Testing group composition within the studio learning environment

Kristine E. Callan,¹ Bethany R. Wilcox,¹ and Wendy K. Adams¹

¹*Department of Physics, Colorado School of Mines, Golden, CO 80401*

The Introductory Mechanics course at Colorado School of Mines uses a hybrid lecture-studio model. In studio, students work in groups of three through scaffolded problems and experiments for two hours twice per week. Research shows that rich learning can take place in small peer group settings with appropriately designed activities, but it is unclear to what extent group composition in terms of ability or gender impacts physics learning. To explore this question, we assigned half of each studio class to groups with mixed incoming physics ability (using FMCE pre-scores), and half to groups with matched incoming physics ability and gender. We evaluate the performance of each group type according to the students' scores on the FMCE (pre and post), common course exams, the CLASS, and a survey about their studio groups. We found neither composition type superior by any measure; but did learn that pace is a concern, group communication is sometimes limited, and that only 1% of students request a different group than assigned.

PACS numbers: 01.40.Fk

I. INTRODUCTION & BACKGROUND

We are interested in improving learning by investigating how group composition, with respect to ability and gender, can impact individual student performance. Interactive engagement techniques are now frequently used in college physics classrooms, which by definition require peer interaction; however, research results are mixed on how best to compose groups in this setting.

The seminal studies on group composition in physics were performed by Heller and colleagues [1, 2]. These studies utilized cooperative group learning, where students were assigned specific roles, when solving context rich problems. To increase group interdependence, a group problem solving portion was added to exams. Within this context, they found that groups of three are ideal. Though they did not report on comparisons of individual problem-solving outcomes, they also found that heterogeneous ability groups (a high, medium, and low student) produced better group problem solutions than homogeneous medium or low ability groups. The reason cited was that these groups often lacked necessary content knowledge to solve the challenging context rich problems. With regard to gender, they determined that all female or female majority groups produced better problem solutions than did two male one female groups. This work on group composition demonstrated that students in cooperative groups produced superior problem solutions compared to students in the traditional course, except in the area of conceptual understanding.

Recent work by Ding [3] and Harskamp [4] also found that female high school students interact in more productive ways and demonstrate higher quality of discussion and improved learning when working on physics problems with other females compared to male-female groups. Female-female groups also performed equal to or better than male-male groups in each of these studies.

The bulk of the literature on group work is focused on K-12 classrooms [5–7]. These studies find slightly different results on homogeneous vs. heterogeneous group

composition than Heller, but similar results regarding gender. A key outcome of these studies is that the benefits of group work depend on the type of activity and difficulty level; merely placing students together is not sufficient for promoting substantive gains in achievement.

Evidence for homogeneous vs. heterogeneous composition by ability is mixed. A synthesis by Slavin finds consistent positive effects for homogeneous grouping in reading and math for all ability levels [5]. However, Lou's [6] and Webb's [7] meta-analyses find that homogenous grouping is superior for medium-ability students but not for low-ability students. Webb found that narrow-range heterogeneous ability groups (medium to low or medium to high) create superior learning for all ability level students compared to wide-range heterogeneous groups.

All studies find, on average, superior performance and improved group interaction for females in female only groups. Arguably, there are continuums of gender identities and possible interaction types in a group environment, some of which are demonstrated more often by females than males (and vice versa). Differing from Heller, Webb found that females were at a disadvantage in groups with two females and one male. Gender effects were also shown to be stronger when activities were more complicated.

Not all group activities elicit improved learning. Recent work has focused on effective group interactions in order to better understand these group composition effects. Webb analyzed interactions to provide evidence for why learning occurs in some group compositions more than others. Her meta-analysis of 17 research studies in K-12 math showed that: giving explanations is positively correlated with achievement, receiving explanations is often not correlated with achievement, and receiving non-responsive feedback (just the answer or no response) is negatively correlated with achievement. In addition she summarized nine different types of explanation that elicit learning. These are consistent with work in PER promoting conceptual understating (e.g., activating prior knowledge, using specific examples to illustrate concepts, etc.).

