Check for updates

Gender Differences in Teamwork and Leadership within Virtual Engineering Laboratories

Marcel Schade, Claudius Terkowsky, Konrad Boettcher*, Stefan Inerle, Nils Kaufhold, Swetlana Herbrandt and Rajeenthan Sritharan

DOI: 10.1002/cite.202300242

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

To adapt to changes in digital transformation and related Learning and Working 4.0, laboratories in engineering education must leap forward and employ well-thought-out pedagogical designs. This includes digital laboratories that in return change social dynamics and interaction patterns of students. This contribution takes the first step of examining gender-specific teamwork and leadership of groups conducting a digital laboratory in fluid mechanics. Self and external evaluation of perceived teamwork of team members (TMs) and leaders (TLs) was analyzed. Furthermore, possibilities for improvements were evaluated. A mixed-methods analysis was performed with respect to gender and role to assess whether differences occur. Conclusions show considerable homogeneity between genders and roles and identify communication as most important point of improvement.

Keywords: Engineering education, Gender differences, Leadership, Teamwork

Received: December 18, 2023; *revised:* July 15, 2024; *accepted:* July 17, 2024

1 Introduction

Engineering education is currently facing the challenge of integrating Learning and Working 4.0 in order to catch up with the accelerating development of the digital transformation in industry as well as overall progress of society. Digitalizing laboratory courses bears several advantages such as making the invisible visible [1], increasing experimental freedom [2], and making remote, time-independent learning possible. Instructional frameworks like constructive alignment (CA) [3] help to structure learning experiences in a meaningful way. Even though, hands-on work cannot be readily replaced in remote or virtual laboratories, this void can be filled with new intended learning outcomes (ILOs). For instance, this can be realized by employing a higher level of autonomy of inquiry learning [4]. Felder and Brent [5] point out that detailed script-based experiments provide students with misleading experiences because professional life in engineering is about dealing with ill-defined problems while having incomplete information. Existing problems shall be solved quickly and independently. This conflicts massively with cookbook-like manuals that are often used to verify knowledge that has already been acquired in the preceding lecture.

The 13 fundamental objectives of the undergraduate engineering laboratories by [6] or the expanded list for Work 4.0 by [7] can be used as a baseline to develop a broad range of addressable ILOs. Still, Borrego and Cutler [8] state that the assessment of CA in engineering and related fields is underreported. This especially applies to interdisciplinary work, ethics as well as the design of learning environments and digital laboratories [9–11]. May [12] and Jahnke [13] emphasize that CA as well as its evaluation need to reflect the technical and social/socio-technical aspects of laboratories. Updating the pedagogical part of engineering education and laboratories requires looking further than employing well-thought-out educational designs; it must acknowledge where society stands and incorporate the requirements for society to move on.

Beddoes and Panther [14] report that teachers in engineering education rarely (if ever) think about teamwork, how teams are built, encouraged to work together, and what makes the work of several yield effective and efficient

[1]Marcel Schade https://orcid.org/0000-0001-9997-0768,
[1]Dr. Claudius Terkowsky
https://orcid.org/0000-0002-4571-5451, [3]Dr. Konrad Boettcher
https://orcid.org/0000-0002-7856-5234
(Konrad.boettcher@tu-dortmund.de), [2]Stefan Inerle,
[1,3]Nils Kaufhold https://orcid.org/0009-0002-4146-5480,
[2]Dr. Swetlana Herbrandt
https://orcid.org/0000-0002-4350-0748, [1]Rajeenthan Sritharan
[1]Center for Higher Education, TU Dortmund University, Vogelpothsweg 78, 44227 Dortmund, Germany.
[2]Statistical Consulting and Analysis, Center for Higher Education, TU Dortmund University, Vogelpothsweg 78, 44227 Dortmund, Germany.
[3]Department of Chemical and Biochemical Engineering, TU Dortmund University, Emil-Figge Str. 66, 44227 Dortmund, Germany.

outcomes. Especially considering gender and minorities, apparent differences can be spotted. Beddoes and Panther [14] find massive lacks in foresight, planning, and in some cases even action against better knowledge concerning teachers in engineering education setting up tasks and projects revolving around mixed-gender teams working together. Thus, a first step of creating assessments of laboratories engaging in Learning and Working 4.0 shall address the social and socio-technical elements of laboratory work.

