

Does group composition impact group scores in two-stage collaborative exams?

Sonya Sabourin¹, Tamara Kelly², and Colin Montpetit¹

¹Department of Biology, University of Ottawa, Ottawa ON; ²Biology Department, York University, Toronto ON

Introduction

Recently, two-stage exams were introduced in two large enrolment second year genetic courses at two different universities. The two courses follow similar formats and use course learning outcomes, activities and materials developed by both instructors. Two-stage exams are those in which students first write an exam individually, followed immediately by a second stage in which they write the same, or similar, exam as part of a small group. Exam grades comprised 85% individual mark and 15% group mark. Overall, student exam grades were improved by the group portion of the exam, however, the extent of score improvement varied between groups, and for several teams, group scores were lower than the members' individual scores. Past studies have highlighted the importance of the highest performing member of a group, but also indicate that other factors may be taken into account as well, such as average and heterogeneity of team members^{1,2,3}. As we try to provide effective collaborative learning experiences through two-stage exams, the goal of this project is to identify and assess the importance of factors that may improve or hinder outcomes for both groups and students.

How can this inform teaching and learning practices?

These results help to quantify the added value of two-stage collaborative testing and justify the importance of keeping undergraduate students engaged in active learning. While group composition may have an impact on group outcomes, we suggest that students be allowed to form their own teams and be made aware that despite the collaborative portion of the exam, they remain responsible for their own success. Although individual academic performance is the main outcome of interest in traditional courses, the value of collaborative learning extends well beyond the classroom. Consequently, we hope these results will encourage undergraduate-level educators to use collaborative testing, as well as other teamwork-based activities, in their own large-enrolment courses. Future research considerations include qualitative and quantitative assessment of student interactions during collaborative tests, as well as comparing individual and group answers to specific questions (e.g. Concept Inventory items) on each test to determine whether performance differences are related to particular concepts.

Group Outcomes

What are the effects of group composition on group performance, and how do these change throughout the semester?

Model

Grade_{ij}

$$= (\gamma_{00} + u_{0j}) + \gamma_{0j}Test + \gamma_{10}Average + \gamma_{20}Best + \gamma_{30}Heterogeneity + \gamma_{40}Fixed + \gamma_{1j}Test * Average + \gamma_{2j}Test * Best + \gamma_{3j}Test * Heterogeneity + \gamma_{4j}Test * Fixed + e_{ij}$$

for *i*th group at *j*th test. All γ_{ij} are coefficients (fixed effects), u_{ij} are the group-level random effects and e_{ij} are residual errors between observed group grades and values predicted by the model.

Table 1. Description of factors included in two-level model.

Level	Factor	Description (units)	Prediction
Dependent variable	Grade	Mark obtained on group test (%)	--
Group-level (i) independent variables	Average (γ_{10})	Average of grades obtained on individual test within each team (%)	Grade increases as Average increases.
	Best (γ_{20})	Best individual score within team (%)	Grade increases as Best increases.
	Heterogeneity (γ_{30})	Standard deviation of grades obtained on individual tests within each team (%)	Grade is affected by heterogeneity of group
Test-level (j) independent variables	Fixed (γ_{40})	Groups who work together for the whole semester (FALSE=0, TRUE=1)	Fixed groups perform better than dynamic groups
	Test (γ_{0j})	Midterm 1 (A); Midterm 2(B); Final Exam (C)	Group grade varies from one test to the next
Interactions between Test and other variables	Test*Ability(γ_{1j})	--	Effect of Ability increases throughout the semester
	Test*Best (γ_{2j})	--	Effect of Best decreases throughout the semester
	Test*Heterogeneity (γ_{3j})	--	Effect of Heterogeneity does not vary throughout the semester
	Test*Fixed(γ_{4j})	--	Effect of Fixed increases throughout the semester