Inspired by the prior literature, we aim to provide all students equal opportunities to interact in ways that are positively correlated with learning (i.e., giving explanations and receiving responsive feedback). In this paper, we present a study designed to measure the impact of group composition on the learning and perceptions of different populations of students in the Introductory Mechanics classes at the Colorado School of Mines.

II. CONTEXT & METHODS

All data for this study were collected from two consecutive semesters of Introductory Mechanics at Mines. This calculus-based course is required for all students regardless of their major and serves 450-550 students each semester. Mines, a highly selective engineering school, has a generally high-performing and well-prepared student population. This is reflected in the relatively high average Force and Motion Conceptual Evaluation (FMCE) [8] pretest score of 45%. Each course is roughly 31% women and 69% men. The majority of class time for this course is dedicated to “studio,” which meets for 2hrs, twice per week. Each studio section can accommodate 108 students in 36 groups of three, aided by five undergraduate teaching assistants (TAs) and one instructor. In studio, groups work at their own pace through a series of scaffolded physics problems, which help prepare them to complete a Problem Solving Skills Development or Experimental Skills Development exercise.

Historically (prior to 2016), students in Introductory Mechanics have been assigned to groups comprising one student from the top third of their studio section, one from the middle third, and one from the bottom third, as measured by their FMCE pre-test score or most recent exam grade. Groups were reassigned after each exam, and no attention was paid to gender composition.

To study the impact of group composition, we modified our approach: Students were assigned permanent studio groups of three once the final drop deadline for the course had passed, and each group was asked to set and reflect on their group norms (a set of behaviors they all agree are important to follow when working together in studio) four times during the semester. The course was randomly divided into two populations. For half of the students, group assignment was based only on FMCE pretest performance, with each group including one student from each ability level (i.e., high, medium, and low), with no consideration of student gender, as was done prior to 2016. We refer to groups in this population as Mixed. For the other half of the students, group assignment was based on both FMCE pretest performance and gender. Groups in this population were all single gender, and included members of adjacent ability levels. In other words, groups could contain both low and medium students (in any distribution) or high and medium students (in any distribution), but never both low and high students. We refer to groups in this population as Matched, meaning they are single gender and have less variance in incoming ability than the Mixed groups.

Groups were physically distributed in the classroom so as not to call attention to the grouping structure, particularly with respect to the existence of multiple single gender groups. Informal conversations with the TAs suggested that the distribution was successful as many TAs did not pick up on the grouping scheme, even as they circulated through the room helping all groups.

To inform continuous improvement of this course, we regularly utilize two standardized, research-based assessments – the FMCE and the Colorado Learning Attitudes about Science Survey (CLASS) [9]. The FMCE was administered on paper during class time at the beginning and end of the semester. The CLASS was administered online as a part of the first and last homework assignments of the course. In both cases, students received full credit for participation. Student responses to these assessments, along with their course exam performance, constitute the quantitative measures used to evaluate the impact of group composition on student performance. The quantitative data set presented here includes scores from two semesters ($N = 987$).

In addition to quantitative measures of students’ performance, we were also interested in students’ feedback on their experiences within the group learning environment of the studio. This feedback was collected only in the second semester of data collection via a survey administered twice in the semester as part of the required homework assignments in which students received full credit for completion. While this survey included seven Likert style prompts targeting students’ experiences within their group, the analysis here will focus on students’ responses to a final open-ended prompt – *Is there anything about how your studio group operates that you wish was different? Explain in the textbox.*

Based on the literature, we hypothesized that students from the Matched groups would, on average, outperform students from the Mixed groups. We expected that both the gender and ability-level composition would increase the frequency of the types of interactions that are positively correlated with learning.

Students’ responses to this question were qualitatively coded to identify emergent themes. Initial coding categories were developed by a single coder (BRW) to capture major themes in student responses. These categories were then operationalized and a second coder (KEC) independently coded 25% of responses to ensure the coding structure was valid and robust; initial percent agreement was greater than 75% for all student responses. The two coders then met to discuss their codes, refine code definitions, and resolve discrepancies. Final percent agreement between coders was greater than 90% for all codes.

