Discussion Paper No. 61

Williams Project on the Economics of Higher Education
Denison Gatehouse
Williams College
Williamstown, MA 01267

http://www.williams.edu/wpehe

PEER EFFECTS, GENDER, AND INTELLECTUAL PERFORMANCE AMONG STUDENTS AT A HIGHLY SELECTIVE COLLEGE: A SOCIAL COMPARISON OF ABILITIES ANALYSIS

George R. Goethals
Williams College

October, 2001
DP-61

Abstract

A study was conducted to examine peer effects among undergraduates at Williams College, a highly selective four-year liberal arts school. Specifically, the study explored whether students would perform better writing about newspaper articles they read and discussed in academically homogeneous or heterogeneous groups of three. In homogeneous groups all three students were from either the top half or bottom half of their class on academic ratings assigned at the time of admission. Heterogeneous groups included students from both the top and bottom half of their class. The results showed that students in the top and bottom half performed similarly overall, but that students performed better in homogeneous groups, whether those homogeneous groups were made up of students in the top half or the bottom half of their classes. This pattern of results was stronger for men subjects than women subjects. The results were interpreted in terms of the principles of social comparison theory (Festinger, 1954).

Peer effects, gender and intellectual performance among students at a highly selective college: a social comparison of abilities analysis[1]

Recent research by Winston and his colleagues (Goethals, Winston, & Zimmerman, 1999;Winston, 1997) suggests that colleges and universities, if they can, manage their enrollments so as to maximize student quality. Their ability to manage in this way depends on their wealth and their ability to use that wealth to offer a highly desirable product, education at a top-flight school. Schools in this fortunate position do not sell their educational product to all customers who are willing to pay. They limit supply and select only the most promising students. One reason they do this is their belief in peer effects. The term peer effects refers simply to the impact the other students have on an individual's educational experience. It is thought that students are likely to get a better education if their fellow students, their peers, have higher degrees of academic talent. In that sense, one of the things that students buy when they attend college is the other students who form the peer environment. Students are both customers and a key component of the product they are buying. That is, one key aspect of the technology of producing higher education is a "customer-input technology."

What evidence is there for peer effects, for the idea that peer quality makes a difference in student education? What qualifications complicate a simple story about the value to individual students of having more rather than less talented fellow students? Clearly there is a good deal of evidence that young people are influenced by peers and

* The author would to thank Henry Bruton, Gordon Winston and David Zimmerman for their wisdom and advice in planning and conducting the research reported here, and O'Neil Campbell, Jared Carbone, Sally-Thomas Daigneault, Chin Ho, Laurie Hurshman, Melissa Murphy, Rachel Seys, Lida Ungar, and Peter Walke for their help in collecting and analyzing the data. Funding for this paper was provided by the Andrew Mellon Foundation through the Williams Project on the Economics of Higher Education.

some evidence that peer effects in education operate at elementary and secondary school levels (e.g., Coleman et al., 1966). Recent studies also support the idea of peer effects among college students (Goethals, 2000; Hoxby, 2000; Sacerdote, 2000; Stinebrickner & Stinebrickner, 2001; Zimmerman, 1999). These studies report specific findings of peer effects at Berea College, Dartmouth College, and Williams College, as well as effects aggregated across many other schools. (It is worth noting that Berea is a very different school from Dartmouth and Williams. Its emphasis on admitting students with substantial financial need distinguishes it.) Two of the papers above (by Goethals and by Stinebrickner & Stinebrickner) also indicate that women are more likely to be positively influenced by superior peers than men. In fact, the Goethals' experiment with Williams College students suggests that during the course of a brief discussion of newspaper articles, men are likely to disengage from academically superior peers. The purpose of this study is to explore peer effects further, specifically to investigate how both men and women at an elite college (Williams) are differentially affected by academically superior, similar, or inferior peers.

