American Psychological Association

Journal of Educational Psychology

0022-0663/19/$12.00

2019, Vol. 111, No. 1, 91–103
http://dx.doi.org/10.1037/edu0000269

Analytical Assessment of Course Sequencing: The Case of Methodological Courses in Psychology

Laura Betancur, Benjamin Margolin Rottman, Elizabeth Votruba-Drzal, and Christian Schunn
University of Pittsburgh

Small differences in course sequencing may have broad effects on undergraduate science learning. In the current research, we developed an analytical approach for assessing questions about course sequencing using educational data sets, and we applied it to questions about the Psychology major. This study examined the relationships between student achievement (grades) in psychology courses taken before and after methodological courses. We used a longitudinal institutional dataset involving thousands of students across seven cohorts, and control for demographics, SAT achievement, and prior psychology GPA. We found that two courses were especially important: Achievement in statistics and research methods courses related to grades in subsequent advanced seminars, lab courses, and overall psychology GPA. Additionally, relations between research methods achievement and topical course grades were stronger when those courses were taken after versus before research methods, further reducing the likelihood of hidden third variable explanations. The same was not true for most other introductory courses, although it was found for biopsychology, which may be because biopsychology (which also includes neuroscience) is relevant across many areas of psychology, similar to research methods. These correlational findings suggest that requiring students to take research methods and biopsychology early on in the major, and ensuring success in these courses, may enhance subsequent learning. More broadly, this research provides a template for data-based approaches to course sequencing questions within any undergraduate major.

Educational Impact and Implications Statement

This study investigates the roles of statistics and research methods courses as part of the psychology major. It found that achievement in these courses, particularly research methods, is especially predictive of achievement in subsequent psychology courses. These results have implications for the design of the psychology curriculum, as they raise the possibility that ensuring student success in methods courses, and encouraging students to take methods courses earlier, may help students succeed in topical psychology courses.

Keywords: course sequencing, psychology major, research methods, statistics, identical elements

Course sequencing is sometimes a difficult task when designing college science majors. In some circumstances the proper sequencing of courses may be clear-cut, for example, it is necessary to take Organic Chemistry I before Organic Chemistry II. However, often there are courses (e.g., related to independent research and analytic skills) that do not strictly need to be taken prior to topical courses, but nevertheless could confer a learning benefit if taken earlier in a student's career. Furthermore, competing theoretical arguments could be made for placing different courses first in the sequence, such as skills development first (Gagne & Briggs, 1974) versus interest development first (Hidi & Renninger, 2006). In the current research, we take an empirical approach to examining statistical relations between achievement in various psychology courses and methods courses. Additionally, we consider whether these associations are stronger when topical psychology courses are taken before versus after methods courses.

This research has implications at three levels. First, the findings establish empirical patterns of inter-course transfer, a conceptually important topic in education that is rarely studied. Second, the approach we developed may be useful for researchers and practitioners in other disciplines when it comes to considering the implications of different course sequencing decisions. We use two different types of regression analyses to capture the relationships between the achievement across multiple courses, and these two types of analyses could be applied in other majors. Third, our findings provide important information about course sequencing for other psychology departments with an undergraduate major that is characterized by a similar course structure.

This article was published Online First April 16, 2018.

Laura Betancur, Benjamin Margolin Rottman, and Elizabeth Votruba-Drzal, Department of Psychology, University of Pittsburgh; Christian Schunn, Learning Research and Development Center, University of Pittsburgh.

Correspondence concerning this article should be addressed to Laura Betancur, Department of Psychology, University of Pittsburgh, 210 South Bouquet Street. Pittsburgh, PA 15260. E-mail: lab193@pitt.edu

Questions About Course Sequencing in the Psychology Major

In the current research, we are particularly interested in the role of methods courses, especially statistics and research methods, given that they are often viewed to be important for the development of critical reasoning skills used in topical science classes (American Psychological Association, 2013; Dunn et al., 2010). See Figure 1 for a definition of different kinds of courses. Methods courses are different from topical courses because their primary goal is to teach the skills of a discipline rather than the body of knowledge in a discipline or some combination of methodological skills together with a body of knowledge. In the sciences, methods courses focus on the scientific processes used to create knowledge and test knowledge claims. The courses include more than just skills; methods courses also teach students concepts such as the concept of construct validity, or the concept of multiple regression, although understanding these concepts is inextricably linked to skills (e.g., the skill of critiquing the construct validity of a study, or the skill of choosing multiple regression as the appropriate statistical test and running and interpreting the regression). In contrast, topical classes focus more on learning concepts that are not tied to methodological skills. Research and policy in K–12 science education has strongly argued for integrated methods throughout topical courses (National Research Council, 2012). But at the college level, stand-alone methods courses and stand-alone topical courses remain common practice (Hofstein, & Mamlok-Naaman, 2007; National Research Council, 2006). In the context of such stand-alone courses, we address the question of sequencing.

In psychology, there are two common forms of stand-alone methods courses: psychology research methods and introductory statistics. In this study, we examine sequencing of these methods courses in the major, and consider how early in the major students should take methods courses. At many universities (Stoloff et al., 2010), there is considerable flexibility in whether psychology students take topical courses before or after research methods and statistics. Students could take the statistics course in their first year and the research methods course early in the second year, leaving many topical courses for later. Alternatively, students could take many topical courses first, leaving statistics to the end of the third year, and research methods to the beginning of the fourth year. In practice, there are students who follow each of the extremes and many variations in between. For this reason, we sought to collect empirical evidence about the role of methods classes within the major.

Reasons why Learning in Methods Courses May Transfer to Topical Courses

The APA Guidelines for the Undergraduate Psychology Major suggests that the "methods and statistics core requirement . . . needs to be taken toward the end of the first 2 years to provide the proper research orientation for later advanced classes" (American Psychological Association, 2013, p. 11). An influential report by Dunn et al. (2010) went further and recommended that these courses be taken soon after completing introduction to psychology. Although not explicitly stated, presumably these authors believed that statistics and research methods are important because they teach skills, which may be useful for topical courses. Based on past detailed task analyses with support of computational models and expertise studies of research methods in psychology (Schunn & Anderson, 1998, 1999, 2001), these skills could include understanding statistical results (such as main effects and interactions, p values, regression weights, effect size, and how to interpret statistical results); understanding the principles of experimental design (such as what a 2×3 factorial design is, or what a "within-subjects" design is, and the strengths and weaknesses of different designs); understand what a confound is at a high level and being able to identify possible confounds when reading a research article; critiquing the construct validity of a questionnaire used in a study; and so forth. Psychology textbooks and lectures for topical

