

Learning from Your Classmates: A Multi-Method Assessment of Classmate Peer Effects in First-Year Core Courses at Three Colleges

Jeffrey Parker, Reed College

Jon Rivenburg, Reed College

Jay Beaman, Lewis & Clark College

Neal Christopherson, Whitman College

With essential contributions from:

James Grant and Benjamin Westervelt, Lewis & Clark College

Jan Crouter and Ruth Russo, Whitman College

Nigel Nicholson, Reed College

This paper grew out of an earlier study of classmate peer effects at Reed College, which was undertaken by Jessica Hoel with Parker and Rivenburg under the sponsorship of the Bernard Goldhammer Collaborative Grant Program. The current research is part of a project supported by an Outcomes and Assessments Grant from the Teagle Foundation. Nigel Nicholson, Ruth Russo, Ben Westervelt, Jay Beaman, and Neal Christopherson are members of the Teagle project research team and contributed, directly and indirectly, to this research. The project also benefited from outstanding research assistance over several years from numerous students at all three colleges, including Nick Chandler-Klein, Kathryn Deane, Christa Gibbs, Ranganai Gwati, Joe Hand, Ben Keefer, David Kling, Joshua Lingerfelt, Morgan Matthews, Ramon Mitchell, Daniela Morar, Dylan Smith, Michael Turner, Izarra Varela, and Liza Wickline.

Introduction

Our study explores classmate peer effects—the effects of classmates on a student’s learning—in required first-year core courses at three selective Northwest liberal-arts colleges. We examine classroom peer effects with two complementary approaches, one based on quantitative evidence and the other based on instructor interviews. The statistical analysis looks for the effects of measurable classmate characteristics (such as admission credentials) on measurable academic outcomes (grades and persistence). The interviews provide a rich contextual background that illuminates the results of the statistical analysis and suggests important channels of peer effects that we were not able to explore quantitatively with the available data.

To preview our main findings, the instructors we interviewed were strongly convinced that peer effects are prominent in discussion-based core courses. However, the most striking consensus was that peer effects related to students’ attitudes and classroom behavior were more important than those associated with their abilities. While having high academic aptitude of the kind that the SAT attempts to measure was viewed as a positive peer characteristic by the instructors, most were clear in emphasizing behavioral rather than ability-related peer attributes. Our quantitative analysis was restricted to student characteristics available in the college databases, which was restricted to such admission measures as SAT scores, high-school records, and admission-office ratings. We found no effect of these aptitude-based classmate characteristics on student learning at any of the three schools. This suggests that future peer-effect research might be productively directed at collecting data on the attitudes and personalities of incoming students and examining the effects that these characteristics have on classmate learning.

Peer Effects

Peer effects are influences of an individual’s peers on his or her behavior. They may occur in many contexts; for example, co-workers’ productivity may influence one’s own, neighbors’ consumption decisions may cause one to consume more or less, one may be tempted to drink or smoke more if one’s friends do. We are concerned with how peer effects among college students affect their academic achievement.

Leaders of colleges seem to believe that peer effects are important. Colleges that are in a position to do so admit students selectively, picking those thought to have the most academic potential. Many colleges attempt to improve the academic quality of their student bodies by offering merit-based financial-aid awards to students who are especially promising. Among several possible

explanations for the pervasive aspiration to selectivity and the frequent practice of merit aid, peer effects relate most closely to the pedagogical mission of a college.¹ To the extent that having “better peers” gives students a better education, the goal of attracting an excellent student body aids attainment of the goal of educating students as effectively as possible. In the presence of positive ability-based peer effects, Rothschild and White (1995) demonstrate that offering merit aid to excellent students can enhance the efficiency of the education system by compensating students for the “services” that they provide in enhancing their peers’ learning.

The student interactions leading to peer effects occur on many levels: roommates, dorm-mates, friends, lab partners, classmates, colleagues in activities, and myriad other interactions. Some student peer effects are broad and occur at the level of an institutional as a whole. For example, an excellent overall student body allows the faculty to offer challenging courses that elevate the level of students’ learning. Institution-wide peer effects may even extend beyond a student’s contemporary peers, as for example when the value of his or her degree is enhanced by the past success of the college’s alumni.

In contrast, some student peer effects occur at the individual or class level. If a student’s roommate has poor study habits, it may distract the student from studying and lower his or her academic performance.² Classmates who contribute clever insights to class discussions or who are effective study partners may improve the learning of others in the class.

In order to distinguish statistically between the effects of peers’ characteristics and the effects of a student’s own characteristics, it is necessary to find situations in which peers are assigned either randomly or according to observable criteria that can be used as controls.³ Most studies of classroom peer effects have focused on primary education, presumably because it is more feasible to undertake controlled studies with random peer assignment in elementary schools where the both curriculum and students’ academic backgrounds are relatively homogeneous. Among the studies examining classroom peer effects among younger students are Henderson, Mieszkowski, and Sauvegeau (1976), Epple and Romano (1998), Hoxby (2000), Zimmer and Toma (2000), Boozer and Cacciola (2001), Gaviria and Raphael (2001), Vandenberghe (2002), Hanushek et al. (2003), McEwan (2003), Rangvid (2003), Robertson and Symons (2003), Zimmer (2003), Angrist and Lang (2004), Dills

¹ Among the other rationales for these phenomena, high-ability students might be easy or more fun to teach and their subsequent success might bring prestige (and financial support) to the college.

² For example, Stinebrickner and Stinebrickner (2008) find that students whose roommates brought a video-game system to campus studied less and earned lower grades than those whose roommates did not, controlling for the student’s own characteristics.

³ See Manski (1993).

(2005), Lefgren (2004), Ammermueller and Pischke (2006), Burke and Sass (2006), Ding and Lehrer (2006), and Kang (2007). Although the evidence is not unanimous, many studies find support for the hypothesis that young students learn more when they have peers of higher ability.

Higher education peer effects are difficult to measure because it is usually infeasible to do a controlled experiment with randomized classmate selection. Even “natural experiments” in which students’ peers are selected randomly or according to observable criteria or criteria unrelated to academic outcomes are uncommon. Most higher-education studies have examined roommates because the process of roommate selection for first-year students often conforms to the natural experiment conditions above. Studies based on residential peers include Sacerdote (2001), Zimmerman (2003), Winston and Zimmerman (2004), Stinebrickner and Stinebrickner (2006), Foster (2006), Lyle (2007), Kremer and Levy (2008), and Stinebrickner and Stinebrickner (2008). These studies have found some evidence that more able roommates are beneficial; the most recent studies have suggested that the mechanism of roommate peer effects may operate through study time and alcohol use, which are more closely related to the attitudes and personalities of roommates than to their innate academic ability.

Project Overview

As noted in the introduction, this study of classroom peer effects involved two related methods of analysis: a qualitative study consisting of conversations with experienced core-course instructors and a quantitative study of the statistical relationship between measurable characteristics of classmates and measurable student outcomes. We view these two components as strong complements.

The quantitative analysis allowed us to look for formal, statistical evidence of classmate peer effects. However, the data that could be collected limited the student characteristics and outcomes that could be examined statistically.