In this study, teamwork is focused. A digital laboratory on fluid mechanics was used to evaluate the perceived teamwork of groups with randomly selected leaders to create a better understanding of team ratings. Differences concerning gender and leadership are focused and discussed. This is supposed to help teachers to better understand tendencies and possible biases when members of a team assess themselves or are assessed by their peers. This is especially important if peer-assessment influences the grading of large group projects spanning over several weeks – as it is the case in a subsequent course of the students of this study – where the teachers' insight into the group dynamics and development is hindered by the scope of the project itself.

2 Literature and State of the Research

Behavioral differences based on gender have been a longstanding focus in sociology for decades. In this context, a distinction between male (M) and female (F) students is drawn. Nevertheless, the field has become progressively intricate in light of societal advancements. Initially, it is essential to note that the effect size (Cohen's P) of the variable 'gender' concerning learning is $P = 0$ [15]. While existing literature on teamwork, leadership, and gender does not explicitly address the completion of digital laboratories, the only discernible distinction lies in the fact that students do not engage in hands-on activities within the physical laboratory but rather work on a computer.

Eagly and Carli [16] report in a meta-study a significant difference in styles of leadership. However, whether differences basing solely on gender truly exist, is still strongly debated as studies lead to inconclusive results. Jonsen et al. [17] suggest that research on gender differences in management are not conclusive "due to a nonlinear combination and interaction of causes – including stereotypes, expectations, actual gender differences, individual differences, and cultural context – …" [17, p. 564]. To approach this, leadership must be understood in the context of gender differences. Rowan-Kenyon et al. [18] find that leadership is coined and exerted differently. In a "male notation" it includes managing the teamwork, setting up and running meetings and overseeing the progress while the "female notation" revolves around facilitating collaborative tasks, taking responsibility for team dynamics and contributing to the team(-synergy). This correlates with [19] reporting that male team members (TMs) tend to rate female team leaders' (TLs') task-oriented leadership rather poorly while rating their leadership in terms of group-bonding and intergroup relationships highly. Stump et al. [20] and Rosch and Collier [21] agree with this notion as they found that female TMs are more engaged in learning collaborations and more motivated to lead groups than males are. The leadership-self-efficacy of male and female does not differ according to [21]. This is supported by De Paola et al. [22] who report on a study conducted on leadership at an Italian university. They found female led teams to be performing significantly better. However, male and female leaders themselves do not perform significantly different in terms of grades. The authors remark regarding the better team performance: "[W]e find that the positive effect of a female led team is mainly driven by the better performance of team members (TMs), with female team leaders (TMs) who react more to female leadership, while the performance of female leaders is slightly worse" [22, p. 20]. The authors attribute this to a higher effort by females to organize and lead the team at the expense of time they could have used for self-improvement otherwise.

Regarding evaluation, De Paola et al. [22] found that male TMs tend to evaluate female leadership worse than female TMs. Lastly, the study states that female TLs are aware of their effort as leaders but lack some awareness on their efforts' effectiveness. Thus, basing on effort put in only, they tend to evaluate their peer's efforts worse than male TLs do. These assessments are in good agreement with Strom and Strom [23]. They report in a study on High School students in the US that self and external evaluation of teamwork in study groups of female TMs is higher than that of male students. This is supported by Rosch et al. and Kaufman et al. [19, 24] concluding in their studies about leadership and gender in engineering education that females are rated higher than males in self- and external evaluations. However, both also report that the difference between self- and external rating with respect to gender is not significant.

Furthermore, Kaufman et al. [24] report that inflated self-ratings, in case the peer-assessments are relevant for grading, are quite rare. On the contrary, deflated self-ratings are much more prevalent and should be accounted for by weighing external ratings as grades awarded by the teacher strongly correlate with the peer-assessment. Strom and Strom [23] also remark the importance of heterogeneous grouping to encourage learning from one another, raising female awareness on their skills. Furthermore, knowing that peers evaluate each other also motivates students to participate and make an effort due to the personal responsibilities. However, teamwork dynamics are often disregarded in higher education as Beddoes and Panther [14] report in an interview study with teachers in different fields of engineering education. They found that teamwork is desired, but criteria for fruitful and healthy teamwork are –sometimes against the teachers' better knowledge – not followed.