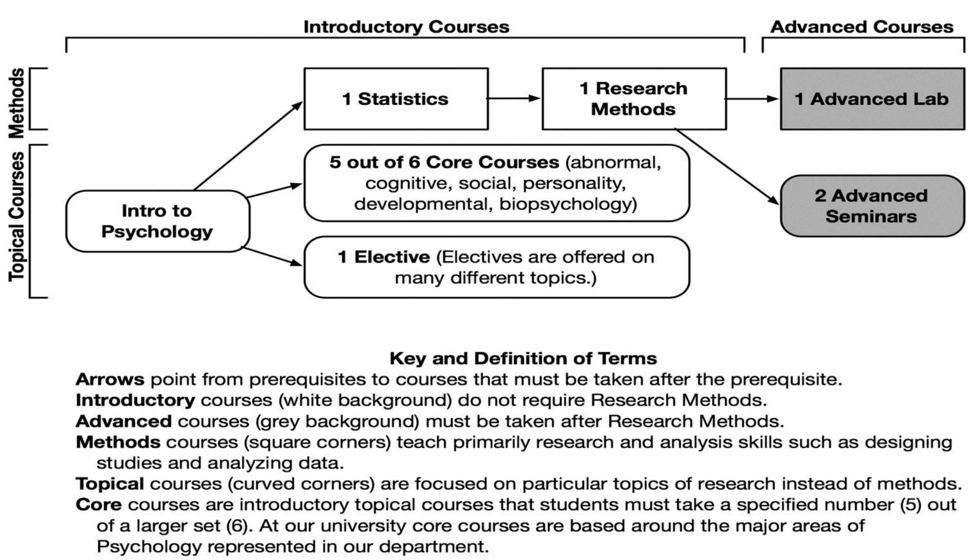

Figure 1. Prerequisite structure for the courses in the psychology major at the University of Pittsburgh. Arrows go from prerequisites to subsequent courses.

courses often present experimental results (rather than just summarizing phenomena). Therefore, understanding material presented this way requires students to have acquired basic methods knowledge and skills, even though they are often only explicitly taught in methods classes, not topical classes. Although the underlying reasoning for recommending that research methods be taken early in the major is plausible, neither the APA guidelines nor Dunn et al.'s recommendations are based on empirical findings in the literature.

Transfer across courses can be understood in terms of Identical Elements Theory (IET). IET was first proposed in the early days of behaviorism (Woodworth & Thorndike, 1901) to explain why practice in a task generally produced improvements in only that task, not improvements in even closely related tasks (e.g., estimating areas of rectangles improved without improving estimating accuracy for other shapes). This first formulation focused on transfer of particular behaviors (i.e., the unit of analysis in behaviorism). More recently, in cognitive formulations of IET, the identical elements were conceptualized as concepts and skills that could be used in other situations that drew upon these concepts and skills (Schunn & Anderson, 2001; Singley & Anderson, 1989). This conceptualization explained transfer across situations with no shared specific behaviors, such as across programming courses that shard no specific commands (i.e., Singley & Anderson, 1989) or from debugging in computer programming to debugging errors in written instructions given to movers (Klahr & Carver, 1988). Closer to the focal learning content in the current project, formal models based on IET have been used to predict the ways in which expert researchers solve novel problems in their domain (Schraagen, 1993) or expert researchers in one domain transfer skills to solve problems in another research domain (e.g., from social psychology to cognitive psychology; Schunn & Anderson, 1999, 2001).

The IET perspective provides a mechanistic explanation of why methods courses can provide foundational skills of data interpretation and data analysis that will confer learning benefits in topical courses when those later courses make regular use of those skills. Note that this transfer will likely be asymmetrical because the skills are explicitly taught in the methods courses but only used (or at least less likely to be explicitly taught) in the topical courses. Within science education research, methodological courses that teach research design and data analysis are believed to help develop scientific reasoning and critical-thinking skills (Lawson, 1999; McLean & Miller, 2010; Mill, Gray, & Mandel, 1994; VanderStoep & Shaughnessy, 1997). These skills, in turn, are believed to be important for learning psychology (Williams, Oliver, Allin, Winn, & Booher, 2003), in line with the extensive literature on the role of science practices and nature of science in science learning (Lederman, 1992; Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002; National Research Council, 2012). More specifically, Dunn et al. (2010) argue that the benefit of taking methods courses is that

> success in undergraduate psychology courses often requires competence in research design as well as data analysis and interpretation—the primary learning objectives of methodology courses. Moreover, without familiarity with the basics of scientific methodology, students' appreciation of all later course material in the major is seriously limited. (p.55)

According to this reasoning, taking and doing well in research methods and statistics should benefit learning in topical courses to the extent that the topical courses include tasks that build off of knowledge and skills learned in statistics and statistics.

However, we acknowledge that taking methods courses may influence subsequent learning in other ways aside from IET. Taking methods courses and doing well versus poorly could also change a student's motivation to study psychology, especially in terms of changing interest and self-efficacy in psychology, which are discussed in the general discussion.

Reasons Why Learning in Methods Courses May Not Transfer to Topical Courses

Although it may seem intuitive that additional exposure to statistical and research methods skills would improve achievement in topical psychology classes, the benefits may be quite small and not large enough to justify requiring these courses early in a students' academic careers. In general, transfer of learning is frequently limited (Nokes-Malach & Mestre, 2013), and forgetting can be substantial across semesters (Ruben & Wenzel, 1996). Furthermore, from an IET perspective, statistics and research methods skills would not help much in courses that focus primarily on the ultimate *findings* of a study rather than on the *methods* and data inference steps in a study, or in courses in which students are not required to read original research that describes the methods and the process of scientific discovery. Even in advanced courses, which are more likely to focus on the process of scientific discovery (especially in psychology, which makes heavy use of primary research findings in undergraduate textbooks and lectures), the emphasis on scientific process could vary substantially from course to course and instructor to instructor. The importance of statistical skills for subsequent course achievement may be particularly questionable since students typically do not perform statistical analysis in most courses. Having a good understanding of statistics could help students understand the results of studies, although again, whether these skills actually are used on tests and graded assignments likely varies across classes and instructors.

Research on curriculum coherence in K–12 education has shown the importance of consistency of topics taught within and across grades (Fortus & Krajcik, 2012; Schmidt, Wang, & McKnight, 2005), and U.S. K–12 science standards have been revised to emphasize a gradual build-up of core disciplinary ideas (National Research Council, 2012). However, research on sequencing of topical courses has not found strong support for changes to more logical orderings: Greater prior exposure to physics does not appear to improve achievement in chemistry, nor does greater exposure to chemistry improve achievement in biology (Sadler & Tai, 2007). But those larger disciplinary leaps across sciences may have different effects than coherence within a science discipline. Across disciplines, differences in notations, degree of emphasis on concepts, and other instructional differences may minimize the learning benefits of better sequencing.