The qualitative study, though it did not allow the same degree of formality, enabled us to move beyond the limitations of the data. The qualitative examination allowed us to peek inside the peer-effects box to explore a range of peer-interaction phenomena, albeit through the lenses of particular professors’ recollections. Thus, one aim of the qualitative study was to provide a richer context through which to interpret the statistical results. Another was to identify promising aspects of peer effects that might be approached statistically through collection of additional quantitative data.

The statistical analysis is summarized briefly below, followed by a more detailed examination of the interview part of the study. More details of the statistical analysis can be found in Parker et al. (2010).

Focus on first-year core courses

We examine the effects of classmates in required, first-year, core classes at Lewis & Clark, Reed, and Whitman Colleges. Several reasons motivate the choice of core classes:

- These courses are required of all entering first-year students, so we need not worry about students with particular characteristics choosing to take (or not to take) the courses. We can therefore claim to be studying the entire (non-transfer) student populations of the three schools rather than an idiosyncratic subset.
- Student self-selection of classmates, whether explicit or implicit, may introduce bias into quantitative and qualitative assessment of peer effects. When students and their peers are similar in dimensions such as interests, motivation, and ability that cannot be perfectly measured (and thus for which we can't perfectly control), it is impossible to distinguish statistically between the effects of a student's own characteristics and the effects of the peers. Because of this problem, most statistical analyses of peer effects in higher education have looked at roommate rather than classmate peers. In the core courses that we study, students are assigned into sections of roughly homogeneous size by a process that is largely or wholly random. This allows us to avoid one of the most difficult problems in examining peer effects.
- Core classes involve teaching students the *process of learning* in a central way alongside the teaching of course content; students learn the reading, writing, and discussion skills it takes to be a college student as well as learning the core material. This claims for these courses a central position in the curriculum, which in turn suggests such core courses have an outsized influence on a student's academic career.
- Classmate peer influence seems likely to be a significant factor in class discussions, which are the norm in these core courses. While there are certainly other contexts in which such students learn (study groups, dorm discussions, etc), class discussion figures more prominently in such classes than in homework- and/or lecture-oriented classes. Classmate interaction should thus be more observable in these classes, both qualitatively (by the professor) and quantitatively (because the data can be gathered specifically on the class).

Looking at core courses naturally leads us to emphasize the kind of peer effects that arise in classroom discussions. This is not to deny that there are other forms of significant peer interaction, and we did gather some incidental evidence about non-core courses from some of our interview subjects. However, extending the analysis beyond the core-course environment lies beyond the scope of the present study.

Courses Studied

The study looks at three selective liberal-arts colleges located in the Pacific Northwest: Lewis & Clark College and Reed College, both in Portland, Oregon, and Whitman College in Walla Walla, Washington. Although there are important differences among the schools, all three are small, selective, and strongly residential, all three offer curricula dominated by traditional liberal-arts disciplines, and most importantly for our purposes all three require first-year students to take a discussion-based core course with a common syllabus. Also crucial for our analysis was access to adequate historical records of individual student characteristics and achievement, which was facilitated by the institutional research staff at all three colleges.

The core courses we study vary in topical content and to some degree in format (although all emphasize student discussion), but they share common goals. Central aims of each school's core course include:

- Teaching students to read carefully and think critically,
- Teaching students to write clearly, and
- Teaching students to learn from and contribute to an intellectual conversation.

In each case, the syllabus for the course was either identical across sections or had extensive shared content. In some cases, exams or paper topics were common across sections. Table 1 summarizes some characteristics of the three core courses.

Statistical Evidence on Peer Effects

Our statistical approach to examining peer effects uses a regression equation of the form

$$y_i = X_i\alpha + P_i\beta + u_i, \tag{1}$$

where y_i is an academic outcome measure for student i , X_i is a set of controls for student i 's own ability, and P_i measures characteristics of student i 's peers. The coefficient vector β measures the effects of the peer characteristics on the student's outcome, controlling for the student's own ability.

The most common peer-effects hypothesis is that some or all students benefit from having more able peers. This is tested by including one or more measures of peer ability in P and examining whether the corresponding elements of β are positive. Other peer-effect hypotheses are less common but can be examined within the framework of equation (1). For example, by including a measure of the dispersion of peer abilities in P one can test whether students gain or lose from having a wider range of abilities among peers. By including a measure of the number (or share) of peers who have ability levels similar to student i , one can test the “like-me” hypothesis that students benefit from having peers that have a similar ability level to their own.

This study uses classmates in core courses as the peer group. Our outcome measures are grades in courses other than the core course, which is omitted because variations in peer quality are likely to bias the relationship between learning and core-course grades by affecting the instructor’s curve. We control for admission credentials of the student and for some non-peer elements of the student’s core-course experience—in particular, for core-course instructor and year. Our various measures of peer characteristics attempt to capture relevant properties of the distribution of classmate abilities.

Variables and data

We examine five grade-point average outcome measures to test for core-course peer effects, all of which exclude the core course to avoid confounding peer effects with the effects of the instructor’s grading curve.⁴ First we use overall cumulative GPA, which includes all courses (except core) taken at the college.⁵ Next, since the effects of the core-course experience may be more important during the year in which the core course is taken and the year after, we examine first-year and first-and-second-year GPA.⁶

⁴ Elimination of the “curve” effect in the core course is crucial to the identification of peer effects. If instructors grade on a curve, then students with high-quality core-course peers will tend to get *lower* grades in the core course, even if they learn more through the presence of excellent peers. Excluding the core-course grade prevents this effect from confounding our peer-effect estimation.

⁵ All GPA measures exclude transfer work and work taken on overseas programs not taught by regular college faculty. Courses taken on a pass-fail basis and courses from which a student withdrew are also excluded. All courses are evaluated on a 4.0 scale with 0.3 added (deducted) for a plus (minus) grade, so an A– counts as 3.7 and a B+ as 3.3.

⁶ The first year is taken to be the fall semester in which the core course is begun and the following spring semester. The second year comprises the fall and spring semesters of the next academic year. For a student who took a year off between freshman and sophomore year, only the freshman year would be counted in the

Finally, the core-course experience may have a stronger impact on courses in related subjects or that use related teaching methods than in other courses. We use a narrow and a broad version of a core-related GPA to measure success in these courses. The narrow measure consists of courses in fields directly related to the disciplines in the core course.⁷ The broader measure includes all courses except those in the natural sciences and mathematics.

We examined numerous potential measures of peer ability and numerous aspects of the class distribution of these measures. Because the available univariate measures of student ability (SAT scores, high-school GPA and class percentile, and admission-office ratings) were not highly correlated with one another, we prefer to measure the academic aptitude of a student entering the core course by a linear combination of these univariate measures rather than any single measure. The linear combination we choose is based on a regression of cumulative college GPA on the set of measures, so our academic aptitude variable is a student's predicted GPA conditional on his or her admission-file characteristics.

Missing data

College databases (and indeed colleges themselves) are not operated for the convenience of subsequent analysis by economists. One manifestation of this is the incomplete information available on many students. In some cases this occurs due to external constraints such as high schools that do not report their students' class ranks. Sometimes it results from institutional policies such as making submission of admission credentials such as SAT scores optional. In other cases it happens when students and applicants fail to answer questions about their ethnicity.