Especially gender is not accounted for in teamwork in engineering education as several teachers in [14] report that

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

they do not think about it or even allow for female students experiencing bad teamwork arguing the earlier female students encounter these cases the more time they have to "toughen up" for their career. However, in these cases, students are not provided help or guidance to learn what to do in such situations, leaving issues unresolved. These issues are confirmed by [25] examining team dynamics and working roles taken on by STEM students and found that roles are often segregated by gender, e.g., male students perform hands on tasks and female students take notes and take care of organizational matters. Furthermore, female students often report about less inclusion within the group, respect, and being heard less. This shows the dire need for further recognition, improvement, and development.

Most studies are comparable to this work as they examined the influence of gender on the perception of teamwork and leadership of undergraduate engineering students [18–21, 24, 25]. One study also included postgraduate students and interviewed three teachers [25]. Another study was involved undergraduate economics students [22] and one focused on the perception of high-school students [23]. The social and ethnic backgrounds of the students were examined in four studies [20, 21, 24, 25]. The size of the cohorts varied between 78 [25] and 513 [20].

3 Research Questions

This contribution shall embark on the evaluation of students to find out whether gender differences within the perception of teamwork occur during the joint laboratory work, differentiated with respect to the role students held: TL or TM. Furthermore, possible discrepancies between self and external evaluation must be covered. The following null hypotheses are covered:

1. The self-evaluation of female students is not poorer than the self-evaluation of male students.
2. The evaluation of female teamwork by her peers does not differ from the evaluation of male teamwork by his peers.
3. Female students' evaluation of their TL's teamwork does not differ from male students' evaluation of their TL's teamwork.
   a. The evaluation of a female TL by a male TM is not poorer than the evaluation by a female TM.
   b. The evaluation of a male TL by a male TM does nt differ from the evaluation by a female TM.

The alternative hypothesis for 1. states that the self-evaluation of female students is poorer than self-evaluation of male students. The alternative hypothesis to 3a. expresses that the evaluation of a female TL by a male TM is poorer than the evaluation by a female TM. For the other hypotheses, the alternative states that there are significant gender differences in each evaluation.

Furthermore, the research question "What improvements in terms of teamwork do engineering students identify for themselves when working in cross-reality laboratories?" was asked to find out which aspects of teamwork in digital laboratories require more attention by supervisors.

4 Method and Materials

This part is structured into the sampling of students, a brief description of the laboratory course and the specific laboratory on which this work is based on. Finally, the questions asked are presented.

a. Sampling

Polled were 100 students (56 male, 44 female) of Chemical and Biochemical Engineering at TU Dortmund University within the winter term of 2021/2022. The experiment is part of a non-graded laboratory module held in the 3rd semester of study. The evaluation was conducted after the laboratory has been completed and was not required to pass the ungraded laboratory.

b. Laboratory Course

The laboratory experiment aims at providing a more autonomous learning environment utilizing scenario-based learning (SBL) [26, 27] to provide the opportunity for leadership to be necessary and thus developed. A Desktop-VR experiment was provided for download. Further information on development and implementation of the laboratory can be found in [26]. No script was provided and no pre-experimental colloquium was conducted. On the first day of processing time, the students received an unclearly-worded e-mail roughly outlining the students' tasks. In that e-mail, one student was selected to be TL and organize the group's time and working schedule as well as to correspond with the laboratory instructor. The laboratory was set up in an SBL approach with an open inquiry to foster the autonomy of the students. The instructor assumed the role of a supervisor within a company and the students were regarded as newly employed engineers remotely controlling a digital twin of a physical real setup.

The students are expected to extract all subtasks from the e-mail and to organize their conduction of the laboratory by themselves. They are supposed to measure a pump's characteristic curves and to implement a model based on their acquired knowledge from the course Fluid Mechanics I to compare these sets of data. The suction height and emptying time of a tank shall be determined. Considering the students being in their 3rd semester, their expertise in engineering was rather limited. Thus, they had to search for appropriate literature themselves with little prior knowledge on how literature search works. There were also fish swimming in the tank, but the students were not informed about this in advance. The emptying of the tank hid an ethical dilemma as the students were told, they remotely controlled a physical setup which implies the students would kill real fish inside the tank.