Prior Empirical Evidence on the Role of Methods Courses and Open Questions

Regarding statistical and research methods in particular, studies have examined the effects of ordering research methods and statistics courses relative to each other; there is evidence that integrated statistics and research methods courses produce long-term learning benefits (Barron & Apple, 2014). Other studies at the college level have shown that concurrent laboratory experiences improve topical learning (e.g., in introductory chemistry, Matz, Rothman, Krajcik, & Banaszak Holl, 2012), but this effect may be one of making phenomena concrete rather than one of exposure to methods and statistics. Very few studies have tested the effects of methods/statistics courses on later disciplinary science courses, which is more directly relevant to the issue of transfer of skills across courses. Freng, Webber, Blatter, Wing, and Scott (2011) found that students' grades in research methods and statistics predicted grades in advanced psychology courses. In addition, students who delayed taking statistics and research methods had lower grades in subsequent courses. We sought to replicate and extend these results in a number of ways.

First, Freng et al. (2011) used data from just 129 students, and the only demographic variable included in the analysis was sex. In general, curriculum sequencing effects are difficult to study because it is unfeasible to run a randomized controlled trial in which students are randomly assigned to take methods courses before or after topical courses, precluding strong causal inference as to the role of these courses. However, there are other approaches to understand the sequencing effects in general, and the role of methods courses in particular through larger data sets with more information about potentially confounding variables and more advanced statistical techniques. In the current study, we used academic data from 2,313 students from seven cohorts. In analyzing a much larger sample across many cohorts, we can rigorously control for many demographic variables without overfitting problems, and we can test the effects across a diverse pool of faculty teaching courses. Additionally, we include a wide variety of demographic and prior academic achievement factors in our analysis to strengthen the internal validity of our findings by reducing omitted-variable bias. Student self-selection into sequencing of courses is likely to be related to prior academic achievement as well as other characteristics of students. Thus, the failure to control for important characteristics that are correlated with course sequencing and subsequent grades may strongly bias results. Second, we analyzed associations between academic achievement and statistics and research methods separately to understand the unique contributions of each; Freng et al. only analyzed the influence of these two courses together. Third, Freng et al. only studied the influence of statistics and research methods on advanced courses. Although advanced courses often focus on interpreting and criticizing empirical research studies (Dunn et al., 2010), even introductory courses in psychology include detailed descriptions of primary research studies. For completeness, we studied the influence of statistics and research methods on introductory-level courses, advanced seminar courses, and advanced laboratory and methods courses.

The current study addressed the following two specific aims. First, using regression and structural equation modeling, we considered how achievement in statistics and research methods predicts achievement in introductory, advanced seminar, and advanced lab courses while controlling for a range of student characteristics, including demographics, prior achievement before entering the university (via the SAT, the academic abilities examination used for selecting students into many U.S. universities), and prior achievement at the university (via the grade point average [GPA]). Second, we performed two additional analyses to rule out threats to validity. We establish clear temporal precedence by testing whether taking different methodological courses before topical psychology courses relates to higher grades in subsequent courses (but not prior courses). We also show that performing well in methodological courses, rather than just taking them, is more strongly associated with higher grades in subsequent courses than with grades in earlier courses. Overall, we provide a model for how university science departments can rigorously mine institutional data to address critical policy questions pertaining to undergraduate education.

The Structure of the Psychology Major at the University of Pittsburgh Compared to U.S. Universities More Broadly

Although the structure of the psychology major curriculum and individual classes at our university is not identical to majors at all other universities, our major is fairly typical on a number of dimensions. Figure 1 summarizes the requirements as well as the definitions of different types of courses.

First, at the University of Pittsburgh, introduction to psychology, except in rare instances, is a prerequisite for all other courses. Across the U.S., Intro Psych is a prerequisite for higher-level courses in 94% of programs (Norcross et al., 2016).

Second, students must take five out of six "core" courses that cover introductions to the following areas of psychology: social, personality, developmental, cognitive, abnormal, and biopsychology/neuroscience. These courses are indeed called "core" at our university, and also fit the definition of "core" course from Stoloff et al.'s (2010) and Norcross et al.'s (2016) criteria that students must take some courses out of a set but have some amount of choice about which courses to take. These six courses are some of the most frequently required topical courses in the U.S.; history and systems, learning/conditioning, and sensation and perception are also frequently required, and are offered as electives at the University of Pittsburgh (Norcross et al., 2016; Stoloff et al., 2010).

Third, students must also take introduction to statistics, and then research methods; these two courses are required for declaring psychology as their major and for taking advanced courses. The statistics course is offered through the Statistics department, not Psychology, and teaches basic principles of probability and statistics (univariate and bivariate statistics such as *t* tests, chi-squared tests, 3-group ANOVA, and correlation, but typically not multivariate regression). The research methods course emphasizes research design and analysis issues through lecture, and has students design, implement, analyze, and write up an independent research project through scaffolded activities in a lab (i.e., only a small component is cook-book lab activities). At our university, students can take any of the core courses, before, concurrently, or after statistics and research methods. Across the U.S., some sort of methodology course is required for the major in 98%–99% of psychology departments (Norcross et al., 2016; Stoloff et al., 2010). The most common sequence (34%) of methods courses is a statistics course followed by a research methods course, like at our university, or an integrated statistics and research methods course (12%), although Stoloff et al. (2010) mention that there are many other patterns of requirements including sequences of statistics, research methods, experimental psychology, and topical laboratory courses.

Fourth, students must also take two advanced seminar-style courses, which focus more on primary research centered around a particular topic, and typically have class sizes of around 35 students or smaller. These classes must be taken after research methods. Fifth, students must take one advanced laboratory or methods course, which involves either conducting a research project or learning about a specific methodology; for concision, we call these courses "Advanced Labs."[1] Lastly, students can optionally take a number of other types of courses such as an externship, directed research, and honors thesis, which are not displayed in Figure 1. Stoloff et al.'s (2010) and Norcross et al.'s (2016) surveys did not analyze the prevalence of all these sorts of courses in the same way that we define them at our university, although many universities offer similar sorts of courses.

In sum, although there is wide variability in the structure of the psychology major at different institutions, the major at the University of Pittsburgh is quite typical, which increases the generalizability of our findings.

Method

This research was approved by the University of Pittsburgh Human Research Protection Office.

Sample

Academic data from an administrative database were obtained for 2,720 undergraduate students taking psychology courses at the University of Pittsburgh. Included students were enrolled in the undergraduate psychology research methods course between the fall of 2009 and the spring of 2016, a period during which the requirement for psychology majors was not revised. Enrollment in research methods signals that a student intends to take advanced courses and is almost certainly planning to declare psychology as their major. Four hundred and seven students were dropped from the sample because their statistics grade, a prerequisite for research methods, was not available, leaving 2,313 students; this can happen if they took statistics at another institution (68% of the dropped students) or placed out of statistics by taking the advanced placement (AP) course in high school and doing well on the standardized final AP Statistics exam (32%). These students were dropped because they did not have a grade in statistics to predict subsequent achievement.