Approximately 70 percent of the students in our sample have complete data for the variables in which we are interested. Missing-data problems are not uncommon in economics and in many applications we can simply ignore the missing cases and estimate our regressions using the sample that is complete, provided that the selection of available cases did not introduce bias. The missing-data problem is more severe for us, however, because our goal is to characterize the distribution of student abilities in an entire class. If we have data on all but one or two of the students in a class section, our picture of the class distribution of abilities is likely to be quite accurate. But with one-third of the average class section missing, the description of the class ability distribution based solely

first and second years. We chose this definition to maintain temporal proximity of the courses whose grades are measured to the core course.

⁷ We solicited the opinions of the core-course chairs at each campus in making these judgments. The courses deemed to be closely related were largely courses in history, literature, philosophy, religion, the arts, and related fields.

on complete cases would be useless. In order to obtain a predicted GPA value for all or nearly all students in each class, we must employ one or more of several methods for estimating models with missing cases.

For most instances of missing data, such as a high-school not reporting class rank, we can safely assume that the data are “missing at random”—the fact that the variable is missing is not related to the value of the variable. In this case, the method of multiple stochastic imputation is appropriate.⁸ The one notable exception is SAT scores at Lewis & Clark College. Because Lewis & Clark accepts but does not require SAT scores from applicants, we think that students with low SAT scores are more likely to be missing from the data set than students whose SAT scores are high. With data “not missing at random,” we must estimate separate regressions for two sub-samples, one with SAT scores and one without.

Statistical results

Before testing the effects of specific characteristics of classmate distributions, it is useful to examine whether our outcome measures tend to be correlated among members of individual class sections. Analysis of variance applied to our GPA outcomes consistently shows strong statistical significance for core-course section, indicating strongly that *something* about the core course experience does have a significant effect on academic performance in other courses.⁹

Of course, the core-course effect could be due to many causes, including the effectiveness of the instructor. But even controlling for the identity of the instructor, year of entry, and individual student characteristics, the core-course section is still statistically significant at two of the three colleges.¹⁰

Clearly, something about the core-course classroom seems to be affecting our outcome measures in a statistically significant way. The next step is to search for any characteristics of the peer distribution that might be the source of that effect. We were unable to find any measurable characteristics of the classroom distribution of academic credentials that had a statistically (or practically) significant effect on our outcome measures.

Table 2 shows the estimated coefficients on classmate mean predicted GPA. The dependent variables in these regressions are the five GPA outcomes shown in the left column. In addition to

⁸ See Rubin (1996) for more information about this method. Our statistical results were generated using Stata 11, which incorporates algorithms based on Royston (2005).

⁹ Of the fifteen tests statistics (for five outcome measures across three schools), only one has a probability value above 0.01. For details, see Table 5 of Parker et al. (2010).

¹⁰ These results are in Table 6 of Parker et al. (2010).

peer mean aptitude, the regressions include controls for the student's math and verbal SAT scores (except Lewis & Clark no-SAT sample), high-school GPA, high-school percentile, admission rating (not available for Whitman), and batteries of dummy variables for gender, financial-aid status, ethnicity, foreign/domestic student, year of entry, core-course instructor, and area of major (to control for differing grading standards across majors). The coefficients on the student admission-credential controls (not shown in Table 2) have the expected signs and strong statistical significance.

None of the estimated coefficients on the peer mean aptitude is statistically different from zero; many are negative and many are positive. From these estimates, it appears that any peer effects resulting from mean classmate abilities in the core courses have no effect on students' subsequent grades.

Moreover, these results fairly represent the large number of variations that we estimated. We examined the effects of the dispersion (standard deviation or interquartile range) of the peer distribution, the representation of ability quintiles in the class distribution, and many other measures without finding any attribute of the ability distribution that had a systematic effect on student performance. We examined peer ability distributions based solely on SAT composite scores and on admission-office ratings (where available); once again, there was no evidence of ability-based peer effects. We performed regressions for sub-samples of students separated by predicted academic ability, gender, and other characteristics without finding any significant effects of classmate peers. For more details on these tests, see Parker et al. (2010).

Evidence from Instructor Interviews

As discussed above, the instructor interview part of our study was intended as a complement to the statistical analysis. We hoped that the interviews would help explain the mechanisms underlying any peer effects that appeared in the statistical analysis and suggest productive future directions for statistical research. Given that our regressions suggest that there are no important peer effects associated with classmate ability, we look to the interviews to explain why.

Interview methodology

We gathered information for the qualitative study through interviews with core-course instructors at all three schools. Our aim was to conduct ten interviews at each campus. We succeeded in conducting ten personal interviews at Reed and Whitman, plus one additional instructor at Whitman who was unable to keep the interview appointment but answered questions in writing. At Lewis & Clark, we were able to complete only eight of the ten scheduled interviews due

to illness and scheduling conflicts. Thus, we have in-depth responses from 29 core-course instructors at the three schools, in addition to the insights of the three team members who regularly taught and chaired these courses.

In selecting the interview subjects from the pool, we sought breadth at each school in age, sex, and home department. Although most of our interviewees had taught the course many times, we also wanted a few younger faculty members who were relatively new to the course to see if their “freshness” led to any useful insights. While most interviewed faculty members were tenured, a few were untenured assistant professors and some were non-tenure-track instructors who have specialized for many years in teaching the core course. To achieve field breadth, we interviewed instructors from a wide variety of academic departments.

In choosing the size and composition of the project team to conduct the interviews, we considered several important criteria. First, it seemed essential to have at least one person from the interviewee’s own college present to provide a local context for the questions and a familiar face to the project. Second, we wanted to have at least one person take part in *all* of the interviews to assure that a similar interview protocol was followed for all sessions and so that responses across individuals and across colleges could easily be compared. Finally, we wanted the interview team to be small enough to allow the interviews to take place in faculty offices.

These criteria seemed best served by an interview team of two, consisting of the principal investigator (for consistency and comparison) and the project team member from the local core-course faculty (for local expertise and familiarity).

To allow a mid-course refinement of the procedure and questions, we conducted the interviews in two batches. Three faculty members at each school were interviewed during March and early April 2007 in a set of “pilot” interviews.

Our goal was to complete ten interviews at each college. We had a pool of ten or eleven candidates at each school. The interview procedure used in the pilot interviews worked well and only slight changes to the questions were deemed desirable. We then proceeded with the remaining final batch of interviews in May 2007.

We wanted the interview subjects to have the opportunity to reflect in advance of the interview on their experience with peer interactions and to consult any records that might aid their recollection. To facilitate their preparation, we sent a set of advance questions to each interviewee; the questions are shown below:

- What do you think are the most important collective characteristics of the students in a class section that help students learn from each other?
- Can you think of particular class sections that worked well to promote the understanding of class members? What were their characteristics? Were these class sections equally effective

for all members of the class or did the effectiveness vary across student characteristics (strong/weak, male/female)?

- Can you think of 2 to 3 students over your teaching career who seemed to promote the understanding of other students? What were their characteristics?
- Were any of these students effective in promoting a positive class atmosphere but were not themselves outstanding scholars?
- What specific student behaviors have most enhanced the learning of peers? What behaviors have detracted the most from peer learning?
- If you had to name a single “model” student from the class, whom would you name? Why?