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Due to this setup, it was examined whether students would draw a connection between an abstract theoretical digital twin and the real-world implications of an actual digital twin. This topic was discussed during the reflection session following the students' presentation to enhance their awareness of hidden or abstract ethical dilemmas they may encounter in a highly digitized industry 4.0. Further details on instructional design and laboratory content are available in [26, 27]. More details on the ethical dilemma are offered in [27].

After the processing time ended, the students handed in a presentation for the board of managers. After possible revision, the students filled in an questionnaire asking them to rate their perceived performance. The students presented their results to the supervisor who after asking content-related questions ended the scenario and lead a focus group discussion on experiential learning. Lastly, a second questionnaire polled students to assess how well they could transfer acquired understanding, determined their perception of ILOs as well as their own accomplishments and development, and asked for free-text responses as to how they would have improved on their performance and what they would have wished for data collection [26].

The second questionnaire included quantitative 7-point Likert-scale questions on teamwork distinguishing between TL and TM:

1. How do you evaluate the teamwork of the TL?
2. How would you rate the teamwork of TM 1 (first name alphabetically)?
3. How do you evaluate the teamwork of TM 2 (first name alphabetically last)?

The data was analyzed using the statistics software R 4.2.3 [28]. Due to the data not being normally distributed, hypotheses were tested using the Brunner-Munzel test [29]. The test calculates the tendency of a group (A) to yield larger values than another group (B). This may be interpreted as whether a random person drawn from group A yields a greater value than a random person drawn from group B. The effect size (P) is determined by calculating the probability of this being the case.

To gain more insights as to what issues may have to be overcome, the free text question "What could I or my TMs have done better concerning teamwork?" was added.

The answers were evaluated by inductive qualitative content analysis according to [30]. Inductive categories are created based on the available content in which the answers are clustered. Applicable codes were set to be as small as a single word and as large as an entire answer. It is also possible that one answer contains several codes, certain phrases may even be coded into two separate categories.

5 Results

Results are divided into quantitative results obtained by statistical testing of hypotheses formulated in Sect. 3 and qualitative results obtained by analysis of free text answers.

Table 1. Statistical values obtained by the Brunner-Munzel test for the null hypotheses.

Hypothesis	p-Value	Test statistic	df	Effect size P	KI(P)
1	0.167	0.973	93.726	0.551	[0.447, 0.654]
2	0.552	0.595	186.832	0.552	[0.447, 0.654]
3	0.325	0.99	92.803	0.554	[0.446, 0.663]
3a	0.888	−1.244	27	0.377	[0.175, 0.58]
3b	0.71	0.377	26.33	0.377	[0.333, 0.742]

a) Quantitative Results

Received answers were not normally distributed. All questions followed the same pattern of evaluating perceived teamwork with distinctions made in sublevels concerning, role, gender as well as self- or external evaluation. Tab. 1 presents the results of each null hypothesis introduced in Sect. 3 including its empirical significance level (p), the Brunner-Munzel test statistic, degrees of freedom (df), effect size (P), and confidence interval of the effect size ($CI(P)$). Regarding the p-values, all hypotheses cannot be rejected to a significance level of 0.05. Thus, regarding (1), no significant difference between male and female self-evaluation regarding one's own teamwork is observed. Regarding (2), the external evaluation of perceived teamwork of male and female TMs, regardless of their specific role shows no significant differences. Regarding (3), the evaluation of the teamwork of TLs, regardless of their gender, differences whether male or female TM evaluate the TL are not significant.

To further specify this and investigate whether the effect holds true when differentiating between the TL's gender null hypotheses 3a and 3b were evaluated. Null hypothesis 3a yields a comparatively low effect size of $P = 0.377$. Thus, in this sample, the teamwork of female TLs is not evaluated worse by male TMs compared to female TMs. On the contrary even, one could argue that this null hypothesis may be worded towards the wrong direction. It may be possible that male TMs tend to evaluate a female TL's teamwork better than female TMs. Lastly, null hypothesis 3b switched the gender of the TL and this hypothesis cannot be rejected as well. Thus, no significant differences can be found between male and female TMs evaluation of the teamwork of male TLs.

b) Qualitative Results

In total 36 students (21 male, 15 female) answered the optional free-text question as to what they or their group could have done better in terms of teamwork. It should be noted that the ratio of male and female students between the all polled students and those who answered the free text