Results

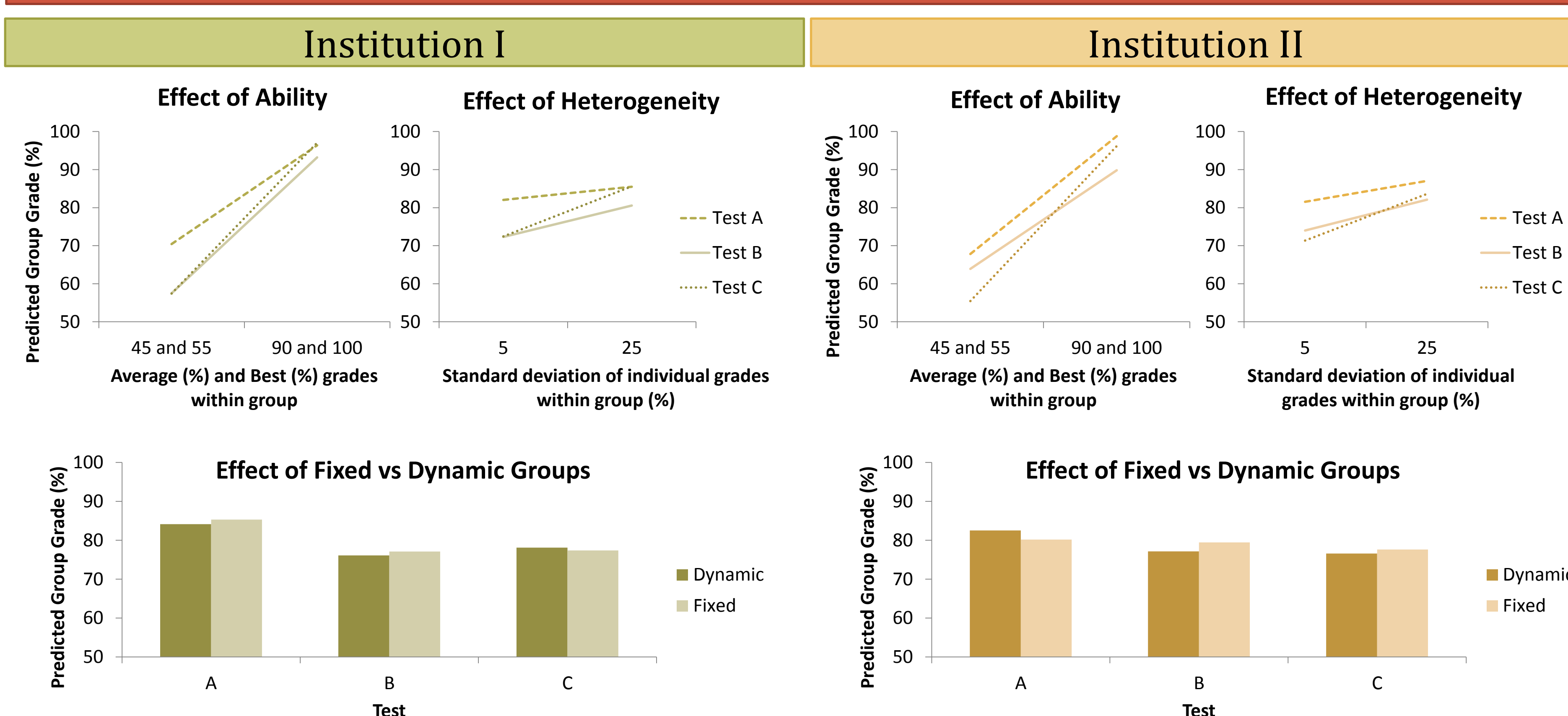


Figure 1. Effect of various group composition factors on group grade. With low and high values of each factor, we can use the model to make predictions concerning group grades and use these to evaluate the scope of each factor's impact on group grade. Effects of Average and Best are combined under Ability due to high correlation between the two factors. Results from both institutions are presented separately, although similar trends are observed in both classes.

Conclusions

- Having at least one strong student in the group can increase the group's performance, especially on tests B and C;
- Whether the groups are fixed or dynamic throughout the semester does not have a significant effect on their performance;
- Heterogeneity of group members also has little effect on performance;
- The group outcomes are similar for both institutions.