Interviews were conducted either in the interviewee’s office (at Reed and Lewis & Clark) or in the core-course team member’s office (at Whitman). We began the interview by telling the subject that we would ask permission before quoting their remarks and asking for permission to record the interview, which was granted by all.

We then described our interests with the following text: “We are interested in how the characteristics of a student’s peers affect success in learning. We wish to define success *relative to each student’s own ability or potential*, so successful peer interaction allows a student to learn more than they might be expected to otherwise. We are particularly interested in how peers’ personalities or abilities work to promote or hinder learning success relative to their own potential. These effects could be the result of individual actions on the part of a peer or could occur through larger group or full-class interactions.” We concluded the introductory phase of the interview by asking the instructor to avoid identifying individual students by name, using pseudonyms or other mechanisms to preserve anonymity.

The body of the interview consisted of a conversation prompted by our prepared questions. Most questions were posed by the local core-course team member with the principal investigator recording detailed notes and following up or filling in with occasional questions.

The questions that guided our interviews are shown below with a rough topical organization:

- *Introductory Questions*
 - *What criteria do you think are most important in assessing student learning in the core course? In other words, what outcomes will a successful student have achieved through the course? Are these success criteria reflected in students’ course grades?*
 - *Do you think that the quality of his or her peers affects a student’s learning?*
- *Successful and Unsuccessful Class Sections*
 - *What collective characteristics of the students in a class section do you think are the most important in helping students learn from each other?*
 - *Can you think of particular class sections that worked well to promote the learning of class members? How would you characterize the mix of students in these classes? Why were they effective?*

- *Were there particular class sections that worked particularly poorly in promoting learning? How would you characterize their mix of students? Why were they ineffective?*
- *Were these class sections equally effective or ineffective for all members of the class or did the effectiveness vary across student characteristics?*
- *Beneficial and Detrimental Peers*
 - *Can you think of two or three students over your teaching career whose presence in the class seemed to promote the understanding of other students? What were their characteristics?*
 - *Were the students who were effective in promoting a positive class atmosphere also themselves outstanding scholars?*
 - *Can you think of two or three students over your teaching career who benefited from peers more than others? What were their characteristics?*
 - *What specific student behaviors have most enhanced the learning of peers?*
 - *What behaviors have detracted the most from peer learning?*
- *Goals of Students in Core Course*
 - *What do you think are the main goals students hope to achieve in the core course? Do they vary across students? To what extent do you think classmate peer interaction is promoted by having similar educational goals?*
- *Peer Effects and Specific Groups of Students*
 - *Have you observed differences between male and female students in how they learn from each other?*
 - *Does the presence of members of minorities groups, international students, transfer students, or older students affect how the class learns?*
- *External Effects on the Core Classroom*
 - *Do the nature of the classroom and/or the time of day have significant effects on student learning and peer interaction?*
- *Mix of Abilities*
 - *In thinking about how the mix of student intellectual abilities affects learning, do you think that a diverse mix of abilities is better for weak students or that they learn best when most of their peers are close to their ability? How about for stronger students?*
- *Peer Interaction Outside of Class*
 - *Because the core course is taught from a common syllabus, there are opportunities for peer interactions among students from different sections. Have you observed such interactions? How beneficial were they for the students?*
- *Final question*
 - *What other things might we have asked or that you think are important in better understanding classmate effects on student learning?*

These questions were used by the interviewers to prompt the instructors to explore the broad question of peer effects, not as a strict rubric for scoring responses or bounding the conversation. During the interviews, instructors often went outside the boundaries of the specific question that was asked, either anticipating later questions or raising issues that were not covered in our set of

questions. We encouraged this and adjusted our subsequent questioning to avoid repeating topics that had been covered earlier.

After performing and analyzing the interviews, we invited the interview participants and faculty representatives from other colleges and universities with common-syllabus core courses to discuss our tentative findings at a conference, held at Reed College in May 2008. The conference featured enthusiastic discussions among about forty participants, leading us to refine our understanding of peer effects.

Interviews with instructors: Results and analysis

ARE PEERS IMPORTANT?

Yes. The vast majority of interviewees felt that peers can have a strong positive or negative influence on their classmates' learning. One stated, "I think the chemistry of an individual section has *everything* to do with whether or not some students go far." A remarkable number of the interviewed instructors admitted, however, that they had not thought much about peer effects prior to our study.

Most of the instructors were able to come up with clear examples of individual students who had influenced their classmates and examples of individual sections that were more and less successful. One example was from an instructor who had taught multiple sections in the same semester. A student from the weaker section who happened to attend the stronger class would perform at a higher level than he or she normally did in the regular section.

ATTITUDE VS. APTITUDE

Given the prominence of aptitude measures in our statistical work, we were naturally very interested in how a student's actual or potential academic ability, as measured by the student's grades or admission credentials, connected with his or her influence as a peer. Do students gain from having more able peers who achieve higher grades and enter with more impressive credentials? Do they gain from peers who are close to their own level of academic standing? Is there some other complementarity between peers?

However, a strong consensus emerged from our interviews that characteristics of the student's personality and attitude had more effects on peers than did the student's raw academic ability as measured by grades or admission numbers. While most agreed that the most beneficial peers had been average-or-better students, it was notable that in only a few cases were the stellar

peers also the strongest when measured on the usual grading scale. For example, the most outstanding peers were often not excellent writers.

The importance that the instructors attached to students' personalities may reflect the emphasis that they placed on classroom discussions in assessing peer effects. A central premise of class discussion is that all students can learn from the group: that the interaction of individuals with varying backgrounds and opinions provides a richness of understanding that is not possible to achieve individually. Thus, instructors consistently lauded peer behavior that advanced the *collective* expression of ideas and denigrated behavior that impeded it.

Academic ability surely affects the content of a student's classroom contributions: smarter and better prepared students would be expected to have greater insight on the readings and to thus make comments that have greater academic merit. While the contributions of these students to the discussion may advance the class toward academically profound answers, they do not necessarily advance the active participation of their peers in the conversation and so may not contribute to achieving the richest and deepest *collective* responses to questions. Indeed, they may have quite the opposite effect if their contributions are voiced in a superior, dismissive, belittling, or intimidating tone.

The instructors we interviewed consistently ranked as the most effective peers the students whose contributions to class discussions seemed to open up space for the contributions of others. While these students were often academically very strong, we heard about many students who were not academically outstanding but who seemed to have a strongly positive effect on their peers. A central conclusion from our analysis of the interviews is that instructors think attitude affects the quality of peer effects more than aptitude.

GOOD PEER BEHAVIOR

There were many peer characteristics that were mentioned by many or most of the core-course instructors we interviewed. The most obvious (and probably universal) beneficial behavior was simply attending and being prepared for class. In small discussion sections, absences are easily noticed by both instructors and peers. Students who fail to participate because of inadequate preparation also drag the class down, as do those who attempt to dominate despite being unprepared.

As noted above, most of the positive peer characteristics related to students' personalities and their attitudes toward the course and toward academic study in general. Many of these characteristics overlap and one would often expect to find students exhibiting one characteristic to have others as well.