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

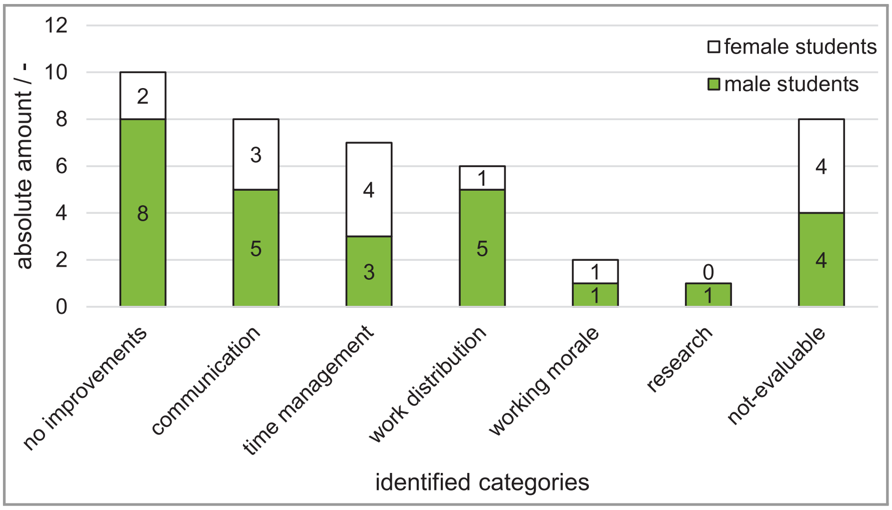

Figure 1. Absolute amount of codes attributed to each identified category by male and female students.

question is quite similar. Six categories have been created inductively into which 34 (23 by male students, 15 by female students) identified codes have been clustered. Replies were coded onto several categories if eligible. A seventh category collects answers that were not evaluable (ambiguous or not answering the question). Ambiguous responses like "–", "X", "/" could not be distinguished between "no answer" or "no improvement". Answers that missed the question are, e.g., "Being less stressed out myself", "I was a lot slower than my TMs. I got to get faster." The structuring differentiates for global amounts, and gender. Fig. 1 shows the total amount of codes differentiated by gender. Tab. 2 displays the relative amount of codes per answers of each gender. Firstly, the identified categories show how broadly the term teamwork may be defined.

The overall most common category is "no improvements". It is striking that about every third male student answered that while not even every fifth female student did. The students remarked positive aspects like even work distribution or synergy effects of group work instead of critiquing issues or simply stated that everything was fine the way it was.

The next most common category is "communication", which is rather evenly distributed between male and female students. The relative frequency the answer arose lies in a similar range with about 27 % and 21 % for male and female students, respectively. Besides the short answer "(better) communication" which was replied three times, students also remarked being more active in scheduling, discussing more often, making plans and arrangements more quickly.

Regarding the relative frequency time management was brought up, female students were considerably more concerned with time management as a part of the teamwork that should be improved. Short answers like "time management", "scheduling" were referred to four times. Further remarks were being more active in scheduling (c.f. communication), starting earlier and timely planning.

The fourth category is "work distribution". The category is mostly referred to by male students – absolute and relative frequency – with answers like better distributing work and offering each other help more often, quicker planning and distributing or criticizing another TM for not being active enough.

Working morale was only referred to twice as students – one male and one female – replied that they could improve their commitment and working approaches. Lastly, one male student replied that the group should do more research to improve their teamwork.

6 Discussion and Outlook

Regarding the quantitative results, all five null hypotheses posing no significant differences within the evaluation of teamwork, could not be rejected. Considering the effect sizes for the hypotheses 1, 2, 3, it can be observed that they are close to 0.5. For the data at hand, only small gender differences regarding leadership and the perception of teamwork within the digital laboratory were observed. The absence of significant differences between self- and external ratings between male and female students is in good agreement with [19, 24].

However, with an effect size of $P = 0.377$, null hypothesis 3a slightly deviates from this pattern of small effect sizes. Thus, for the observed sample, male TMs seem to tend to evaluate a female TL's teamwork better than female TMs do which is in agreement with [19, 22–24]. This could be due to male TMs valuing the higher effort female TL put into the social stability and organization of their group, as [22] stated, higher than female TMs do. Combining this with

Table 2. Relative distribution of the identified categories differentiated by gender (36 students, 21 male, 15 female, 34 evaluable codes; absolute frequencies in brackets, 23 by males, 11 by females).

