* 12:650525. doi: 10.3389/fpsyg.2021.650525
- Schreier, M. (2012). Qualitative content analysis in practice. London: SAGE.
- Schürmann, V., Spahn, T., Marquardt, N., and Bodemer, D. (2023). Evaluating peer collaboration in higher education: Behaviorally anchored rating scales. In: eds. P. Blikstein, J. van Aalst, R. Kizito, & K. Brennan. *Proceedings of the 17th International Conference of the Learning Sciences*. International Society of the Learning Sciences. 1904–1905.
- Schürmann, V., Marquardt, N., and Bodemer, D. (2024). Conceptualization and measurement of peer collaboration in higher education: a systematic review. *Small Group Res.* 55, 89–138. doi: 10.1177/10464964231200191
- Schürmann, V., Bodemer, D., and Marquardt, N. (2025). Applying debriefings in the context of higher education: How joint reflection fosters students' collaborative problem solving. *Soc. Psychol. Educ.* 28:62. doi: 10.1007/s11218-024-09991-3
- Schwab, D. P., Heneman, H. G., and DeCotiis, T. A. (1975). Behaviorally anchored rating scales: a review of the literature. *Pers. Psychol.* 28, 549–562. doi: 10.1111/j.1744-6570.1975.tb01392.x
- Slade, M. L., Burnham, T. J., Catalana, S. M., and Waters, T. (2019). The impact of reflective practice on teacher candidates' learning. *Int. J. Scholarship Teach. Learn.* 13:130215. doi: 10.20429/ijstl.2019.130215
- Steinwachs, B. (1992). How to facilitate a debriefing. *Simul. Gaming* 23, 186–195. doi: 10.1177/1046878192232006
- Tan, S. Y. (2021). Reflective learning? Understanding the student perspective in higher education. *Educ. Res.* 63, 229–243. doi: 10.1080/00131881.2021.1917303
- Tannenbaum, S. I., and Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Hum. Factors* 55, 231–245. doi: 10.1177/0018720812448394
- Tao, P. K., and Gunstone, R. F. (1999). Conceptual change in science through collaborative learning at the computer. *Int. J. Sci. Educ.* 21, 39–57. doi: 10.1080/095006999290822
- Tseng, H., Wang, C.-H., Ku, H.-Y., and Sun, L. (2009). Key factors in online collaboration and their relationship to teamwork satisfaction. *Q. Rev. Dist. Learn.* 10, 195–206. Available at: <https://lailima.hawaii.edu/access/content/user/jfifield/645project/hungwei.pdf>
- van Braak, M., Giroldi, E., Huiskes, M., Diemers, A. D., Veen, M., and van den Berg, P. (2021). A participant perspective on collaborative reflection: video-stimulated interviews show what residents value and why. *Adv. Health Sci. Educ.* 26, 865–879. doi: 10.1007/s10459-020-10026-7
- Veine, S., Anderson, M. K., Andersen, N. H., Espenes, T. C., Soyland, T. B., Wallin, P., et al. (2020). Reflection as a core student learning activity in higher education - insights from nearly two decades of academic development. *Int. J. Acad. Dev.* 25, 147–161. doi: 10.1080/1360144X.2019.1659797
- VERBI Software (2019). MAXQDA 2020 [computer software]. Berlin: VERBI Software.
- Vizcarrondo, T. (2021). Reflecting on reflections: debriefing in public relations campaign classes. *J. Public Relations Educ.* 7, 188–199. Available at: <https://aejmc.us/jpre/wp-content/uploads/sites/25/2021/05/Vizcarrondo-.pdf>
- Wildman, J. L., Nguyen, D. M., Duong, N. S., and Warren, C. (2021). Student teamwork during COVID-19: challenges, changes, and consequences. *Small Group Res.* 52, 119–134. doi: 10.1177/1046496420985185
- Wilson, L., Ho, S., and Brookes, R. H. (2018). Student perceptions of teamwork within assessment tasks in undergraduate science degrees. *Assess. Eval. High. Educ.* 43, 786–799. doi: 10.1080/02602938.2017.1409334
- Xu, S. (2024). Groupwork and collaborative learning: Chinese university students' struggles and strategies. *High. Educ. Res. Dev.* 43, 227–242. doi: 10.1080/07294360.2023.2215174
- Yang, T., and Choi, I. (2023). Reflection as a social phenomenon: a conceptual framework toward group reflection research. *Educ. Technol. Res. Dev.* 71, 237–265. doi: 10.1007/s11423-022-10164-2
- Yin, R. K. (2018). Case study research and applications. Design and methods. London: SAGE.
- Yukawa, J. (2006). Co-reflection in online learning: collaborative critical thinking as narrative. *Int. J. Comput.-Support. Collab. Learn.* 1, 203–228. doi: 10.1007/s11412-006-8994-9

Appendix

TABLE A1 Student teams' self-assessed collaboration during the debriefings.

		PA	COOR	MONR	HTR	GSI	CCG	ENRC	PAAC	ME	IJP	SUM
Team 1	D1	4	3	3	5	4	5	4	5	5	4	42
	D2	4	4	4	5	4	5	5	5	5	4	45
	D3	3	4	3	5	4	4	5	5	5	3	41
	D4	2	2	2	4	2	2	2	3	2	2	23
Team 2	D1	5	3	4	5	3	3	4	2	2	2	33
	D2	5	5	5	5	4	3	3	4	3	3	40
	D3	3	4	4	5	4	3	3	4	3	4	37
	D4	4	3	3	5	4	4	3	4	4	3	37
Team 3	D1	5	5	4	5	5	3	4	5	5	5	46
	D2	4	4	5	4	4	5	5	5	5	4	45
	D3	5	5	4	5	4	3	3	5	5	3	42
	D4	5	5	4	5	5	5	5	5	5	5	49
Team 4	D1	5	3	1	5	2	3	4	5	1	3	32
	D2	5	3	1	5	2	3	4	5	1	5	34
	D3	5	3	1	5	2	3	4	5	1	5	34
	D4	5	5	5	5	5	3	4	5	5	5	47

D1-D4, Debriefing 1–4; PA, Planning activities; COOR, Coordinating; MONR, Monitoring and reflecting; HTR, Handling tools and resources; GSI, Gathering and sharing information; CCG, Creating a common ground; ENRC, Elaborating and negotiating to reach consensus; PAAC, Establishing a positive atmosphere and cohesion; ME, Managing emotions; IJP, Individual and joint participation.