Social comparison and peer effects. Festinger's theory of social comparison has provided a useful social psychological framework for understanding peer effects (Festinger, 1954; Suls & Wheeler, 2000). It begins with the proposition that people have a drive to evaluate their opinions and abilities. Decades of research have shown that people compare themselves with others on many other personal characteristics, such as income, attractiveness, and health, but the theory's original emphasis on opinions and abilities is extremely relevant to a consideration of peer influences among college students (Suls & Miller, 1977; Suls & Wills, 1991; Wood, 1996). Festinger argued that

people evaluate their opinions and abilities through comparison with other people and that they can make much more stable evaluations by comparing with other people who are similar. An important consequence of the need for similar others to satisfy evaluation needs is strong pressure within groups toward uniformity of opinions and abilities. When opinions are at issue the pressures toward uniformity are unalloyed, and there is discussion until talk has produced uniformity, or until those with deviant opinions are rejected from the group, usually with some degree of hostility. When abilities are being evaluated, pressures toward uniformity combine with pressures toward being better than others. Individuals compete until a ranking evolves, marked by differences within a narrow range. Those with highly different ability levels become defined as non-comparable – comparison with them ceases – although they are not rejected in a hostile way, as is the case for opinions. They simply cease being a part of the individual's reference group, and they are largely ignored. In short, pressures toward uniformity in groups produce marked opinion and ability homogeneity, if not uniformity, through changes in opinion or ability levels or through adjustment of group boundaries.

There are several important implications for peer effects. Often, discussion and competition in groups can lead to productive exchange and learning. But such interaction may affect more able and less able students differently. Let us first consider the case of the less able student interacting with superior peers vs. similar peers. Superior peers could foster or enable superior performance for less able students. On the other hand, less able students may cease comparing with those that they perceive to be significantly more able. They may regard superior peers as noncomparable, or even intimidating, and

therefore discontinue engaging with them in a productive exchange of ideas. They may do better with similar, less able peers.

The studies by Goethals (2000) and Stinebrickner and Stinebrickner (2001) suggest that there may be a difference between men and women in their reactions to superior peers. In an experimental study, Goethals (2000) found that women but not men benefited from superior peers. They seemed to be open to the ideas of their superior peers, and often behaved in ways that facilitated others' participation in discussion. Stinebrickner and Stinebrickner (2001) found "compelling evidence of peer effects in first semester grades" for women, but not men, at Berea College (p. 8). They speculated that women may be more accepting of roommates with different backgrounds. These results suggest that in a similar study women but not men will show positive peer effects from interacting with superior peers vs. similar peers. On the other hand, Sacerdote's (1999) study with Dartmouth students found that roommates in the top 25% on academic indices lift one's own grades, and no gender differences were reported. Zimmerman (1999) also found peer effects, with a 100 point increase in roommate's verbal SAT being associated with a small but statistically significant increase in one's own grades. Zimmerman found no gender differences. Taken together, these data suggest that we can expect positive peer effects to be produced by more academically able peers, at least among women.

What about the impact of peer interaction on more able students? Are there reasons to be concerned about negative peer effects, such that high ability students are negatively influenced by peers of lower ability levels? If high ability peers do help lower ability students, do they do so at some cost to themselves? While this is possible, it is

also possible that high ability students benefit from helping low ability students, due to a "teaching effect" identified by Zajonc (1976). Zajonc suggests that children often benefit from teaching younger siblings, siblings who are intellectually less advanced. What evidence do we have relevant to these possibilities? Sacerdote finds that students in the bottom 25% on academic indices have no impact on their roommates' grades. Top students can help but bottom students seem not to hurt. In contrast, Stinebrickner and Stinebrickner find that high income female students "paired with a low income student may be less likely to remain in school than the high income student who is paired with a high income student…(p. 11)" This finding does not speak to how students of different ability levels affect each other, it does alert us to the possibility of negative peer effects. Also, Zimmerman finds that students in the middle of the SAT distribution can be negatively affected by roommates in the bottom 15% of the SAT distribution.

These considerations suggest that it would be important to study both homogeneous and heterogeneous groups to see how higher ability students fare with similar and inferior peers, and how lower ability students fare with similar or superior peers. In general, based on the findings by Goethals and Stinebrickner and Stinebrickner, we expect that women will show more benefit from interacting with superior peers than men. On the other hand, in considering whether either men or women will be affected positively or negatively by inferior peers, we have little basis for prediction. Zajonc work suggests a potential benefit. We believe that this possibility should be assessed.

The present study explores these questions using a methodology very similar to that used by Goethals (2000). It works as follows: groups of three undergraduates at Williams College spend forty minutes in the context of a psychology experiment reading

and then discussing articles from The New York Times, and then twenty minutes writing about what they've learned from reading and discussing the articles. The academic ability levels of the students and thus the group composition is controlled. The quantity and quality of what students write is assessed by trained coders. While opportunities for interaction, and therefore peer effects, are limited in this setting, its experimental nature does offer the advantage of controlling group composition. Selection effects are eliminated. The question is whether the setting provides enough basis at a school where nearly all students have very high academic ability for students to react to whatever similarity or difference in academic ability does exist in their groups. Since our earlier study using this methodology did find peer effects, it was used again.