- **Attendance and careful preparation for class** were at or near the top of most lists. Indeed, this positive characteristic was obvious enough that few instructors felt the need to dwell on it.
- **Maturity, leadership, and concern for the progress of the class** were mentioned in various forms by many instructors. Students who are genuinely concerned with how the class is going and who have the personal tools to take the lead in facilitating peer interaction seem to be priceless assets in a discussion-based classroom. Self-confident, humble, energetic, empathetic, and cheerful were mentioned as desirable traits often seen in these students.
- **General intellectual curiosity** was almost universally listed as a desirable peer characteristic. Students who like to think about abstract questions and learn for learning's sake tend to bring a positive influence to class discussions. They are good role models for others who are less convinced of the merits of the intellectual enterprise. Several interviewees used the term “playfulness” with ideas to describe beneficial peer behavior.
- **Respect for the opinions of other students** was another near-universally mentioned characteristic. Students who believe that their peers' insights are valuable usually listen to them carefully and respond thoughtfully and in a way that invites further comment. This kind of participation leads to improved class dynamics. One instructor noted that students feel like they are taken seriously when one of their peers (or the instructor) subsequently refers to or builds on their comments, even if they end up disagreeing. Such behavior validates the student's contribution to the discussion and encourages further participation.
- **Demonstrated enthusiasm for the core course and its subject matter** was often described as important. Even students with general intellectual curiosity are not necessarily excited about the subjects that dominate the reading lists of core courses. Science majors, in particular, were sometimes identified as (at least initially) uninterested in the humanistic core fields. Excitement about the subject makes good peers in at least two ways: (1) such students are very likely to come to class well prepared and (2) they may convey their enthusiasm to others. Some instructors perceived the pervasive local attitude toward the core course to be quite negative; these instructors were likely to view enthusiasm for core as a positive characteristic. (The degree to which instructors were concerned about the prevalent student attitude toward the core course varied considerably across schools.)
- **Openness to new ideas and approaches** was mentioned by many instructors. Those with a willingness to revise their opinions—to see through the comments of others that their initial

ideas were mistaken—stand to gain more from peer interaction and may contribute more to it. By contrast, closed-minded or dogmatic students often find little basis for productive discussion with their peers.

- **A cooperative rather than an overly competitive attitude** toward academics was highlighted by some interviewees. Cooperative students were more interested in advancing the class conversation toward an improved shared understanding; overly competitive students often attempted to “score points” through their comments or to convince the class to adopt their opinions, which can shut down further discussion. One instructor characterized cooperative peers simply as “civil people.”
- **Willingness to speak in class** was important, though most instructors felt that shy students could contribute positively without making frequent contributions. “Extroverts that are not narcissists” was how one characterized the ideal peer personality. Several instructors noted that students with backgrounds in acting were often classroom catalysts because they were uninhibited participants. Our interviewees’ opinions about the quiet peer are discussed in more detail below.
- **Taking chances** was viewed as important by many instructors. When recalling excellent peers, many described situations in which a student “exposed himself” by “being puzzled out loud” or “talking through an idea in front of the other students.” By putting such an idea or question in front of the class, the student provided a focus for productive conversation among his or her peers. One instructor pointed out that by being uncertain or even wrong, these students could relieve the pressure that other students may have felt to be brilliant all the time.

DETRIMENTAL PEER BEHAVIOR

To a large extent, the behaviors described as destructive were the opposite of those presented above as good peer behaviors. However, it is important to distinguish between students who merely fail to engage in beneficial behaviors and students who engage in behaviors that instructors reported as having actively negative influences on peers.

- **Dismissive or judgmental behavior** was a common characteristic of a destructive peer. Highly perceptive comments by high-ability students can be positive or negative depending on the tone with which they were delivered. Students who perceive themselves to be better than their peers sometimes discourage discussion in a number of ways: by talking down to the rest of the class, by attempting to “score points” by demonstrating their intelligence, by

making comments with a “tone of authority,” or by attempting to engage in a dialog with the instructor to the exclusion of peers.

- **Rigidity of views** may present difficulties. Good discussion involves being willing to admit that you might be wrong, which is difficult for some students, especially on sensitive topics. For example, some instructors reported that students with strong religious views had particular difficulties discussing religious texts dispassionately.
- **Rejection of peer learning** was reported by some instructors, with the prevalence varying widely across colleges. Such students wanted professors to tell them the “right” answers rather than engaging in a group conversation. Several instructors described a student attitude that might be represented as “I’m not paying all this money to listen to a bunch of 18-year-olds.”
- **Excessively loquacious** peers (sometimes unintentionally) could stifle participation by other students. Some such students always have to be the first to speak, others keep talking endlessly (often about issues only tangentially related to the text under discussion), and some repeat the same point over and over. One instructor referred to such a person as “annoying boy.” (This behavior was more commonly reported among male students than females.) Most instructors reported that they attempted to manage such situations with a private conversation with the student. One mentioned a correlation of such disruptive behavior with specific learning disabilities.
- **Active or ostentatious disengagement** that “exudes boredom” was reported by some instructors. Examples included ignoring the class discussion, looking out the window, hostile body language, checking text messages or email (or surfing the Web), and in one case performing conspicuous acts of personal grooming during class. These actions were distracting to other students and could foster or contribute to a more widespread negativity toward the class.

While most instructors were able to relate examples of detrimental behavior, few reported situations where this behavior had “destroyed” the class’s effectiveness. In some cases, the instructor reported that he or she had succeeded in managing the behavior; in others the other students in the class had marginalized the offending student. When, however, the detrimental behavior was attributed to severe social awkwardness, students often shrugged it off fairly readily.

THE QUIET PEER AND THE “NICE” PEER

A special category of peer is the student who consistently avoids contributing to class discussions but who does not display active disengagement. Nearly every instructor had taught many quiet peers and talked about them at some length. Some students are shy by nature and find it very difficult to speak in front of others. Others have been taught to be passive in the classroom setting. One common example was international students whose educational and cultural backgrounds have conditioned them not to speak in class.

All instructors agreed that there were always a few quiet students in a class section. Most believed that this was not usually a problem. However, there seemed to be a tipping point at which too many quiet peers made it difficult or impossible to sustain a productive discussion. Many veteran instructors had experienced at least one such section in their years of teaching.

Most instructors hesitated to find fault with quiet peers and seemed generally sympathetic. Many described positive behavior by quiet peers: active listening, taking careful notes, and showing clear, if silent, signs of engagement. Some noted that the quiet student’s occasional contributions were often given disproportionate attention and impact by his or her peers: “when the wallflower finally speaks up, it is more appreciated.” One instructor felt that the quiet but engaged students often learned the most from their peers, carefully writing down everything that was said in class and using it in their subsequent papers.

Another special category of peer that was described by some instructors is the student who is “too nice” to disagree with a statement made by another student. The prevalence of this behavior seems to vary dramatically across the three colleges. Although most instructors agreed that civility was a positive and important norm of classroom demeanor, some students seemed to carry this too far. “They don’t want to disturb the waters of the friendly community (so) they ... let points of intellectual disagreement pass or even avoid critically thinking about what someone is saying.”