	No improvements (10)	Communication (8)	Time management (7)	Work distribution (6)	Work morale (2)	Literature search (1)
Female (11)	18.18 % (2)	27.27 % (3)	36.26 % (4)	9.09 % (1)	9.09 % (1)	0.00 % (0)
Male (23)	34.78 % (8)	21.74 % (5)	13.04 % (3)	21.74 % (5)	4.35 % (1)	4.35 % (1)

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

male students' notion of leadership (c.f. [21]), it is possible that male TMs would not have put the same amount of effort into the team bonding and collaborating as their female TL did. Thus, they may be led to appreciate the effort more than a female TM would considering the female TM shares a more similar notion of leadership, and "expects" this kind of leadership. Considering only overall success was important (this laboratory is not graded), teamwork and coordination of tasks were crucial. Regarding hypothesis 3b, the effect size is again close to 0.5 leading to the same conclusion drawn for the hypotheses 1–3 which is also in good agreement with the results of [19, 24].

If no gender biases occur, it could be expected that no relevant gender differences between TMs arise if all students show each other that they are committed to complete the laboratory successfully. In an ungraded laboratory, the assessment is not superimposed with the unwillingness to make one's peers look comparatively poor, nor is there an incentive to downplay another's performance in order to achieve a better grade, even though this effect is shown to occur rarely [24]. However, as stated previously, these groups were not selected at random but could form freely with the only limitation being the size of three. Still, groups of two could be joined by another student introducing another unobserved variable into this research as whether students know each other and chose forming a group deliberately, was not assessed. Thus, assessing a lesser-known student may lead to distortions when compared to assessing a friend, especially if the students evaluate their peer's teamwork, as they perceived it.

The similarities of assessments may also arise due to differences in strictness when it comes to leadership and gender-specific leadership notations as [18] pointed out. Considering the students are of similar age and essentially in the same position as they conduct a mandatory laboratory, one could imagine that a student randomly chosen to lead may take on his or her role more or less strictly or seriously depending on their personality and the groups' dynamics in particular. Furthermore, this agrees with [19] that a TL seeing oneself as TL significantly influences the TMs' perception and recognition of the TL being the leader.

Regarding the qualitative part, male students seem to be pleased with the teamwork within their groups. They identified possible improvements mostly around communication and work distribution. Furthermore, it should be noted that one reply may be coded onto several categories. This is insofar relevant, as this overlap describes interactions within the groups. For instance, poor work distribution may be due to poor communication or vice versa. Whether the work distribution or the communication constitutes the main contribution of the issue could mostly not be discerned by the respective answers due to the answers being short. However, given the assumption of a stronger orientation on results rather than teamwork (c.f. [22]), a case for the work distribution could be made. This is further supported since the third main aspect "time management" is much less prevalent. Thus, this may support the notion that the male students' perspective of improving teamwork is less about the organizational part of teamwork but rather focuses on the results teamwork shall yield. This could be supported by the types of answers given by male students like the quote cited above "I was a lot slower than my teammates. I got to get faster." This is further supported by comparing these results to the general understanding of male and female students as to what leadership actually is.

Comparing this with the results by coding the replies of female students, one could argue that the organizational part of teamwork is more important for female students than the focus on the results as time management, communication make up about 63 % of their codes. However, a certain overlap is present again, lowering the possible effect. However, the limitation of this qualitative analysis has to be kept in mind since the absolute frequencies of several categories were rather small (<6). Thus, the relative frequencies should be understood as a tendency rather than reliable figures. The possibility of a certain preselection of students regarding the qualitative results cannot be excluded. The question was optional which means, students that are more engaged or those which have had particular experiences regarding their groups' teamwork (positive or negative), may have been more likely to respond. However, even mandatory replies may be circumvented by typing out a single character or "no answer", thus a self-selection of participants answering such questions must in many cases be accounted for. Thus, it should not be assumed that the sample of 36 students is representative for the 100 students that completed the laboratory. Still, possible (dis-)advantages that may or may not arise from certain group compositions should be brought to the instructors' attention as it influences a group's performance strongly.