We compare the performance of individual students, male and female, in four different kinds of groups in which students are classified as being in the top half or bottom half of their class academically using data from the college's admission office. The names of the groups refer to the number of students in the group who are in the top or bottom half of the class. The four kinds of groups are called All Top Half; Two Top Half; One Top Half; and All Bottom Half. The average SAT's for students in the top half of the class are approximately 755 Verbal and 735 Math. The approximate averages for students in the bottom half of the class are 655 Verbal and 655 Math. That is, the top half students have an average combined SAT of about 1490 while the bottom half have a combined average of approximately 1310. Both groups of students have higher SATs than the vast majority of test takers. The fact that there is relatively little difference between the two groups relative to the entire SAT distribution makes it less likely that

peer effects will emerge. Any that do, therefore, must be seen as powerful indications of those that might take place with greater differences between groups.

As explained above we expect that lower half female students will perform better after interacting with superior peers that similar peers, but that lower half male students will not similarly benefit. We have no clear prediction as to how upper half students will perform with inferior vs. similar peers. There are reasons for thinking that any of the three logically possible outcomes might occur: some negative impact of inferior peers, no impact, or some benefit from interacting with inferior peers.

Method

Participants

One hundred thirty-eight Williams College first-year students and sophomores volunteered to participate in this study. They were paid $15.00 or received one-hour of extra-credit in an Introductory Psychology course. The study was called "College Students and Public Affairs."

Procedure

Participants were scheduled in groups of three, such that all three participants were in the same class (freshman or sophomore). Participants were greeted by an experimenter who explained briefly that the study entailed reading three articles from the New York Times, discussing those articles as a group, and answering questions about what they had read and discussed. The participants sat at a round table with a microphone in the center. The experimenter explained that they would be observed through a one-way mirror and that their discussion would be video-taped using the microphone and a ceiling-mounted camera.

Participants were given twenty-minutes to read three articles, twenty-minutes for discussion, and twenty-minutes to answer a questionnaire asking about what they had learned from reading and discussing the articles. After giving instructions the experimenter left the room, and subsequently returned twice, first to ask the participants to begin the discussion and then to ask them to stop the discussion and complete the questionnaires. The room lighting was arranged so that when the experimenter was out of the room she was still partially visible in the adjoining room through the one-way mirror.

Academic Ratings

As mentioned above, students were recruited for the study and assigned to groups on the basis of academic ratings assigned by the Office of Admission when students apply. The academic rating is based on students' secondary school grades, the quality of their secondary school academic program, their SAT's, and information in recommendations that seems to reveal academic potential. The academic ratings have been used for many years and are, at Williams, the best available predictors of student grades. While the academic rating predicts student grades better than any of its components, the best single predictor among the components is Verbal SAT.

Materials

Participants read three articles published in the New York Times in the summer of 1999. The first discussed the increasing amount of time that Americans spend at work vs. leisure (So Much Work, So Little Time, by Steven Greenhouse). The second discussed issues in genetic engineering raised by the then recent finding that Princeton scientists had created a genetically smarter strain of mice (Ideas & Trends: Eek!; The Hidden Traps in Fooling Mother Nature, by Nicholas Wade). The third dealt with AIDS

prevention (Focusing on Prevention in Fight Against AIDS, by Lawrence K. Altman). The questionnaire asked students to rate on seven-point scales how much they learned from reading and discussing each article, how interested they would be in reading or discussing such articles in the future, and how much they learned from each of the other two students. It also asked them to write on one page the ideas or information they learned from reading or discussing each article. The page listed the numbers one through ten, to provide space to write ten statements, but said that the reverse side could be used as well.

Coding of written responses

Trained undergraduate raters individually coded each participant's statements of ideas and information. A quantity rating gave credit for each idea or piece of information the participant stated. A quality score rating from one to three was given to each statement on the basis of its specificity, detail and elaboration. A total quality points score and an average quality score per statement were calculated for each article, for the first two articles (the ones that were discussed), and for all three articles. Inter-rater agreement on quantity scores was virtually 100%. For quality scores it was 74% and all disagreements were averaged.