III. RESULTS

Using FMCE pre- and post-test scores, CLASS pre- and post-test responses, combined exam score, and final course score, we compared the performance of students assigned to Mixed and Matched studio groups. We found no statistically significant differences between the

TABLE I. Coding categories and example quotes from students who explicitly stated they would not change anything about how their group functioned or were happy with their group. $N = 194$ and $N = 200$ students fell in this category for the first (S1) and second (S2) survey respectively. Codes are not exclusive or exhaustive but represent the most common themes.

Code	S1	S2	Example quotes
Helping	45	28	I think that my group is doing fine, we help each other and we don't move until each one of us got the point. I like the way we solve the problems together.
Pace/efficiency	18	21	I believe our studio group works efficiently while learning our objectives in studio.

scores of students in Mixed and Matched groups on any of these measures (Mann-Whitney U test [10], $p > 0.05$). This finding held for students at all incoming ability levels. Based on prior literature, we also examined whether this trend varied for men and women. While we found a statistically significant ($p < 0.05$) gender gap in all our quantitative measures except final course score, this gap did not vary significantly based on group composition.

To further explore and explain the null result found in the quantitative data with respect to group composition, we were interested in understanding what other aspects of the studio group environment were most salient, and potentially problematic, for all students independent of group composition. To do this, we used responses to the open-ended prompt given at two points during the second semester of data collection that asked students what they wished was different about their studio group. The major codes identified in students responses to this survey, with example quotes, are summarized in Tables I and II.

Of the 509 students who completed each survey, roughly three quarters (fall $N_1 = 387$, spring $N_2 = 378$) made comments more substantial than “None” or “No comment.” Of these, just over half ($N_1 = 194$, $N_2 = 200$) made comments stating explicitly that they would not change anything about their group or that they were happy with how their group operated. Some of these students ($N_1 = 77$, $N_2 = 57$) who were happy with their groups specifically discussed which aspects of how their group operated were most beneficial (see Table I). The most common justification ($N_1 = 45$, $N_2 = 28$) expressed the sentiment that the group prioritized the learning of the whole group by ensuring that all members got the help that they needed. A second common justification ($N_1 = 18$, $N_2 = 21$) was that the group worked efficiently or progressed through the studio at a good pace.

The remaining students ($N_1 = 193$, $N_2 = 178$) all made comments specifically citing one or more aspects of how their group operated that they wished were different or thought could be improved (see Table II). The most frequent comments ($N_1 = 63$, $N_2 = 59$) related to the pace at which students worked through the studio material. Some of these students wanted their group to work faster in order to get through more of the studio material, others wanted their group to work slower so that they could spend more time understanding the content. Others explicitly discussed how they or their group prioritized speed over understanding when work-

ing in the studio. Additionally, students also mentioned cases in which they perceived one member of their group (sometimes themselves) as working at a significantly different pace, either slower or faster, than the rest of the group. The prevalence of comments about pacing was likely exacerbated by the structure of the studio. In order to ensure that all students work for the entire period, the studios have been intentionally designed to be too long for most groups to finish within the allotted time; however, students are only graded on activities that the majority of the class (90% or more) was able to complete.

Another common class of critiques centered around the amount of communication ($N_1 = 36$, $N_2 = 20$) and collaboration ($N_1 = 27$, $N_2 = 34$) within the group. The majority of these students wanted more communication and/or more collaboration when working through the studio problems. In particular, some students reported that members of their group individually completed the problems and only came together to check the final answer. According to the students, problems with this approach included: leaving students stranded if they did not know how to approach the problem, having difficulty helping group members when everyone's approach was different, and not being able to learn from others' reasoning when that reasoning was not being externalized.

In addition to comments about pace, communication, and collaboration, students brought up a variety of other issues around general group dynamics. For example, some students ($N_1 = 21$, $N_2 = 24$) noted one member of their group (sometimes themselves) was not interacting within the group like the other members. Most often this was a student who was not contributing productively to the group, either through lack of engagement or disruptive behavior. Other students noted that while their group functioned well, they felt the ability level of the group as a whole was too low. Other general issues with group dynamics involved the desire to switch roles periodically, particularly with respect to who operated the computer, and difficulties resolving disagreements.