Students who were dropped were not statistically different from those who were not dropped on most of the 18 demographic characteristics, with only three exceptions according to t test, and only one was big enough to exceed Cohen's (1988) convention for a large effect ($d = 0.80$): students who were dropped started college at an older age $t(1, 3{,}619) = 16.82$, $p < .001$, $d = 0.83$, had less educated parents, $t(1, 2{,}675) = -1.92$, $p = .05$, $d = 0.08$, and were more likely to have attended a public school, $t(1, 2{,}173) = 1.96$, $p = .05$, $d = 0.11$.

Measures

Demographic characteristics. A range of demographic characteristics were included in the analysis. When enrolling, students provide the university with demographic information including gender, race, ethnicity, citizenship status, date of birth, and high school identification. Gender was represented with an indicator variable, with female as the reference group. Race was coded with five categories: White (reference group), Black, Hispanic, Asian, and Other race. An indicator of variable "traditional" versus "nontraditional" student was created to represent whether a student was enrolled at the University of Pittsburgh by age 20 or not (reference).

Multiple indicators of socioeconomic status (SES) were included in the analysis. These were created based on data that were obtained from the 74% of students whose families filled out the Free Application for Federal Student Aid (FAFSA). The indicators included the number of family members, the family's adjusted gross income (AGI) in units of 10,000 U.S. dollars, and parental education. Parental education was represented with a series of variables indicating whether both parents had less than a high school diploma (reference group), a high school diploma, a bachelor's degree, or higher.

Finally, several characteristics of students' high schools by year of graduation were obtained from the Public Elementary/Secondary School Universe Survey Data collected by the National Center for Education Statistics (NCES; 2014). These included an indicator for whether the school was public (reference) or private, the percentage of African American and Latino students in the school, and urbanicity level. Urbanicity was coded as a series of indicator variables indicating if the school was located in a community that was urban (reference group), suburban, or rural. These measures of school characteristics are commonly associated with academic achievement and educational attainment (Billings, Deming, & Rockoff, 2014; Carbonaro & Covay, 2010; Palardy, 2013), so it was important to control for the potentially confounding influence of these factors in the analysis.

Prior academic achievement. Academic achievement prior to college was captured by verbal, math, and writing SAT scores, which were divided by 100. For students who took the ACT instead of the SAT, the university automatically converts ACT scores into SAT scores and records the highest scores if students took both. Table 1 shows the demographic information and mean SAT scores.

Prior academic achievement in college was measured using course grades in students' academic records. These records identified which courses were taken each semester and the resulting grades (0–4 scale: A = 4, B = 3, C = 2, D = 1, F = 0, − removes .25, + adds .25). In addition, the mean and standard deviation of the grades within each course were obtained, allowing us to standardize the grades at the class level to correct for bias in course difficulty and teacher grading.

The cumulative number of psychology credits that a student had taken as well as their psychology GPA were calculated prior to the semesters during which they took Introduction to Statistics and Research Methods. These variables were used as covariates in different models, as explained below.

[1] Many students opt out of taking a true advanced lab, and instead take a course that technically fulfills the requirement but does not focus on the process of scientific discovery. This course was not included in this analysis of advanced labs and methods courses.

Table 1
Demographic and SAT Characteristics and Prior Academic Achievement of the Sample

Variable	*M* or %	*SD*
Sex		
Female	72%	
Male	28%	
Race		
White	75%	
Black	8%	
Hispanic	3%	
Asian	9%	
Other race	4%	
U.S. citizen	97%	
Enrolled by age 20	91%	
Number in family	3.91	1.21
AGI/10,000	11.22	9.79
Parental education		
Below high school	1%	
High school	18%	
College	81%	
High school percent of minority students	14%	17.30
High school percent private	15%	
High school urbanicity		
Urban	13%	
Suburban	66%	
Rural	21%	
SAT Verbal/100	6.14	.78
SAT Math/100	6.13	.75
SAT Writing/100	6.04	.75

Participation in research methods and statistics courses. Two indicator variables capture whether a student took each course before or after research methods and before or after statistics. Additionally, a set of dichotomous variables indicate which statistics course was taken: Basic Applied Statistics (STAT 200, reference group), which is the most common statistics course, Applied Statistical Methods (STAT 1000), or Statistics & Probability for Business Management (STAT 1100). STAT 1000 and 1100 are more challenging courses than STAT 200 and may facilitate learning and achievement in subsequent classes.

Missing Data

Out of the 2,313 students in the sample, 1,246 (54%) had complete data, as often is the case in large educational data sets. The percentage of missing data for each variable in the analyses ranged from 0.2% to 26.5%, and varied depending on the source of information. The percent of missing data was low (0.8%) for the demographic data collected by the university at the time of enrollment. Since not all students fill out the FAFSA, 26.1% have missing FAFSA demographic data. A total of 20.1% of high school data could not be obtained from Public Elementary/Secondary School Universe Survey Data because the NCES does not collect data about international schools and because private schools are not required to report their information. Since this missing data clearly cannot be treated as missing-at-random and collectively constitutes a large fraction of the students, imputation methods were used to include all of the students in the analyses. Following current best practice, missing data were imputed with chained equations (ICE) implemented in Stata 13 to create 20 complete data sets (Royston, 2004, 2005). These data sets were combined and analyzed using "mim" commands in Stata.

Analytic Approach

Mixed effects regressions with subject-specific random intercepts and fixed effects for student-level predictors were used to address the primary aims of the study. This analytic framework was chosen because the outcome variable is student grades in psychology courses. Each student has grades in many courses (repeated measures). Furthermore, since some students do not graduate, have not graduated yet, switch majors, or take additional courses, students have taken different numbers of courses across each category of classes. All analytic models include a range of demographic covariates that tend to be associated with student achievement and educational attainment to control for their confounding associations. Since course selection is not random, it was important to include these variables in the models to reduce bias from potential selection effects. For example, if more socioeconomically advantaged students had a greater propensity to enroll in a particular statistics course, failure to control for SES when examining the influence of the statistics course on subsequent course grades would generate upwardly biased estimates of the course's effectiveness.

Aim 1. The first aim was to test whether demographics, SAT scores, prior psychology GPA, achievement in statistics, and achievement in research methods predict achievement in subsequent psychology classes. We approached this aim by fitting 12 regression models (Tables 2–4), which predict achievement in different types of classes (Tables 2–4) and with different predictors (the four models within each table). Prior psychology GPA was calculated differently for different models—the rule was that the grades that went into Prior Psychology GPA would include grades in psychology courses up until the previous term in which the student took the methodological course that serves as the predictor in each model.