Successful and unsuccessful class sections

A discussion class is more than a simple aggregation of its individual students. While a class filled with students exhibiting good peer behavior (by the criteria above) would likely be a very effective one and a class of destructive peers would almost surely be disastrous, all instructors agreed that the impact of any individual peer or behavior depends crucially on the peer group within which it occurs. A generally enthusiastic and active group can probably ignore or ostracize an actively disengaged classmate who might disrupt a quieter or more marginally engaged class.

Moreover, there may be various “roles” played in a classroom discussion: initiator, lightning rod, advocate, skeptic, joker, curmudgeon, peace-maker, etc. Just as 80 brilliant oboe players do not

make an orchestra, a class full of lightning rods, skeptics, or jokers is sure to fail. A successful core-class section may require not just “good” individual peers, but also a mix of students who can fill the necessary roles. One instructor pointed out that “We can’t have all questioners; we need to have some answerers as well.” Another noted that students seem to settle into the roles of “first to speak,” “first to respond,” “get involved later,” and “not speak at all unless called on” roles.

A mix of student interests (or prospective majors) seems desirable in a core class as well. With a diversity of backgrounds and interests, each student may get a chance to be excited by the material and shine in class on a different day. “X might be great at the philosophical text and less good with the poetry, and that might be Y’s forte.” A diverse set of background experiences was also thought desirable. One instructor said that in an ideal section “you have some kids from small towns, some kids from cities. You have some kids who went to religious schools or are themselves very religious and you have some who like to think of themselves on the cutting edge of atheism.”

The instructor also contributes heavily to the classroom mix of personalities. One role of the instructor in a discussion is to step into an unfilled role: initiating conversation if others do not, making a joke when levity is needed, doubting the consensus if skeptical viewpoints have gone unspoken, or calming conflicts between forceful students.

Instructors, like their students, adopt some roles more naturally than others, so a naturally skeptical instructor would complement a class lacking students to fill this role. Instructors differ in their ability to manage various potential disfunctionalities in the classroom as well. Some instructors thrive on lively, even heated, exchange while others might lose control of a class where discussion became too spirited. Some instructors are more able to draw quiet students into the discussion than others, for whom a quiet group of students might never get started.

Because of the complexities of classroom interaction, we tried to probe beyond individual actions to learn about the group characteristics of successful and unsuccessful class sections. Instructors were generally able to recall specific class sections that they viewed as more (and less) successful than others. When asked what made sections successful, they often described the contributions of individuals in the class, referring to some of the positive individual behaviors described above. Many discussed the negative impacts of individuals’ dysfunctional behaviors as leading to unsuccessful sections.

Analyzing the interviews, the most common definition of a good section seems to be one that can sustain academically productive class discussion involving the majority of students. Toward this end, the most important “role” that instructors commonly invoked for making sections work was the student whose own contributions were of a kind that encouraged others to get excited about the class and join the conversation. A few such positive peers could make a class successful. Those

whose class contributions (either through participation or active non-participation) trivialized the class or discouraged peers from speaking were at the center of most unsuccessful sections.

In addition, some instructors commented on the importance of class bonding. It was not necessary for students to be close personal friends, but feelings of familiarity and acceptance led to “academic friendship” among members of successful classes.

Students were also more satisfied, comfortable, and confident if they felt that theirs was a “good class.” This class self-image can be formed early in the year and may be difficult to reverse if it becomes negative.

A CONCENTRIC-RING MODEL OF PEERS

An image evolved during the course of our conversations that describes the initial composition of a class section in terms of three concentric rings. In the center is an active and engaged group—those naturally “good peers” who demonstrably take the enterprise seriously, model good academic performance, and encourage their peers to join in the classroom intellectual discourse. In the outer ring are any students who are predisposed to be actively disengaged or hostile on the periphery. Lying between them in the middle ring is a potentially large group of initially “uncommitted” students who are not disposed to strongly positive or strongly negative peer behavior.

We discussed above many different characteristics of good and detrimental peers. These underlying student characteristics vary on a multivariate continuum that resists facile reduction to a single measure of “good-peer-ness.” Thus our attempt to depict this complex web of factors in a single dimension must be viewed as a simple abstraction. Moreover, while there was consensus that students vary considerably in their predisposition to be good peers, drawing any kind of borders between good and neutral or neutral and detrimental is clearly arbitrary.

Nonetheless, the concentric-ring model provides a useful framework for describing two closely related phenomena that we heard described by many instructors. The first is the “critical-mass hypothesis” that a class is successful if the central group is large enough to sustain good class discussion with occasional contributions from the middle group and despite potential distractions from the periphery. This hypothesis is consistent with a static view of student behavior that is determined mostly by personality and background and that does not change much within the core class. If enough of the students exhibit good peer behavior, then the class will be successful for most students even if those outside the center never become strong participants themselves.

The second is the “gravitational-attraction hypothesis” that students in the middle ring tend to be drawn over time either into the center or toward the periphery. Successful classes are those in

which the center is large enough (and the periphery small enough) to attract a large share of the class into positive peer engagement. This hypothesis emphasizes the dynamic evolution of student behavior over the duration of the core class in response to the behavior of peers. Whereas the central group sustains the class by itself under the critical-mass model, it does so in the gravitational model by inducing additional students to become active participants.

CRITICAL-MASS HYPOTHESIS

All interviewees agreed that class discussions require a minimum critical mass of active participants to be successful. The number of active contributors that instructors felt was necessary in a class of 15–20 ranged from a low of three to a high of six or more. Those who viewed student participation as being largely determined exogenously by their innate personalities tended to view the success of a class in terms of having enough outgoing students to sustain lively interaction.

Having a critical mass of participants of varying levels of sophistication can serve to anchor the conversation at a level that is accessible to all students (and so may encourage all to participate). The mass may, for example, often “translate” for the benefit of weaker members of the class the input of very strong students, whose contributions might otherwise not elicit responses (and who might therefore cease to contribute). Equally, they may translate the input of weaker students by adding evidence or refining terms, so that the whole class can benefit.

In terms of the concentric-ring model, students in the center group are consistent contributors who can be counted on to be active in class. The peripheral group does not participate positively. We characterized as the middle group those who contributed sporadically, either offering an occasional comment regularly or varying between active and inactive participation depending on the topic, their daily degree of preparation, or other factors.

According to the critical-mass hypothesis, a successful class is one with a sufficient number of students whose engagement and personalities make them active, central participants. Detrimental peer behavior was not frequently discussed by interviewees promulgating the idea of critical mass, so the distinction between the districting peripheral group and the middle was less important.

GRAVITATIONAL-ATTRACTION HYPOTHESIS

For many instructors, the success of the class seemed to depend on *relative* size of the center and the outer ring and on the direction toward which this middle ring of students is pulled. Many interviewees reminded us that the students entering these courses are 18 years old and are experiencing college for the first time; they are “like unopened blossoms” in the phrase of one instructor. They do not know the norms of behavior that are expected in a college classroom and

have not yet formed their academic identities. This means that their behavior is likely to evolve considerably over their first year as these academic identities develop.

