

Illusions Of Gender Effects in Group Exams

Patrick J. Dubois, Sizhe Chen, Jared Stang, Joss Ives

Two-Stage Exams

Two-stage exams involve having students take a conventional exam (e.g. midterm or final) individually, as usual, and then having students form small groups in which to answer the same (or very similar) questions again. The group part realizes many benefits: students get to review answers immediately, social contact helps dissipate exam stress, students learn from each other and practice articulating thoughts. Social and logistical dynamics of the group exam process, however, may magnify or diminish these benefits. These are complex phenomena needing better understanding.

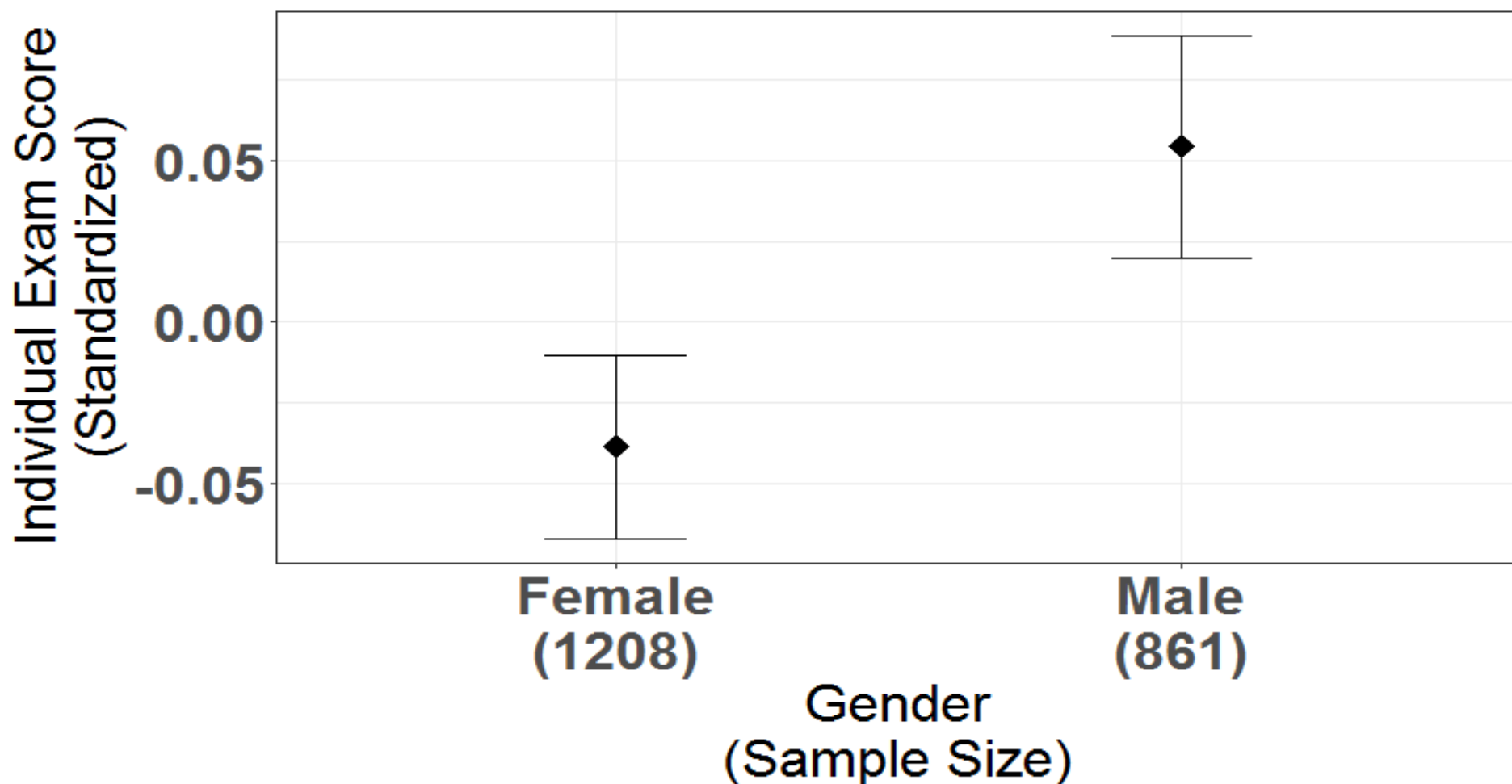
Data from PHYS101

Joss Ives and Jared Stang are leading an inter-disciplinary research team exploring dynamics of the group-exam process. A recurring theme in student exam performance in physics (and STEM in general) is gender. Of 2069 students sampled over three terms, 1208 (58%) identified as female (gender was coded by UBC as a binary variable).