Student Outcomes

What are the effects of individual performance and group composition on student gains, and how do these vary throughout the semester?

Model

Gain_{ijk}

$$= (\pi_{000} + r_{00k} + u_{0jk}) + \pi_{00k}Test + (\pi_{100} + \pi_{10k}Test + u_{1jk})Individual + (\pi_{010} + \pi_{01k}Test + r_{01k})Average + (\pi_{020} + \pi_{021}Test + r_{02k})Heterogeneity + (\pi_{200} + \pi_{201}Test)Highest + (\pi_{300} + \pi_{301}Test)Lowest + e_{ijk}$$

for *i*th student in *j*th group at *k*th test. All π_{ijk} are coefficients (fixed effects); u_{ijk} are the student-level random effects and r_{ijk} are the group-level random effects. Finally, e_{ijk} is the residual error between observed and predicted Gain.

Table 2. Description of factors included in three-level model.

Level	Factor	Description (units)	Prediction
Dependent variables	Gain	[Group grade - Individual grade] / [100% - Individual grade]	--
Student-level (i) independent variables	Individual (π_{100})	Grade obtained by each student on individual portion (%)	Gain decreases as Individual grade increases
	High (π_{200})	Highest student within a group, based on individual mark (FALSE=0, TRUE=1)	High students have smaller gains
	Low (π_{300})	Weakest student in a group, based on individual mark (FALSE=0, TRUE=1)	Lower students have increased gains
Group-level (j) independent variables	Average (π_{010})	Average of other group member's individual grade (excluding student's own) (%)	Gain increases as Average of other group members increases
	Heterogeneity (π_{020})	Standard deviation of grades obtained on individual test within each team (%)	Gain is affected by Heterogeneity of group
Test-level (k) independent variables	Test (π_{00k})	Midterm 1 (A), Midterm 2 (B), Final Exam (C)	Effect of test varies from one test to the next
	Test*Individual (π_{10k})	--	Effect of Individual remains consistent throughout semester
Interactions between Test and other variables	Test*Average (π_{01k})	--	Effect of Average remains consistent throughout semester
	Test*Heterogeneity (π_{02k})	--	Effect of Heterogeneity remains consistent throughout semester
	Test*High (π_{20k})	--	Effect of High remains consistent throughout semester
	Test*Low (π_{30k})	--	Effect of Low remains consistent throughout semester