Tables 2–4. The regressions in Tables 2 predict achievement in all psychology courses with a few exceptions: statistics (taught by the Department of Statistics) and research methods were not included because they are predictors. Additionally, courses such as externships and directed research, which are graded only as satisfactory/not satisfactory, were not included. More details about which classes are included in each of the four models in Table 2 are described below. The regressions in Table 3 predict achievement in advanced seminars. The regressions in Table 4 predict achievement in advanced labs. Students are only required to take one advanced lab course, so the regressions in Table 4 have the fewest observations.

Four regression models within each table. Within each Table 2–4, we present four models that include different predictors to understand the individual and collective predictive value of demographics and SAT, statistics achievement, and research methods achievement. Within each of the four models (columns), all the predictors were added simultaneously, not in steps.

In the first model, only the demographics and SAT scores were included as predictors. The second model tests the predictive value of grades in statistics. In order to maintain temporal separation of grades used as predictors versus grades being predicted, only grades for classes taken after the statistics class were predicted.

Table 2
Models Predicting Grades in All Psychology Courses That Give Letter Grades (Excluding Statistics and Research Methods)

	Demographics & SAT		Statistics		Research Methods		Stats & RM	
Predictors	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*
Female	.17***	.02	.11***	.02	.07**	.02	.06**	.02
Black	−.17***	.04	−.09*	.04	−.05	.04	−.05	.04
Hispanic	.05	.07	.01	.06	−.05	.06	−.04	.06
Asian	−.15***	.04	−.10**	.04	−.07	.04	−.06	.04
Other race	−.09	.06	−.09	.05	−.07	.05	−.07	.05
U.S. citizen	.14	.08	.20**	.07	.16*	.07	.16*	.07
Enrolled by 20	.15***	.04	.08*	.04	−.01	.04	−.02	.04
Number in family	.01	.01	.005	.01	−.00	.01	−.00	.01
AGI/10,000	.00	.00	.01	.00	.01	.00	.01	.00
Parent ed. high school	−.01	.12	.02	.14	.02	.15	.01	.15
Parent ed. college	−.02	.12	−.02	.14	−.01	.15	−.02	.15
High school % minority	−.00	.00	−.00	.00	.00	.00	.00	.00
High school private	−.05	.03	−.01	.03	−.04	.03	−.03	.03
High school suburban	.02	.03	.01	.03	.01	.03	.02	.03
High school rural	.04	.04	.02	.04	.02	.04	.02	.04
SAT verbal/100	.10***	.02	.08***	.02	−.01	.02	−.01	.02
SAT math/100	.06**	.02	−.00	.02	.01	.02	−.01	.02
SAT writing/100	.09***	.02	.06**	.02	.03	.02	.03	.02
PSY credits before Stats			−.01***	.00	−.01***	.00	−.01***	.00
PSY GPA before Stats			.22***	.02				
PSY GPA before RM					.29***	.02	.26***	.02
STAT grade			.30***	.02			.13***	.02
STAT 1000			.12***	.02			.02	.02
STAT 1100			.03	.04			−.00	.04
RM grade					.39***	.02	.36***	.02
Intercept	−1.59***	.18	−1.54***	.19	−.99***	.19	−.83***	.19
Random								
Intercept	−.76	.02	−.91	.02	−1.12	.03	−1.15	.03
Residual	−.47	.00	−.49	.01	−.48	.01	−.48	.01
N	25358		16116		10046		10046	

Note. The "Demographics" model includes all these courses. The "Statistics" model includes the subset of these courses taken after Statistics. The "Research Methods" and "Stats & RM" model predicts the subset taken after Research Methods. Coeff = standardized coefficient.
* $p < .05$. ** $p < .01$. *** $p < .001$.

The Psychology GPA and the number of credits in psychology were entered as covariates to control for each student's general achievement. Again, for temporal separation, the Psychology GPA and number of credits in psychology were calculated for the semester up until when the statistics course was taken. Since students can take one of three statistics courses, they were entered as a series of dummy variables.

The third model tests the predictive value of grades in research methods, separate from statistics. The research methods grade, but not the statistics grade, was used as a predictor, along with the psychology GPA up until taking research methods and the number of credits in psychology before statistics. Only classes taken after research methods were predicted.

In the fourth model, students' grades in research methods and statistics were both entered into the model to estimate the differential predictive value of learning in each course. Psychology GPA up until research methods and the number of credits in psychology up until taking statistics were used as predictors, and only classes taken after research methods were predicted.

Temporal separation of grades being predicted and used as predictors. The models in Table 2 require special attention, because they have a different set of class grades that are being predicted, depending on which predictors are included. In the models in Table 2, all psychology courses that give letter grades (introduction to psychology, core courses, electives, advanced labs, and advanced seminars) are included, except for statistics and research methods, which were not included in any of the models. In the "Demographics" model, all the above-mentioned classes are included. In the "Statistics" model, all the above-mentioned classes are included so long as they were taken after statistics. This decision was intended to keep a temporal separation between grades used as predictors (statistics) and grades used as outcomes (psychology classes taken after statistics). For the "Research Methods" model and the "Stats & Research Methods" model, all the above-mentioned classes were included so long as they were taken after research methods. The covariate Psychology GPA works in a complementary way as the outcome variable. For the "Demographics" model, no Psychology GPA was included because the goal of the model was to predict achievement in psychology courses in college based on demographics and achievement in high school. For the "Statistics" model, Psychology GPA was calculated up until Statistics was taken. For the latter two models, Psychology GPA was calculated up until research methods was taken. These decisions meant that there was a strict separation between grades used in the predictors and grades used as outcomes. For the models in Tables 3 and 4, the covariate of Psychology GPA works the same way, but the outcome variables are simpler because advanced seminars (see table 3) and advanced