Does Gender Affect Exam Scores? Slightly

In the individual portion of the exam, females did, overall, perform slightly lower ($-.09^*$ $[-.18, -.005]$, $t(1828.84) = -2.08$; $d = -0.09$, measured in standard deviations), as shown in the graph below.

Gender Differences in Individual Exam



Gender Composition of Groups

Students self-selected into groups of 3 or 4. This created the following kinds of gender configurations:

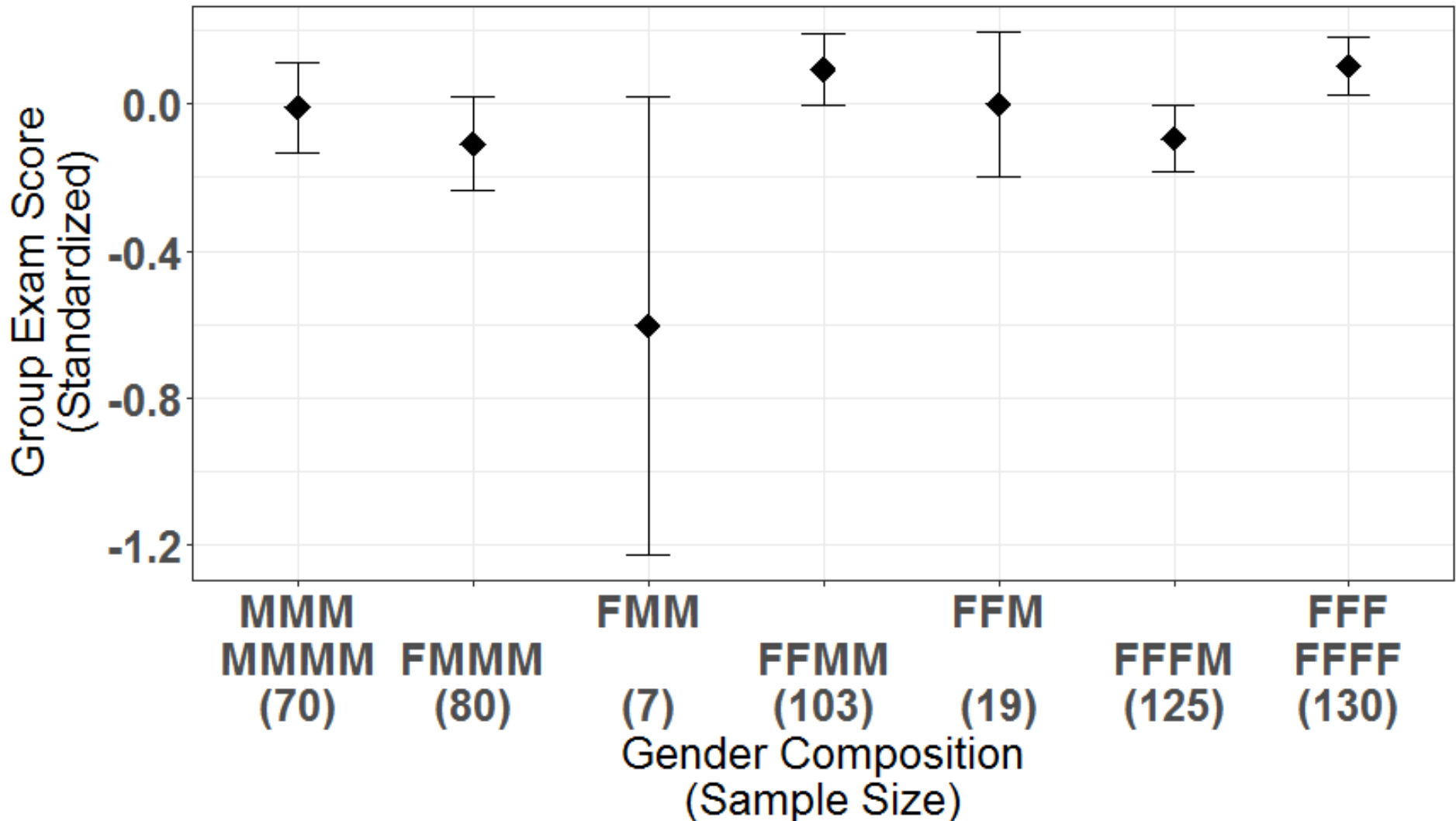
| Configuration | Group of 3 | Group of 4 | N |
|---------------|------------|------------|-----|
| UniformM | MMM | MMMM | 70 |
| SolitaryF | MMF | MMMF | 87 |
| Mixed | – | MMFF | 103 |
| SolitaryM | MFF | MFFF | 144 |
| UniformF | FFF | FFFF | 130 |

To estimate the likelihood of these configurations occurring naturally (i.e. randomly), we ran a simulation in which the same number of students formed the same number of groups of 3 and groups of 4, but chosen purely by chance. Results suggested that students showed a preference for Uniform groups (all the same gender) and avoided Solitary groups having only one member of either gender.