Coding of discussion videotapes

Undergraduate raters coded each statement in the videotape of each discussion. First each rater proposed a written "order of talk" that listed who spoke when on the tape. There was near 100% agreement on who was speaking and any disagreements were resolved through discussion. Then each statement was given a length rating from one to four, depending on whether the statement was less than 5 seconds, from 6 to 10 seconds,

11 to 15 seconds, or greater than 15 seconds. Inter-rater agreement for quantity ratings was 95%. Disagreements were resolved through averaging.

Each statement was also given a quality rating of negative one to three, based on how effectively the statement advanced the discussion, and contributed to the intellectual quality of the discussion. Negative one scores were given to statements that halted or derailed discussion. Zero was given to statements that were neutral or bland, one was given to remarks that advanced the discussion through simple statements or questions, two was given to remarks that were more thought provoking, and three was given to those rare statements that advanced the discussion productively and that were exemplary in thought and expression. The two raters agreed on 73% of the quality ratings. Disagreements were resolved by averaging.

For each participant and each group, total and average quantity and quality scores were calculated. Also, each subject's "peer environment" was calculated by averaging the total quality scores of his or her two fellow participants.

Results

There was a total of 46 groups. Eight were All Top Half, 18 were Two Top Half, 13 were Two Bottom Half, and 7 were All Bottom Half. The unit of analysis was individual subjects, such that there were 24 top half subjects in All Top Half groups, 36 top half subjects and 18 bottom half subjects in Two Top Half groups, 13 top half subjects and 26 bottom half subjects in One Top Half groups, and 21 bottom half subjects in the All Bottom Half groups. All results reported below are statistically significant at the .05 level unless otherwise noted.

Group Differences

Three factor analyses of variance were conducted on the major measures of the study, the three factors being the type of group (Group Type: All Top Half, Two Top Half, One Top Half, All Bottom Half.), the subject's gender (Subject Gender: Male vs. Female), and the subject's academic rating (Subject Rating: Top Half or Bottom Half). On the critical verbal quality measure the only significant effect is a strong main effect for Type ($p < .01$; see Table 1). Subjects in the All Top Half ($m = 18.7$) and All Bottom Half ($m = 21.4$) groups do better than subjects in mixed groups (Two Top Half, $m = 16.2$ or One Top Half, $m = 15.7$). The pattern of means is the same for both Top Half and Bottom Half subjects, such that the Group Type X Subject Rating interaction does not approach significance. Both top and bottom half subjects do better in homogeneous groups. In fact, as discussed further below, one of the most remarkable characteristics of the data is that bottom half and top half subjects differ so little. The pattern of means showing that subjects do better in academically homogeneous groups than academically mixed groups is similar for men and women, although it is stronger for men, such that the Group Type X Subject Gender interaction does approach significance ($p =. 207$). Among men the All Top Half ($m = 20.8$) and All Bottom Half ($m = 22.7$) subjects do significantly better than men in mixed groups (Two Top Half, $m = 15.1$ or One Top Half, $m = 13.9$). For women, those in All Bottom Half groups perform the best ($m = 20.5$). Women in the other three groups perform similarly to each other, but significantly less well than those in the All Bottom Half groups (All Top Half, $m = 16.2$; Two Top Half, $m = 17.2$; One Top Half, $m = 17.0$).

Perhaps the most meaningful way to look at the data is examine how top and bottom half subjects perform when they are in homogeneous groups (that is, both peers in their groups are in the same half of the class as the subject is) or heterogeneous groups (where one or both of the subject's peers is in the other half of the class). For men the results are very striking. Among top half males, those in All Top Half groups have an average written quality score of 20.8 while those in mixed groups, where one or both of their peers are in the bottom half, have means of 16.5 and 14.0, respectively. Among bottom half males, those in the All Bottom Half groups have a written quality mean of 22.7 while those in mixed groups, where one or both of their peers are in the top half, have means of 13.9 and 12.7, respectively. Among women subjects there is not such a performance advantage for those in homogenous groups, though the data generally point in that direction. Among top half female subjects, those in the All Top Half groups have a written quality mean of 16.2 whereas top half women in groups with one or two bottom half peers have means of 17.4 and 15.1, respectively. Among bottom half females, those in All Bottom Half groups have a written quality mean of 20.5 whereas bottom half women in groups with one or two top half peers have means of 17.9 and 16.7, respectively.

As mentioned earlier, another way to look at the data reported above is to note how remarkably similar is the performance of top half and bottom half subjects, in both homogeneous and heterogeneous groups. First, consider top half and bottom half subjects in homogeneous groups. Those in All Top Half and All Bottom Half groups perform much the same, 18.7 vs. 21.4. If anything, the Bottom Half subjects perform better, but the difference is not significant. In heterogeneous groups the story is similar.