In conclusion, this study found no significant difference between the self-evaluation of their teamwork by male and female students. Furthermore, no significant gender differences were detected in the evaluation of the teamwork of one's peers. Male and female TMs do not rate their TL's teamwork performance significantly different. The teamwork of male TLs is not significantly different rated by male and female TMs. All these effects also show rather small effect sizes. Furthermore, the teamwork of female TLs is not perceived to be significantly poorer by male TMs than by female TMs. On the contrary, the effect size in our sample indicates that the teamwork of female TLs may even be perceived better by male TMs than by female ones. This could be attributed to the larger effort in teambuilding and organizing female TLs put into their team. However, further in-depth coverage of such an effect is needed in future work.

The qualitative part found categories reaching from no improvements over communication, time management, work distribution, working morale up to literature research. A slight tendency may be discernible that female students identify possible improvements within organizational aspects of teamwork, especially time management and

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

communication while male students tend to identify aspects of teamwork catering towards the results to be achieved like work distribution and efficiency as possible improvements with the social part of teamwork being less prevalent. This also corresponds to the general notions of leadership male and female students hold. This raises the question, whether specific gender differences in the rating of teamwork — in this and in other studies — is largely influenced by different interpretations and thus expectations on leadership.

Future work will revolve around in-depth analyses of leadership, social dynamics as well as roles within teams to explore this question. An expansion of the research onto the whole laboratory module is possible as well as integration of a larger group work module that in part utilizes self- and external evaluation talks for a more precise assessment of students' performance influencing grading. On this basis an effort can be made to design group work and projects in a way that students are supported and presented a learning environment which caters to learning team working skills rather than just sitting together working without having the opportunity to improve or resolve conflicts properly if they arise.

Acknowledgment

This contribution was written within the scope of the joint project "CrossLab – Flexibel kombinierbare Cross-Reality Labore in der Hochschullehre: zukunftsfähige Kompetenzentwicklung für ein Lernen und Arbeiten 4.0", funded by Stiftung Innovation in der Hochschullehre (funding code: FBM2020-VA-182-3-01130).

Open access funding enabled and organized by Projekt DEAL.

Symbols used

CI(P)	[–]	confidence interval of effect size
df	[–]	degrees of freedom
p	[–]	empirical significance level
P	[–]	effect size

Abbreviations

CA	constructive alignment
F	female
ILO	intended learning objective
M	male
SBL	scenario-based learning
TL	team leader
TM	team member

References

[1] K. Boettcher, A. Behr, C. Terkowsky, *iJOE* **2022**, *18* (*14*), 114–132. DOI: https://doi.org/10.3991/ijoe.v18i14.35099

[2] P. Chan, T. Van Gerven, J. L. Dubois, K. Bernaerts, *CAEO* **2021**, *2*, 100053. DOI: https://doi.org/10.1016/j.caeo.2021.100053

[3] *Teaching for Quality Learning at University*, 5th ed. (Eds.: J. Biggs, C. Tang, G. Kennedy), The Society for Research into Higher Education, McGraw-Hill Education, Berkshire **2022**.

[4] *Labore in der Hochschullehre* (Eds.: C. Terkowsky, D. May, S. Frye), wbv Media, Bielefeld **2020**, 1.

[5] *Teaching and Learning STEM: A Practical Guide* (Eds.: R. M. Felder, R. Brent), John Wiley & Sons, New York **2016**.

[6] L. D. Feisel, A. J. Rosa, *JEE* **2005**, *94* (*1*), 121–130. DOI: https://doi.org/10.1002/j.2168-9830.2005.tb00833.x

[7] K. Boettcher, C. Terkowsky, T. Ortelt, I. Aubel, in *Proc. of the 20th Int. Conf. on Remote Engineering and Virtual Instrumentation (REV)* (Eds: M. E. Auer, R. Langmann), Springer, Cham **2023**. DOI: https://doi.org/10.1007/978-3-031-42467-0_56

[8] M. Borrego, S. Cutler, *JEE* **2010**, *99* (*4*), 355–369. DOI: https://doi.org/10.1002/j.2168-9830.2010.tb01068.x

[9] D. A. Martin, E. Conlon, B. Bowe, *Sci. Eng. Ethics* **2021**, *27* (*5*), 60. DOI: https://doi.org/10.1007/s11948-021-00333-6