Gender and Group Performance

The number of females in a group had no consistent effect on overall group exam score, $r(532) = .04ns [-.04, .12]$.

Gender Influence on Group Exam Score



Does Gender Composition Affect Group Exam Scores? Indirectly

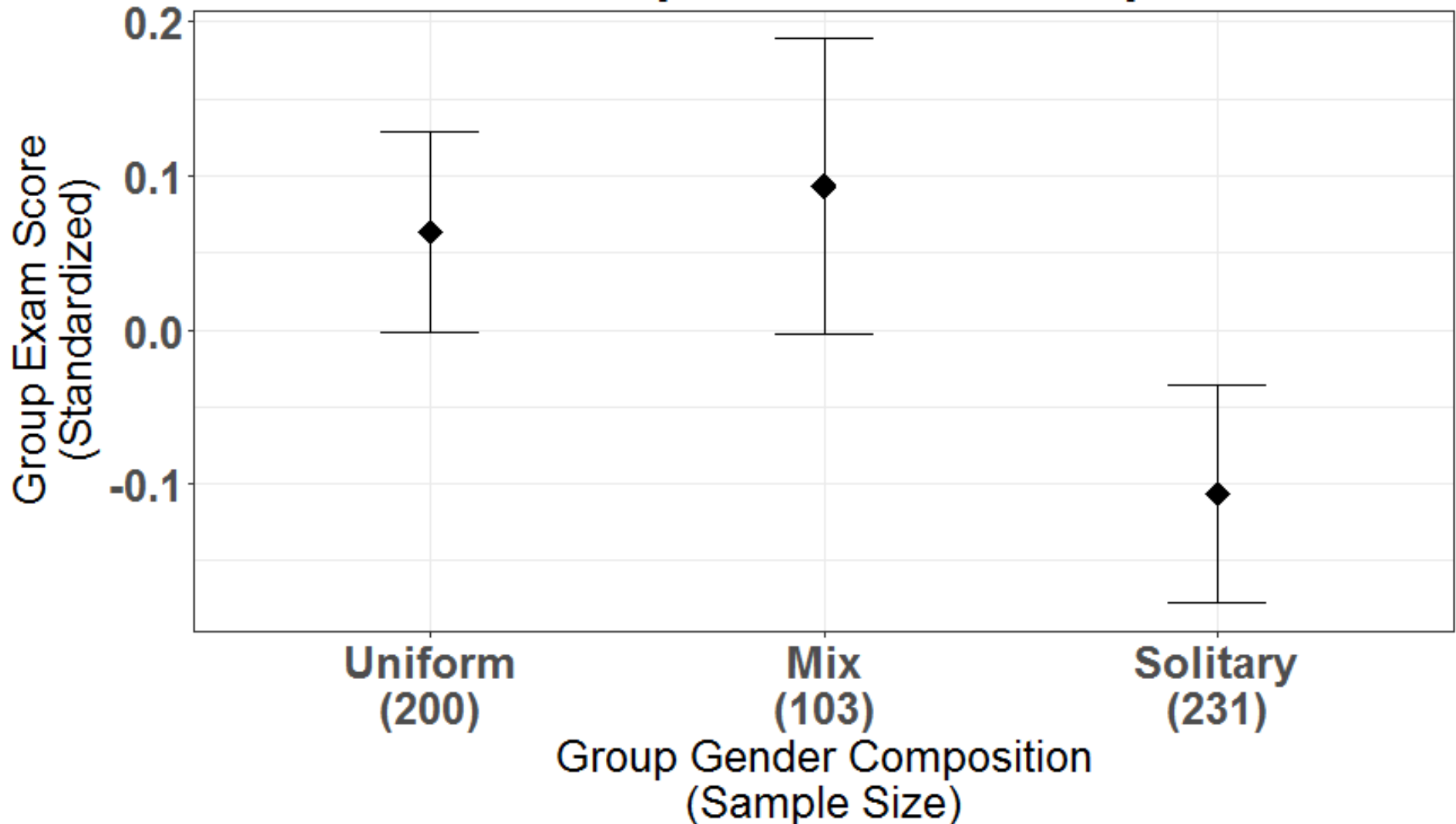
Comparing the mean (standard deviation) of standardized group exam scores from groups of different gender compositions, we find only one comparison significant: groups with only one member of a gender (male or female) tended to show lower performance.