Table 3
Models Predicting Grades in All Advanced Psychology Seminars

Predictors	Demographics & SAT		Statistics		Research Methods		Stats & RM	
	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*	Coeff.	*SE*
Female	.15***	.04	.09*	.04	.06	.04	.05	.04
Black	−.15*	.06	−.01	.06	−.05	.06	−.03	.06
Hispanic	.05	.10	−.04	.10	−.06	.09	−.07	.09
Asian	−.09	.07	−.03	.07	.01	.06	.01	.06
Other race	−.17*	.08	−.14	.08	−.20**	.07	−.19**	.07
U.S. citizen	.02	.12	.08	.12	−.04	.12	−.03	.12
Enrolled by 20	.20**	.07	.12	.07	−.01	.07	−.01	.07
Number in family	.003	.01	.00	.01	−.01	.01	−.01	.01
AGI/10,000	.002	.00	.00	.00	.00	.00	.01	.00
Parent ed. high school	−.05	.19	−.03	.22	.08	.25	.06	.25
Parent ed. college	−.06	.19	−.04	.21	.07	.25	.06	.25
high school % minority	−.00	.00	−.00	.00	.00	.00	.00	.00
High school private	−.04	.05	−.05	.05	−.12*	.05	−.11*	.05
High school suburban	.05	.05	.02	.05	−.02	.06	−.02	.06
High school rural	.03	.07	−.02	.07	−.06	.07	−.05	.07
SAT verbal/ 100	.08*	.03	.07*	.03	.01	.03	.01	.03
SAT math/ 100	.02	.03	−.04	.03	−.00	.03	−.02	.03
SAT writing/ 100	.11**	.04	.09*	.04	.03	.04	.03	.04
PSY credits before Stats			−.02***	.00	−.01***	.00	−.01***	.00
PSY GPA before Stats			.18***	.03				
PSY GPA before RM					.29***	.03	.25***	.03
STAT grade			.27***	.03			.11***	.03
STAT 1000			.12***	.04			.01	.04
STAT 1100			.06	.07			.01	.07
RM grade					.38***	.03	.35***	.04
Intercept	−1.27***	.28	−1.19***	.30	−1.04***	.32	−.87**	.32
Random								
Intercept	−.83	.03	−.95	.03	−1.08	.04	−1.09	.04
Residual	−.51	.01	−.52	.01	−.50	.01	−.50	.01
N	9581		6305		4284		4284	

Note. Advanced seminars require Statistics and Research Methods. Coeff. = standardized coefficient.
* $p < .05$. ** $p < .01$. *** $p < .001$.

labs (see Table 4) can only be taken after both Statistics and Research Methods.

SEM. Next, a structural equations model was used to consider the effect of math SAT, achievement in statistics, and achievement in research methods on achievement in advanced seminars (like in the fourth model in Table 3). In addition to providing a visualization, this model also explicitly shows the relations between math SAT, statistics, and research methods, which is not presented in the regressions. The directions of the arrows in the SEM were developed by the temporal ordering of these variables: SATs are taken before statistics, which is taken before research methods, which is taken before advanced seminars. The outcome variable was conditioned on the demographic information, verbal and writing SAT scores, the number of credits before statistics, and GPA in psychology courses before research methods.

Aim 2. In Aim 2 (see Table 5) we attempted to uncover the influence of statistics and research methods on subsequent academic achievement by more clearly establishing temporal precedence in two ways. First, we examined whether grades in psychology core courses were higher if taken after, compared to before, research methods. Second, we examined whether achievement in research methods better predicts achievement in subsequent core courses compared to prior core courses. In other words, we consider how well achievement in research methods predicts grades in core courses taken before research methods versus how achievement in research methods predicts achievement in subsequent core courses. If learning in research methods has an influence on subsequent courses, we would predict the latter correlation to be higher. Mathematically, this involves testing the interaction between the grade in research methods and the indicator variable of whether the course was taken before versus after research methods. Any correlation between achievement in courses taken before research methods and research methods achievement could be due to general factors such as intelligence and grit. Some amount of the correlation between achievement in research methods and achievement in subsequent courses would also be due to these same general confounding factors; however, a higher correlation between research methods achievement and subsequent as opposed to prior achievement during college would attest to the temporal precedence aspect of the influence of research methods.

We focused this analysis on core courses since students can take these before or after research methods. The students in our sample collectively enrolled in core courses 6,487 times before research methods, and 7,009 times after research methods. We fit a hierarchical linear model with a by-student random intercept, and an indicator variable of whether the class was taken before versus after research methods. Since almost all students took some of these courses before and some of them after research methods, this hierarchical model essentially tests a within-student comparison

Table 4
Models Predicting Grades in All Advanced Labs in Psychology

Predictors	Demographics & SAT Coeff.	Demographics & SAT *SE*	Statistics Coeff.	Statistics *SE*	Research Methods Coeff.	Research Methods *SE*	Stats & RM Coeff.	Stats & RM *SE*
Female	.10	.07	.06	.06	.02	.06	.02	.06
Black	−.10	.15	−.07	.13	−.03	.13	−.04	.13
Hispanic	−.10	.21	−.23	.17	−.55**	.19	−.50**	.18
Asian	−.18	.11	−.14	.09	−.18	.09	−.18	.09
Other race	−.18	.18	−.09	.15	.11	.16	.07	.16
U.S. citizen	−.02	.19	−.01	.16	.15	.16	.12	.16
Enrolled by 20	.08	.13	−.01	.11	−.09	.12	−.08	.11
Number in family	.01	.02	.01	.02	−.00	.03	.01	.03
AGI/10,000	.01	.00	.01	.00	.01	.00	.01	.00
High school	.05	.56	.10	.65	.09	.76	.11	.78
College	.09	.56	.09	.66	.03	.76	.03	.78
High school % minority	.00	.00	.00	.00	.01	.00	.01	.00
High school private	−.04	.07	.04	.07	.03	.09	.04	.09
High school suburban	−.01	.10	.04	.10	.11	.11	.10	.11
High school rural	.02	.13	.15	.12	.16	.13	.18	.12
SAT verbal/100	.07	.06	.03	.05	−.07	.05	−.07	.05
SAT math/100	.08	.05	.02	.04	.05	.04	.02	.05
SAT writing/100	.07	.06	.06	.06	.06	.06	.07	.06
PSY credits before stats			−.01*	.00	−.01	.00	−.01	.00
PSY GPA before stats			.21***	.05				
PSY GPA before RM					.30***	.06	.25***	.06
STAT grade			.26***	.05			.14*	.06
STAT 1000			.18**	.06			.02	.06
STAT 1100			.08	.08			.07	.08
RM grade					.35***	.06	.33***	.06
Intercept	−1.31*	.64	−1.36	.71	−1.15	.82	−1.02	.83
Random								
Intercept	−.87	.05	−1.17	.07	−1.27	.09	−1.30	.09
Residual	−.62	.01	−.64	.02	−.60	.02	−.60	.02
N	2792		1854		1226		1226	

Note. Advanced labs require Stats and Research Methods. Coeff. = standardized coefficient.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

that controls for other unobserved subject-level confounds like intelligence, grit, or math anxiety.