Peer pressure and instructor influences within the core classroom will play an important role in teaching new college students behavioral norms and in helping them to define their identities as students. This is one reason that we might expect strong peer effects to emerge from these classes. As one instructor put it, by the end of the first month the students will have decided “whether it’s going to be cool to play” the academic game.

In terms of the concentric-ring model, if most of the middle ring “buys in” and is attracted toward the center then the majority of the class participates in discussions and the class is likely to be successful. If the middle ring “opts out” toward the periphery then classroom interaction will be strained because the number of active participants is small.

Many factors influence the middle ring of students. Having a sufficient number of actively engaged participants can instill a norm of intellectually active behavior in these wavering, middle-ring students, pulling them into the active group and making a highly successful class. When there are too few students who are initially predisposed toward active engagement, it is harder to demonstrate intellectually productive peer-with-peer interaction, and thus more difficult to establish active discussion as a valuable norm. Moreover, if active peers are too sparse, they may be seen by the majority of students as exemplary in a way that does not encourage imitation—they may be viewed as simply smarter or better readers—and instead shuts down conversation.

With a solid and sizable center establishing a centripetal norm of academic engagement, it may be possible for a class to shrug off the centrifugal influence of actively disinterested or even hostile members. However, if a large proportion of the class is disengaged and non-participatory then silence, indifference, or rejection of the class can become accepted as the norm, leaving the small central group as the lone active participants in a less-than-successful class. Some instructors reported that one sufficiently disruptive student had sometimes been enough to derail an entire class.

External affinities of middle-ring students with center or periphery students may also affect whether they become active participants. One instructor noted a class with several athletes in which excellent peers drew their teammates into positive roles.

Conclusions from faculty interviews

Despite the diversity of responses that we heard from our interviewees, we were able to draw tentative conclusions about some aspects of peer effects in the core courses at our three institutions.

- **Peers matter.** Although many had not explicitly considered peer effects prior to our study, all the instructors we interviewed believed that peers affect students' learning. They were particularly aware of peer effects through participation in classroom discussions, which affects fellow students' learning both of content and of the process of academic discourse.
- **There are specific characteristics that make good peers.** There was near-universal consensus about the individual characteristics and behaviors through which good peers enhance students learning.
 - **Personality and attitude are crucial.** The most important peer characteristics seemed to be one associated with students' personalities and their attitudes toward learning and toward the class. The most stellar peers were often A- or B+ students, but they always behaved in ways that stimulated others' interest in the class and invited their active participation in discussions.
 - **Good academic quality is also important.** The best peers were never C students and often had A or A+ ability.
- **A critical mass of strong peers is important.** Many instructors highlighted the importance of having a critical mass of positive peers who were capable of elevating the quality of discourse to a high level. That is, the composition of the class as a whole matters as much as the individual characteristics of members.
- **Classmate role models are very important.** Many students do not come into core class with established academic identity. They establish it partially in the core class, and partially on the basis of role models in that class. A central group of good peers can convince others that it's "cool to play" the academic game, drawing them into the center and promoting a successful class.
- **There are diverse roles to be played in a successful classroom,** so a diversity of personalities and backgrounds is good. Even though most instructors could describe an "ideal peer," it was not always clear that one needed, or even wanted, a whole class full of them. A class was more than a simple sum of the individuals in it. While in an "ideal class," all students would be careful readers, excited about the course, and civil in demeanor, such an ideal class would also need diversity: both questioners and answerers, both skeptics and consensus-seekers, both quick responders and careful, deliberate thinkers.
- **Males and females tend to play different roles in the classroom.** We heard many instructors characterize distinct academic personalities that they associated with male and female students. Males were more aggressive and tended to focus on content; females were

more cooperative and tended to be more concerned with process. Excellent peers who were identified by instructors were about equally divided between males and females. Those identified as destructive peers were more often male.

Conclusions

Combining the evidence from our statistical analysis and the instructor interviews, we have strong evidence that within a selective college, having a class of students with strong admission credentials does not automatically lead to better learning by an individual student in the class, given his or her own characteristics. This, in itself, is a strong conclusion. It casts doubt on the pedagogical benefit of targeting merit-aid money toward students solely on the basis of academic credentials.

Of course, our study is limited by the fact that nearly all of the students we studied have relatively high credentials. Thus, our conclusion may hold only within the usual range of applicants and might not apply if a college were able to use merit aid to significantly raise its entire student body. Nonetheless, it does suggest that a class drawn disproportionately from the upper tiers of a college's normal application pool is not likely to lead to greater learning for any individual student in the class than one drawn evenly from the distribution as a whole.

The interviews lead us to believe that the reason for our weak statistical results is that we are measuring the wrong peer characteristics. If peer attitudes and personalities have more important influences on classmate learning, then we should be more successful in finding statistically significant effects if we can measure those characteristics. College databases have only fragmentary information that might help us examine the effects of peer attitudes and personalities.

If early-decision (ED) applicants are "better fits" for the college and have more positive attitudes, then it is possible that having more classmates who applied ED would be beneficial. We tested this hypothesis and found no positive evidence that more ED classmates affected our outcome measures.

For two of our schools, we had a general admission-office rating, giving an overall assessment to the applicant's file. This rating is based partially on SAT scores and high-school record, which are part of our database, and partially on letters of recommendations and student essays, which we cannot directly quantify. By taking a "residual" from a regression of the admission-office rating on our other measured applicant characteristics, we were able to construct an indirect measure of the quality of the letters and essays, which might be a better gauge of student attitude and fit for the college. Again, we found no statistically significant effect of having classmates with higher admission-rating residuals.

In order to test for attitude-related peer effects statistically, it will probably be necessary to have direct data on incoming students' personalities and/or attitudes about college. This could be done by surveys of entering students, either using instruments specially designed for this purpose or perhaps by extracting relevant information from the freshman survey designed by the Cooperative Institutional Research Program (CIRP) and administered by many colleges.

A small pilot study of three sections of Humanities 110 at Reed College was conducted in 2009–10. Although the sample was too small for statistical analysis, the results from before-and-after surveys of students and instructors strongly corroborated the conclusions we drew from our interview analysis. When asked to identify the peer behaviors that aided their learning, students overwhelmingly selected the same attitude-based behaviors that our interviewed instructors suggested were most important and consistently viewed academic prowess as less important.

The end-of-year survey asked instructors to identify which students in the class had engaged in various peer behaviors. Those students who had engaged in several of the behaviors that students felt were beneficial to learning were then identified as “beneficial peers.” Looking at the pre-course attitudes reported by the beneficial peers, there was a strong tendency for the beneficial peers to have had (1) a positive attitude toward the humanities course and (2) a reported pre-course expectation that they would participate actively in class discussions.

The results of this small survey study suggest that a larger study that might measure students' pre-college attitudes and expected class participation could shed additional light on the nature of peer effects. In particular, such a study over multiple years could allow the examination of subsequent student outcomes in relation to the distribution of peer attitudes and the reported quality of the core-course peer interaction.