It is also worth noting several codes that did not appear very often. For example, with permanent groups, we anticipated that students would request to change or pick their groups; however, only a few students requested this ($N_1 = 10$, $N_2 = 6$). In addition to this relative lack of push back from the students, we asked the TAs in an end of semester survey whether and why they felt keeping a single group all semester was beneficial for the students'

TABLE II. Coding categories and example quotes from students whose responses included one or more aspects of the way their group functioned that they wished was different. $N = 193$ and $N = 178$ students fell in this category for the first (S1) and second (S2) survey respectively. Codes are not exclusive or exhaustive but represent the most common themes.

Code	S1	S2	Example quotes
Pace	63	59	Sometimes when I am really not understanding a problem, it is hard to have the entire group stop when they are ready to move on.
Group dynamic	47	49	I would like everyone to participate in the conversation. Sometimes it feels like people do not want to be there and they just sit there not doing the work.
Communication	36	20	I wish that everyone would be vocal about when they don't understand something, so our group doesn't have to back track information or steps.
Collaboration	27	34	I sometimes think it would be better for us to work together on the equations than just do them alone.

learning and group dynamics. Many of these TAs were returning TAs who had previously worked in this course when groups were rotated after each exam. Overall, 11 out of the 13 TAs who responded to the survey agreed with the statement that keeping permanent groups allowed students to develop better group dynamics. Justifications focused almost entirely on the fact that keeping the same group throughout the semester both allowed and forced students to find ways to engage productively with one another and establish consistent group norms.

IV. DISCUSSION

We found that, within the student population and instructional context of Mines' Introductory Mechanics course, group composition with respect to incoming ability and gender was not significantly associated with quantitative measures of student performance or perceptions. However, we caution against over-generalizing these results to contexts and populations that are significantly different than Mines. As a highly selective, engineering focused school with a relatively high incoming ability level, it is possible that even our heterogenous ability groups did not have a wide enough range in abilities to fully distinguish them from the matched ability groups. Additionally, student feedback indicated that group interaction was limited for some groups. Research on effective group interactions indicate the importance of providing explanations and receiving responsive feedback. The studio materials used in the course (e.g., problem solving and experimental activities) are also fairly structured and may not elicit sense-making discussions as often as we would like for all groups. Group composition may be

more important in other educational contexts. For example, WKA has used matched groups in her courses for several years. Her group activities focus on conceptual exercises and more open-ended challenge labs. In this context, group observations find students using 7 of the 9 types of explanation that elicit learning as articulated by Webb. Future work should explore the generalizability of these results through replication of this study in other educational contexts and at other institutions.

Student feedback on the group studio environment showed that when students were happy with their groups, the most common justification was that members of the group helped one another and made sure that no members were left behind. Alternatively, common concerns from students who were not completely happy with their group centered around the pace at which the students moved through the material and the desire for additional communication and collaboration between group members. Together, this suggests that for a subset of students, the pace and/or materials characteristic of the studio are not promoting the type of collaborative group environment they are looking for. Alternatively, contrary to our expectations, very few students spontaneously requested that they be able to pick their group or periodically change groups. Together this feedback will inform future refinement of the structure and explicit framing of the group learning environment in our studio courses.

ACKNOWLEDGMENTS

Thank you to the students, faculty, and TAs involved in the course. All authors contributed to the writing and BRW performed the majority of the statistical analysis.

-
<p>[1] P. Heller et al., Am. J. Phys. 60, 627 (1992).
 [2] P. Heller & M. Hollabaugh, Am. J. Phys. 60, 637 (1992).
 [3] N. Ding & E. Harskamp, J. Special Ed. Tech. 15, 331 (2006).
 [4] E. Harskamp, Ed. Res. 50, 307 (2008).
 [5] R. E. Slavin, Rev. of Ed. Res. 57, 293 (1987).</p> | <p>[6] Y. Lou et al., Rev. of Ed. Res. 66, 423 (1996).
 [7] N. M. Webb, J. Res. Math Ed. , 366 (1991).
 [8] R.K. Thornton & D.R. Sokoloff, Am. J. Phys. 66, 338 (1998).
 [9] W. K. Adams et al., Phys. Rev. ST-Phys. Ed. Res. 2, 010101 (2006).
 [10] H.B. Mann & D.R. Whitney, JSTOR , 50 (1947).</p> |
|---|---|