This model also includes all the other demographics and SAT covariates. It also includes indicator variables for the six core courses. Even though the course grades were standardized, we included these dummy variables because some of these courses tend to be taken earlier on (especially Developmental) and some later (especially Cognitive Psychology and Biopsychology); this timing is likely a product of both availability of the courses when registering and differential preferences for these courses. Including dummy variables also controls for additional differences in these courses' grades such as differences in skew. These indicator vari-

Table 5
Model 1: Grades in Core Courses Predicted by Whether the Course Was Taken After vs. Before the Predictor Course. Model 2: Grades in Core Courses Predicted by Whether the Course Was Taken After vs. Before the Predictor Course, Grades in the Predictor Course, and the Interaction Between the Two

Tested model	Research Methods Coeff.	Research Methods *SE*	Abnormal Coef.	Abnormal *SE*	Biopsychology Coef.	Biopsychology *SE*	Cognitive Coef.	Cognitive *SE*	Developmental Coef.	Developmental *SE*	Personality Coef.	Personality *SE*	Social Coef.	Social *SE*
	Course used as predictor													
Model 1:														
After course indicator	.03*	.01	.03*	.02	.07***	.02	.04*	.02	.07***	.02	.07***	.02	.07***	.02
Model 2:														
After course indicator	.01	.01	.03	.02	.05**	.02	.04	.02	.08***	.02	.07***	.02	.07***	.02
Course Grade	.48***	.01	.47***	.02	.44***	.02	.43***	.02	.45***	.02	.50***	.02	.52***	.02
Course Grade* After	.13***	.01	.02	.02	.26***	.02	.02	.03	.03	.02	.04	.02	.03	.02
N	11028		7745		7156		5497		8013		7167		7448	

Note. Coeff. = standardized coefficient. These models also included the standard set of covariates as well as indicator variables for the class type.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

ables were not included in Table 5 because they are only meant as covariates to account for additional variance in grade distributions across the core courses, but do not add interpretability to the results. Additionally, we tested achievement in each of the six core courses before versus after each of the other five core courses. The reason for this analysis was to examine whether there is generally a developmental trajectory in which students develop a better understanding of psychology concepts and language about findings and/or greater motivation to learn psychology as they take more classes in psychology. It may be that any core course conveys a learning benefit if taken early rather than research methods in particular.

Results

Aim 1

Results of analyses predicting achievement in psychology courses by demographics, SAT scores, scores in statistics, and scores in research methods are presented in Tables 2, 3, and 4, depending of the type of course being predicted. Here we focus on the main trends in the findings across all the models and types of class.

Model 1: Demographics and SAT. Overall, demographics and SAT did have significant effects and therefore are important to consider when examining effects of course taking. Results in the first column of Tables 2 and 3 indicate that being female is associated with gains of 0.17 *SD* (see Table 2) in all psychology courses and 0.15 *SD* for advanced seminar courses (see Table 3). African American students had lower grades in all psychology courses (−0.17 *SD*) and advanced seminars (−0.15 *SD*). Being enrolled at a university by age 20 is associated with gains of 0.15 *SD* for all psychology courses and 0.20 *SD* for advanced seminars. Achieving higher SAT scores across verbal, math, and writing were associated with higher grades in all psychology courses. SAT verbal and writing, but not math scores, were significant for advanced seminars. None of the SATs were significant for the advanced labs. None of the family and high school characteristics were related to grades, presumably because those demographic variables act through academic achievement during high school (measured by SAT). Table 4 indicates that none of the demographic characteristics were associated with grades in advanced laboratory and methods courses. However, the power for these analyses is much smaller because students are only required to take one advanced lab.

Model 2: Statistics. Results in the second column of Tables 2–4 indicate that grades in statistics are positively related to grades in subsequent courses, even after controlling for all the demographic variables plus students' GPA and the number of credits before the statistics course. An increase of one standard deviation in the statistics grade is associated with an increase of 0.30 *SD* in all psychology courses after statistics, 0.27 *SD* in advanced seminars, and 0.26 *SD* in advanced laboratory courses. Students who took STAT 1000, a more advanced introductory statistics course, obtain better grades (0.12–0.18 *SD*) than students taking the standard STAT 200 course. Additionally, taking statistics later in one's academic career is associated with lower subsequent grades. Each additional psychology credit taken before statistics is associated with a 0.01–0.02 *SD* decrease in grades in subsequent psychology courses. For reference, most courses are 3 credits, so delaying statistics by 3 courses (a typical term load of psychology courses for majors) is associated with roughly a 0.10 to 0.20 *SD* decrease in grades in subsequent courses.

Model 3: Research Methods. Results of the third column in Tables 2–4 indicate that higher research methods grades are associated with higher grades in subsequent courses, even with all the included controls. A one standard deviation higher grade in the research methods course related to 0.39 *SD* higher achievement in all subsequent psychology courses, 0.38 *SD* higher grades in advanced seminars, and 0.35 *SD* higher grades in advanced laboratory and methods courses. Most of the demographics become insignificant in this model, suggesting they may affect later courses through research methods achievement.

Model 4: Statistics and Research Methods. Results of the fourth column in Tables 2–4 indicate that grades obtained in research methods and statistics have differential effects on later grades. Research methods grades are more strongly related to subsequent course achievement than grades in statistics; the effect of statistics is still significant but considerably attenuated compared to Model 2. The association between taking the STAT 1000 course as opposed to STAT 100 and psychology grades also disappears.

SEM. Figure 2 presents the findings from the structural equations model, which predicts achievement in advanced seminars (similar to Table 3). The main points of interest are the considerable relation between achievement in statistics and research methods and the negligible direct influence of math SAT on grades after controlling for achievement in statistics and research methods. Furthermore, the relatively strong relations between statistics and research methods, and between research methods and grades, means that a considerable proportion of the relation between statistics and grades in advanced seminars can be explained by research methods; this effect is also seen in the smaller effect of statistics grade in Model 4 versus 2 in Table 3.

Aim 2

Table 5 presents the analyses for Aim 2; the effects for research methods are in the first column. There is a small effect such that grades in core courses are slightly higher after versus before research methods (0.03 *SD*). However, we note that this same finding holds for all the core courses. For example, grades for the other five core courses are higher after taking personality than

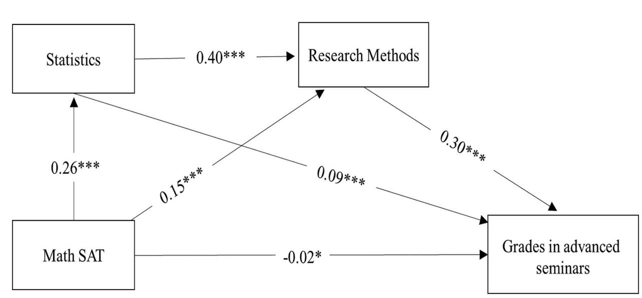

Figure 2. SEM results of Math SAT, Statistics, and Research Methods on Grades in advanced Psychology seminars subsequent to Research Methods.

before (0.07 *SD*). This means that it is a general learning or developmental effect, not something unique to research methods.