[10] A. Van den Beemt, M. MacLeod, A. Van de Ven, *JEE* **2020**, *109* (*3*), 508–555. DOI: https://doi.org/10.1002/jee.20347

[11] J. G. Halls, C. Tomás, J. S. Owen, *EJEE* **2022**, *47* (*3*), 373–393. DOI: https://doi.org/10.1080/03043797.2021.2009775

[12] D. May, *iJOE* **2020**, *16* (*3*), 4–26. DOI: https://doi.org/10.3991/ijoe.v16i03.12849

[13] I. Jahnke, *IJIET* **2023**, *43* (*3*), 330–343. DOI: https://doi.org/10.1108/IJILT-05-2022-0105

[14] K. Beddoes, G. Panther, *EJEE* **2018**, *43* (*3*), 330–343. DOI: https://doi.org/10.1080/03043797.2017.1367759

[15] J. Hattie, Routledge, London, New York **2010**. https://www.visiblelearningmetax.com/Influences (Accessed on October 20, 2023)

[16] A. H. Eagly, L. L. Carli, *Leadersh. Q.* **2003**, *14* (*6*), 807–834. DOI: https://doi.org/10.1016/j.leaqua.2003.09.004

[17] K. Jonsen, M. L. Maznevski, S. C. Schneider, *EDI* **2010**, *29* (*6*), 549–572. DOI: https://doi.org/10.1108/02610151011067504

[18] H. T. Rowan-Kenyon, A. Coso, A. K. Swan, *JWM* **2012**, *18* (*2*). DOI: https://doi.org/10.1615/JWomenMinorScienEng.2012003031

[19] D. M. Rosch, D. A. Collier, S. M. Zehr, *JOLE* **2014**, *13* (*2*), 96–124. DOI: https://doi.org/10.12806/V13/I2/R5

[20] G. S. Stump, J. C. Hilpert, J. Husman, *JEE* **2011**, *100* (*3*), 475–497. DOI: https://doi.org/10.1002/j.2168-9830.2011.tb00023.x

[21] D. M. Rosch, D. Collier, *JOLE* **2013**, *12* (*2*). DOI: https://doi.org/10.12806/V12/I2/103

[22] M. De Paola, F. Gioia, V. Scoppa, *IZA Discussion Paper No. 11861*, **2022**. DOI: https://doi.org/10.2139/ssrn.4193376

[23] P. S. Strom, R. D. Strom, *Educ. Res. Eval.* **2011**, *17* (*4*), 233–251. DOI: https://doi.org/10.1080/13803611.2011.620345

[24] D. B. Kaufman, R. M. Felder, H. Fuller, *JEE* **2000**, *89* (*2*), 133–140. DOI: https://doi.org/10.1002/j.2168-9830.2000.tb00507.x

[25] S. Male, M. Marinelli, L. V. Kim, in *Proc. of the 8th Research in Engineering Education Symp.* (Ed.: B. Kloot), Research in Engineering Education Network, 292–301, **2019**.

[26] K. Boettcher, C. Terkowsky, M. Schade, *EJEE* **2023**, *48* (*5*), 949–971. DOI: https://doi.org/10.1080/03043797.2023.2182184

[27] C. Terkowsky, K. Boettcher, S. Grünendahl, in *Research and Innovations in Virtual Laboratory Practices in Engineering Education (Education Innovation Series)* (Eds.: P. Churi, P. Joshi), Springer Singapore, in press.

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[28] R Core Team (**2023**): R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.; Ara T (2022). brunnermunzel: (Permuted) Brunner-Munzel Test. R package version 2.0. https://CRAN.R-project.org/package=brunnermunzel.

[29] E. Brunner, U. Munzel, *Biom. J.* **2000**, *42* (*1*), 17–25. DOI: https://doi.org/10.1002/(SICI)1521-4036(200001)42:1%3C17::AID-BIMJ17%3E3.0.CO;2-U

[30] P. Mayring, in *Approaches to Qualitative Research in Mathematics Education* (Eds.: G. Kaiser, B. Sriraman), Springer, Dordrecht **2015**, 13.

 © 2024 The Author(s). Chemie Ingenieur Technik published by Wiley-VCH GmbH

15222640, 2024, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cite.202300242 by Technische Universitaet Dortmund, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License