Top half subjects and bottom half subjects in the Two Top Half groups perform similarly, 17.0 and 14.7, respectively, and top half and bottom half subjects in the One Top Half groups average 14.6 and 16.2, respectively. The only notable departure from this pattern is that top half males in Two Top Half groups perform better than bottom half males in those groups, 16.5 to 12.7.

Our focus has been on how men and women perform in groups where, in terms of academic ratings, their two peers are both similar, both different, or one is similar and one is different. We also considered whether their peers are men or women. There is no suggestion in the data that the gender composition of the peers matters. While the numbers are too small to give us much confidence that gender composition really does not matter, we will not consider the issue further here. There is simply not any signal in this set of data warranting consideration.

Discussion

The results give strong support to the idea that peers matter. If peer academic ability exerted no influence, we would not find the large effects of group composition that we did find. In the laboratory setting we have created, both high and low academically rated students are affected by their peer environment. When both of their peers have academic ratings similar to their own they perform significantly better than if one or both of their peers are different. Top half students do better when their peers are both in the top half, and bottom half students do better when their peers are both in the bottom half. In the context of finding no overall difference between the performance of top half and bottom half students, the impact of the peer environment is remarkable.

How do these results fit with other studies of peer effects among college students, particularly those of Goethals (2000), Sacerdote (2000), Stinebrickner and Stinebrickner (2001), and Zimmerman (1999)? Can we make sense of them theoretically? What policy implications do they have? While the methodological approach of these laboratory studies differs dramatically from the econometric approach of the other studies, two findings in the lab studies echo those of the econometric studies. First, both Stinebrickner and Stinebrickner and Zimmerman find hints that peer effects can be negative. The Stinebrickners find that staying in school can be negatively affected by low income peers. Zimmerman finds that low SAT peers can negatively affect the academic performance of students in the middle SAT range. Second, the Stinebrickners find positive peer effects for women only. The lab data show that men do not respond well to academically superior peers, while women sometimes do.

On the other hand, these results are strikingly different from the other studies in not showing that top students can positively affect less able students. All three of the econometric studies show positive effects of high ability peers on students with less ability. In the first but not the second of our two lab studies we show positive effects for women, and there are no such effects for men. Instead, our results suggest a positive result from interacting with similar peers, for both top half and bottom half students. Further research will be needed to provide a clearer sense of the conditions under which students of a given ability level can benefit from interacting with those of a higher ability level.

Theoretically, the results make good sense in terms of social comparison theory. People in groups who are similar in ways that are relevant to the tasks of the group

should interact most effectively. When the group is homogeneous both opinion and ability comparison should be enhanced and group members should engage most effectively in the task of discussing and then writing about the assigned articles. The current results strongly support this hypothesis. We found in this study and in Goethals (2000) that Williams students with higher and lower academic ratings perform, overall, about the same. That is, the rated academic ability of individuals does not in itself affect student performance in this setting. On the other hand, group composition, with respect to academic rating, is crucially important. If the group is homogeneous, engagement and performance is enhanced.

There are several findings which clearly support the idea that individual subjects influence the engagement of their peers. One comes from an analysis of the videotaped discussions. Each subject received a total verbal quality score, based on the sum of the quality ratings of his or her verbal utterances. There is a highly significant correlation between individual subjects' total verbal quality scores and the total verbal quality scores of his or her peers ($r = .39$, $p < .001$). Similarly, there is a highly significant correlation between a subject's written quality score and the average written quality score of his or her two group peers ($r = .56$, $p < .001$). If there were no peer influence these correlations would be zero. But clearly, norms emerge within each group as to how seriously to engage in both the discussion and writing tasks. Also, if good discussions emerge in groups, there is a link to how well individual subjects perform in the writing task. The correlation between a group's total verbal quality score (summed across the three group members) and an individual subject's written quality score is significant ($r = .17$, $p < .05$).