Results

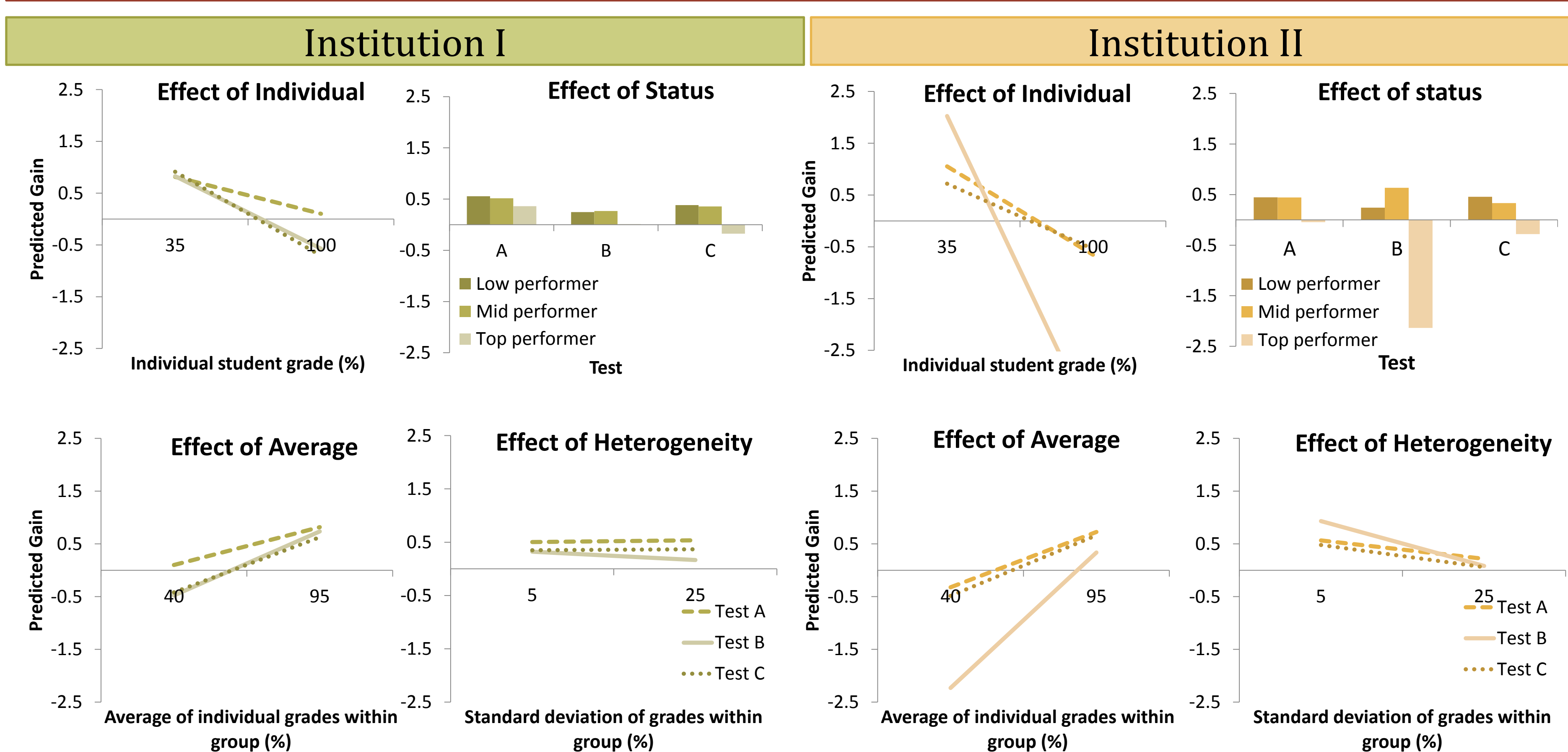


Figure 2. Effect of individual and group factors on student gains. With low and high values of each factor, we can use the model to make predictions concerning student gains and use these to evaluate the scope of each factor's impact on individual student gains. Results from both institutions are presented separately. Results from Test B at Institution II are highly skewed due to a large proportion of students obtaining high individual scores: 25.9% of students obtained over 90% mark. However, of these, only 8.9% obtained positive gains (group score higher than individual score).

Conclusions

- Students may show negative gains (i.e. group grade lower than individual grade), particularly if they are high scorers and/or the other members of their group are low scorers;
- At Institution I, groups tend to outperform their best member on test A, perform as well on test B and worse on test C; at Institution II, groups tend to perform as well as their best member on test A, but not on tests B and C;
- Effect of Average is similar on tests A and C at both institutions;
- Effect of Heterogeneity is negligible at Institution I, but at Institution II, students may obtain higher gains when working with others who obtained similar marks on the individual exam.

References

1. Nihalani et al. (2010). What determines high- and low-performing groups? The Superstar effect. *Journal of Advanced Academics*, 21, 500-529.
2. Watson, Michaelsen & Sharp (1991). Member competence, group interaction, and group decision making: a longitudinal study. *Journal of Applied Psychology*, 76(6), 803-809.
3. Webb, Nemer & Chizhik (1998). Equity issues in collaborative group assessment: group composition and performance. *American Educational Research Journal*, 35(4), 607-651.

Acknowledgements

SSabourin thanks Prof. Morin and Prof. Lamothe for their precious input on the analysis; TKelly thanks her TAs for wandering into the unknown with her and advocating for two-stage tutorials. She also thanks oCUBE members for fantastic discussions and inspiration; CMontpetit thanks the Teaching and Learning Support Services for their Blended Course Design Grant.