Tables

Table 1

	Lewis & Clark	Reed	Whitman
Title of course	Inventing America	Classical Humanities	Antiquity and Modernity
Duration and syllabus	Two semester courses: Fall is common syllabus; students choose different topic-based sections in spring	Year-long course; few students change sections	Year-long course; very few students change sections
Student credit	4 semester hours each semester	6 semester hours (1.5 Reed units) for each semester	4 semester hours each semester
Contact hours per week	3 hours of discussion; occasional optional lectures	3 hours of discussion; 3 hours of required lecture	3 hours of discussion; occasional optional lectures
Content	Fall: Classical to modern, emphasizing political philosophy relating to founding of U. S. and to groups left out of founding process	Classical philosophy, history, and literature; Greece in fall and Rome in spring	Philosophy, history, and literature, focusing on classical in the fall and early modern/modern in the spring
Section size	Discussion sections of 18–19 (slightly larger in early years).	Discussion sections of 14–17.	Discussion sections of 16–18.
Scheduling of sections	Most at common afternoon time. A few at morning time for in- season athletes and students with afternoon labs.	Sections are offered at many times and days.	All sections meet at common morning time.
Section choice	Students rank-order four preferred sections/instructors.	Students register for section/time without knowing instructor.	Students are randomly assigned to sections by registrar.
Criteria for evaluation	Writing, class participation, and exams.	Writing, class participation, and exams. Single grade for full year.	Writing, oral reports, discussion participation, exams. Semesters graded separately.

Table 2

Dependent variable is GPA (excluding core) in:	Coefficient (standard error) on peer-mean predicted GPA			
	Lewis & Clark		Reed	Whitman
	SAT	No SAT	Full sample	Full sample
All courses	-0.183 (0.119)	-0.144 (0.182)	0.119 (0.151)	-0.0445 (0.110)
First-year courses	-0.215 (0.148)	0.134 (0.231)	-0.0580 (0.174)	-0.0143 (0.129)
First- & second-year courses	-0.201 (0.127)	0.129 (0.196)	-0.0840 (0.161)	-0.0399 (0.113)
Narrow core-related courses	-0.142 (0.161)	0.221 (0.281)	-0.0043 (0.174)	-0.0582 (0.125)
Broad core-related courses	-0.165 (0.122)	0.113 (0.182)	0.0824 (0.161)	-0.0896 (0.112)

References

- Ammermueller, Andreas, and Jörn-Steffen Pischke. 2006. Peer Effects in European Primary Schools: Evidence from PIRLS. In *Working Paper*. Cambridge, Mass.: National Bureau of Economic Research.
- Angrist, Joshua D., and Kevin Lang. 2004. How Important are Classroom Peer Effects? Evidence from Boston's Metco Program. In *Working Paper*. Cambridge, Mass.: National Bureau of Economic Research.
- Boozer, Michael A., and Stephen E. Cacciola. 2001. Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data. In *Discussion Paper*. New Haven, Conn.: Yale University Economic Growth Center.
- Burke, Mary A., and Tim R. Sass. 2006. Classroom Peer Effects and Student Achievement. In *Working Paper*. Tallahassee, Fla.: Florida State University Department of Economics.
- Dills, Angela K. 2005. Does Cream-Skimming Curdle the Milk? A Study of Peer Effects. *Economics of Education Review* 24 (1):19-28.
- Ding, Weili, and Steven F. Lehrer. 2006. Do Peers Affect Student Achievement in China's Secondary Schools? In *Working Paper*. Cambridge, Mass.: National Bureau of Economic Research.
- Epple, Dennis, and Richard E. Romano. 1998. Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *American Economic Review* 88 (1):33-62.
- Foster, Gigi. 2006. It's Not Your Peers, and It's Not Your Friends: Some Progress toward Understanding the Educational Peer Effect Mechanism. *Journal of Public Economics* 90 (8-9):1455-75.
- Gaviria, Alejandro, and Steven Raphael. 2001. School-Based Peer Effects and Juvenile Behavior. *Review of Economics and Statistics* 83 (2):257-68.
- Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2003. Does Peer Ability Affect Student Achievement? *Journal of Applied Econometrics* 18 (5):527-544.
- Henderson, Vernon, Peter Mieszkowski, and Yvon Sauvegeau. 1976. Peer Group Effects and Educational Production Functions. Ottawa: Economic Council of Canada.
- Hoxby, Caroline. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. In *Working Paper*. Cambridge, Mass.: National Bureau of Economic Research.
- Kang, Changhui. 2007. Classroom Peer Effects and Academic Achievement: Quasi-Randomization Evidence from South Korea. *Journal of Urban Economics* 61 (3):458-495.
- Kremer, Michael, and Dan Levy. 2008. Peer Effects and Alcohol Use among College Students. *Journal of Economic Perspectives* 22 (3):189-206.
- Lefgren, Lars. 2004. Educational Peer Effects and the Chicago Public Schools. *Journal of Urban Economics* 56 (2):169-191.
- Lyle, David S. 2007. Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point. *Review of Economics and Statistics* 89 (2):289-299.
- Manski, Charles F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60 (3):531-542.
- McEwan, Patrick J. 2003. Peer Effects on Student Achievement: Evidence from Chile. *Economics of Education Review* 22 (2):131-41.
- Parker, Jeffrey, James Grant, Jan Crouter, and Jon Rivenburg. 2010. Classmate Peer Effects: Evidence from Core Courses at Three Colleges. Portland, Ore.: Reed College.

- Rangvid, Beatrice Schindler. 2003. Educational Peer Effects: Regression Evidence from Denmark with PISA2000 Data. Copenhagen: AKF, Institute for Local Government Studies.
- Robertson, Donald, and James Symons. 2003. Do Peer Groups Matter? Peer Group versus Schooling Effects on Academic Attainment. *Economica* 70 (277):31-53.
- Rothschild, Michael, and Lawrence J. White. 1995. The Analytics of the Pricing of Higher Education and Other Services in Which the Customers are Inputs. *Journal of Political Economy* 103 (3):573-586.
- Royston, Patrick. 2005. Multiple Imputation of Missing Values: Update. *Stata Journal* 5:188-201.
- Rubin, Donald B. 1996. Multiple Imputation after 18+ Years. *Journal of the American Statistical Association* 91 (434):473-489.
- Sacerdote, Bruce. 2001. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics* 116 (2):681-704.
- Stinebrickner, Ralph, and Todd R. Stinebrickner. 2006. What Can Be Learned about Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds. *Journal of Public Economics* 90 (8-9):1435-54.
- . 2008. The Causal Effect of Studying on Academic Performance. *B.E. Journal of Economic Analysis & Policy* 8 (1: Frontiers):Article 14.
- Vandenberghe, V. 2002. Evaluating the Magnitude and the Stakes of Peer Effects Analysing Science and Math Achievement across OECD. *Applied Economics* 34 (10):1283-90.
- Winston, Gordon C., and David J. Zimmerman. 2004. Peer Effects in Higher Education. In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, edited by C. Hoxby. Chicago: National Bureau of Economic Research and University of Chicago Press.
- Zimmer, Ron. 2003. A New Twist in the Educational Tracking Debate. *Economics of Education Review* 22 (3):307-315.
- Zimmer, Ron W., and Eugenia F. Toma. 2000. Peer Effects in Private and Public Schools across Countries. *Journal of Policy Analysis and Management* 19 (1):75-92.
- Zimmerman, David J. 2003. Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics* 85 (1):9-23.