Two other questions arise. In the context of this study, perceived academically-relevant similarities and differences between students have a dramatic effect. Why do they have so much impact, and is it possible to induce students to think of their similarity to or difference from others slightly differently, in a way that might increase academic engagement with dissimilar others. Similarity or difference matters so greatly quite probably because of the enormous importance of academic ability to success in a school like Williams. Especially in the first two years many students are getting their academic feet on the ground and forming self-concepts that include a strong sense of being more or less capable in the intellectual domain. Their comfort-level and engagement in this highly consequential area can affect how they approach it. Do they turn on or turn off? In homogeneous groups it is easy to engage. But in heterogeneous groups there is apt to be discomfort both for those who feel intimidated or outclassed by superiors and those who sense the discomfort and perhaps disinterest of those who are less talented and confident. At Williams there are differences between the more and less academically talented that are very salient. Students in the bottom half of the class are much more likely to be athletes or minorities, groups that sometimes disidentify with the academic enterprise because of relevant negative stereotypes (Shulman & Bowen, 2001; Steele, 1997). To the extent that members of these groups are easily identifiable, and perceive themselves to be, and also perceive themselves to be different and perceived to be different, the chances of disengagement in an initial encounter seem high.

Are there ways that these real and perceived differences can be made less disruptive to productive academic engagement? First, it might be that after relatively little time, students just "get over it" and interact easily over perceived group boundaries.

Data from the roommates studies cited earlier is reassuring on this score. Second, since unstable intellectual self-evaluations might be playing a central role in these experiments, perhaps some intellectual affirmation could alter the dynamic. If students were more generally affirmed in their intellectual identity -- by faculty, administrators, coaches, and alumni -- there might be more fruitful intellectual interaction and engagement among all students. Third, linked to intellectual affirmation, if the similarities rather than the differences between groups of students could be emphasized, the dynamics of interaction between students with different academic promise might be enhanced.

References

Coleman, J.S. et al. (1966) Equality of Educational Opportunity, U.S. Department of Health, Education, and Welfare, Office of Education.

Festinger, L. (1954) A theory of social comparison processes. Human Relations, 7, 117-140.

Goethals, G.R., Winston, G.C., & Zimmerman, D.J. (1999) Students educating students: the emerging role of peer effects. In Forum futures: Exploring the future of higher education, M. Devlin & J. Meyerson, (eds.), pp. 25-45. Forum Publishing

Hoxby, C. (2000) Peer effects in the classroom: learning from gender and race variation: NBER Working Paper No. 7867.

Sacerdote, B. (2000) Peer effects with random room assignment: results for Dartmouth roommates. NBER working paper # 7469.

Shulman, J.L., & Bowen, W.G. (2001) The game of life: college sports and educational values. Princeton, NJ: Princeton University Press

Steele, C.M. (1997) A threat in the air: how stereotypes shape intellectual identity and performance. American Psychologist, 52, 613-629.

Stinebrickner, R., & Stinebrickner, T. (2001) Peer effects among students from disadvantaged backgrounds. Mimeo.

Suls, J.M., & Miller, R.L. (1977) Social comparison processes: theoretical and empirical perspectives. Washington, DC: Hemisphere.

Suls, J.M., & Wheeler, L. (2000) Handbook of social comparison: theory and research . New York: Kluwer Academic/Plenum Publishers.

Suls, J.M., & Wills, T.A. (1991) Social comparison: contemporary theory and research. Hillsdale, NJ: Erlbaum.

Winston, G.C. (1997) Why can't a college be more like a firm? Williams Project on the Ecnomics of Higher Education, Discussion Paper #42.

Wood, J.V. (1996) What is social comparison and how should we study it? Personality and social psychology bulletin, 22, 520-537.

Zajonc, R.B. (1976). Family configuration and intelligence. Science, 192, 227-236.

Table 1

Written Quality Scores for Experimental Groups

Group Type	*Student Type*	**Sex**	**Mean**	**Number**
All Top Half Academic Rating	Top Half Academic Rating	Female	16.2	11
		Male	20.8	13
		Both	18.7	24
	All Students in Group	Total	18.7	24
Two Top Half Academic Rating	Top Half Academic Rating	Female	17.4	20
		Male	16.5	16
		Both	17.0	36
	Bottom Half Academic Rating	Female	16.7	9
		Male	12.7	9
		Both	14.7	18
	All Students in Group	Total	16.2	54
One Top Half Academic Rating	Top Half Academic Rating	Female	15.1	7
		Male	14.0	6
		Both	14.6	13
	Bottom Half Academic Rating	Female	17.9	15
		Male	13.9	11
		Both	16.2	26
	All Students in Group	Total	15.7	39
All Bottom Half Academic Rating	Bottom Half Academic Rating	Female	20.5	12
		Male	22.7	9
		Both	21.4	21
	All Students in Group	Total	21.4	21