| Composition | N | M (SD) | Composition | N | M (SD) | Difference (95% CI, t-test, effect size) |
|-----------------|-----|-------------|-----------------|-----|-------------|---|
| All Female | 130 | .10 (.87) | All Male | 70 | -.01 (1.01) | -.11ns [-.40, .17], $t(124.37) = -0.79$; $d = -0.12$ |
| Uniform Gender | 200 | .06 (.92) | Not | 334 | -.05 (1.05) | -.11ns [-.28, .06], $t(459.76) = -1.25$; $d = -0.11$ |
| Uniform Gender | 200 | .06 (.92) | Half & Half | 103 | .09 (.97) | .03ns [-.20, .26], $t(196.92) = 0.26$; $d = 0.03$ |
| Uniform Gender | 200 | .06 (.92) | Solitary Gender | 231 | -.11 (1.07) | -.17m [-.36, .02], $t(428.99) = -1.77$; $d = -0.17$ |
| Solitary Male | 144 | -.08 (1.00) | All Female | 130 | .10 (.87) | .18ns [-.04, .41], $t(271.61) = 1.63$; $d = 0.20$ |
| Solitary Female | 87 | -.15 (1.18) | All Male | 70 | -.01 (1.01) | .14ns [-.21, .48], $t(154.33) = 0.79$; $d = 0.12$ |
| Solitary Female | 87 | -.15 (1.18) | Solitary Male | 144 | -.08 (1.00) | .07ns [-.23, .37], $t(159.22) = 0.44$; $d = 0.06$ |
| Solitary Gender | 231 | -.11 (1.07) | Not | 303 | .07 (.94) | .18* [.006, .36], $t(458.24) = 2.03$; $d = 0.18$ |

Disadvantage of Groups with a Solitary Gender?

We found that groups with a Solitary gender configuration (only one male; only one female) performed slightly less well on the group exam.

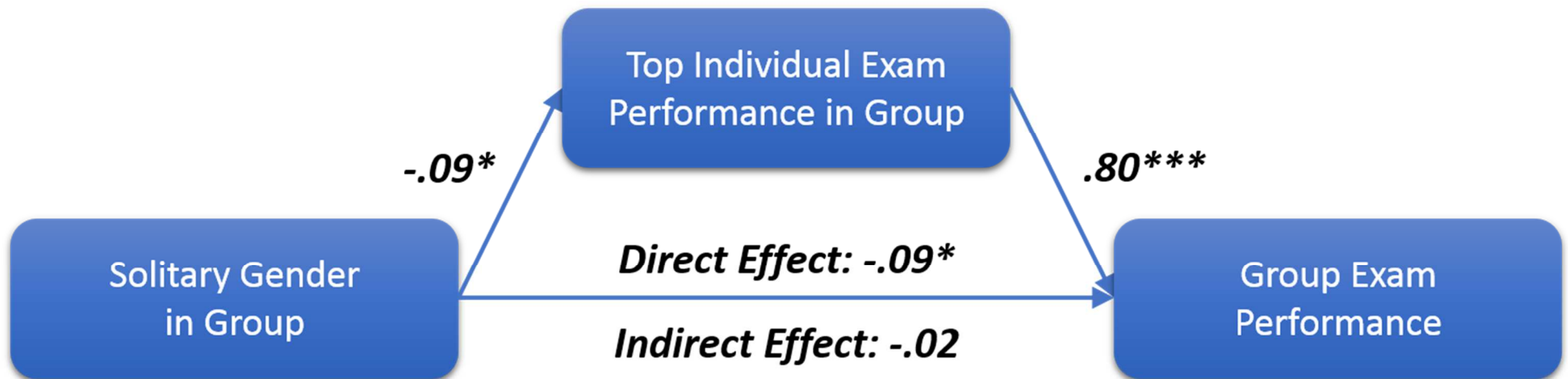
Gender Composition and Group Exam



Effect of Gender Composition is Mediated by Performance Composition

Group exam results can be largely predicted by the individual exam results of those student within the group. The greatest influence comes from the highest performer in the group, i.e. how the group performs can be predicted by how well the strongest student did on the individual exam.

For reasons not currently understood, groups with a Solitary member of one gender also had a lower top performer, with a standardized score difference (between top performers) of $-.10^*$ $[-.21, -.002]$, $t(464.48) = -2.01$; $d = -0.18$.



Mediation effect of gender composition: groups with a solitary gender have lower exam scores because they also have lower top performers.

Summary: Performance Matters More Than Gender

While gender had a slight impact on individual exam behavior, the effect on group exam behavior was more nuanced. Gender itself had no direct effect, but being isolated by gender appeared to. However, that effect was indirect: Solitary-gender groups were also less likely to have higher top performers, which led to slightly lower group exam scores.

Thus, for reasons that are not yet understood, groups having a solitary member of one gender were less preferred by students, and were slightly less likely to have top performers in their group. This incidentally led to such groups showing slightly lower performance in the group phase of the exam.

Acknowledgements:

We gratefully acknowledge the financial support for this project provided by UBC Vancouver students via the Teaching and Learning Enhancement Fund.