Next, we tested whether the predictive value of the achievement in each predictor course (e.g., research methods) is higher for other courses taken after versus before the predictor course. In this model, we added the grade in the predictor course (e.g., research methods) as well as the interaction between this grade and whether another course was taken before versus after the predictor course. There was a significant interaction between achievement in research methods and whether another course was taken afterward versus before (β = .13). Interestingly, this effect was also significant for biopsychology (β = .26) and was even larger than for research methods. A plausible explanation for this finding, in line with Identical Elements Theory, is that biopsychology and neuroscience concepts and methods (e.g., hormones, skin conductance, neuroanatomy, and the biological basis of memory and disorders like autism, dyslexia, anorexia) are relevant to many other areas of psychology, but are usually only explicitly taught in biopsychology, and instructors in other courses may falsely assume that students have already learned these topics. This effect was not significant for any of the other courses, ruling out a general exposure to psychology content as the basis of benefits from research methods.

In sum, it is not merely *taking* research methods or biopsychology that is associated with uniquely higher achievement in subsequent courses. In fact, it appears that merely taking any course is associated with higher achievement in future courses. However, *success* in research methods and biopsychology is the key predictor of success in future courses compared to past courses.

Discussion

We tested the role of research methods and statistics on subsequent psychology coursework. The main findings were that achievement in research methods and statistics both predicted future achievement in other psychology classes, even after controlling for demographics, SAT scores, and prior psychology GPA. Research methods achievement accounted for larger variance than statistics achievement (0.30 *SD* in comparison with 0.09 *SD*; see Figure 2). Delaying statistics, which also results in a delay for research methods, is associated with worse subsequent achievement. Finally, although achievement in core courses tended to be better after taking all other core courses and research methods, *success* in research methods (and biopsychology) is uniquely predictive of success in future courses compared to past courses.

These findings are consistent with Identity Element Theories of transfer and largely agree with prior research, highlighting the important role for statistics and research methods (Freng et al., 2011), and more generally understanding the nature of science and science practices (Lederman et al., 2002; National Research Council, 2012). Students may delay enrolling in methodological classes because of anxiety, which is associated with lower grades (Onwuegbuzie & Wilson, 2003). The current study suggests that this strategy is likely to be counterproductive.

Unexpected Outcomes

There were two unexpected outcomes. First, Asian students tended to perform worse than White students (American Psychological Association Presidential Task Force on Educational Disparities, 2012; Hsin & Xie, 2014), which may reflect English-as-a-second-language issues that could be especially problematic in language-rich psychology courses. Second, success in biopsychology was especially predictive of success in subsequent courses. A plausible explanation for this finding is that the knowledge learned in Biopsychology (e.g., brain regions and neuroscience methods) is becoming increasingly relevant to other areas of study and may be hard to understand without previously taking a course in biopsychology. Within the Identical Elements Framework, this may relate to skills of interpreting neuroscientific studies or it may relate to knowledge of foundational concepts (e.g., brain regions). Thus, it may be beneficial to encourage students to take biopsychology (or neuroscience) early on, and also to implement measures to help ensure success in biopsychology.

Learning About Students Versus Classes

These findings provide important information about both psychology students and classes. For example, consider whether it is desirable to have a strong correlation between achievement in statistics and advanced courses in psychology. On the one hand, a high correlation suggests that these advanced courses build upon and rely upon the content taught in statistics. We presume that many instructors in psychology intend for their courses to build upon students' prior quantitative skills, and low degrees of correlation with statistics or research methods could be a sign that a course does not require students to deeply engage with the research methodologies of studies being discussed in the course. On the other hand, the better job that instructors do in scaffolding student learning to meet the needs of students with diverse backgrounds and levels of prior achievement, the less of a correlation there would be between achievement in statistics and research methods and subsequent achievement. For these reasons, we do not believe that there is an optimal degree of correlation. Considering these sorts of issues may help departments better understand the degree of course overlap and the degree to which they want to require skills learned in one class in subsequent classes.

A similar logic can be used for the findings about biopsychology. One interpretation is that encouraging students to take biopsychology earlier may facilitate learning in other classes. Another interpretation is that given that the other classes are core classes that do not require biopsychology, those classes might be expecting students to have too strong of a background in biopsychology, and it may be wise to incorporate more scaffolding to help students understand the biological and neuroscience aspects that are relevant for those classes.

Limitations

There are four main limitations. The first limitation is that since the study was observational, not a randomized controlled trial, there may be alternative explanations for some of the findings. In particular, intelligence or motivation (especially interest, self-efficacy, and grit) presumably contribute to achievement in all courses. Furthermore, it is possible that students higher on these characteristics might choose to take methods courses earlier. In this way, individual differences could potentially explain correlations between when a student takes methods courses and how well they perform in other courses. The findings that SAT and achievement in statistics become considerably weaker once controlling for achievement in research methods is harder to explain with motivation third variables like interest and self-efficacy.

It is theoretically possible that performance in research methods (or statistics or biopsychology) has a large impact on these factors, which could then impact grades in subsequent classes (e.g., performing well vs. poorly in research methods could increase vs. decrease a student's self-efficacy, which could impact performance in subsequent classes). Future research will have to directly examine these explanations. However, it is important to note that difficult courses (like research methods, statistics, and biopsychology) generally tend to decrease self-efficacy and interest (Osborne, Simon, & Collins, 2003; Urdan & Schoenfelder, 2006), and mean interest levels in psychology can decline with additional coursework (Harackiewicz, Barron, Tauer, Carter, & Elliot, 2000). In any case, since it is possible that motivation variables such as interest and self-efficacy change during research methods or biopsychology as students are exposed to the nitty-gritty details of psychology and students reevaluate their interest level, this possibility would still suggest that research methods and biopsychology are important, but because they influence students' motivation and interest. If this is true, it would be important to ensure that students' interest in and motivation to study psychology are maintained through these challenging courses. In summary, causal inference about the underlying causes of one course on another is challenging because of the possibility of alternative explanations, and it is likely that there are multiple factors including transfer and motivation at work.

The second limitation is that we were only able to conduct the analyses for Aim 2 on core courses. The reason is that students must take research methods before advanced courses; only in exceptional cases are students allowed to take advanced courses before research methods. It is most plausible that the knowledge and skills learned in research methods and statistics will be most relevant in advanced courses that involve critiquing articles and conducting lab-based research. For this reason, it is not surprising that the effects for research methods are quite small on the core courses in Aim 2. Additionally, for Aim 1, although it would be interesting to compare the predictiveness of research methods on advanced seminars and labs (which presumably build more upon research methods) versus core courses, core courses are much larger and often use different grading practices. Advanced courses often have high and narrow grade distributions, so we did not think that it makes sense to directly compare them.

The third limitation is that it would be ideal to have a better understanding of the true overlap between the knowledge and skills taught in different courses, which would complement our empirical analyses. Although this might be possible for a small number of courses, given the large number of courses at our university, the fact that many courses are taught by multiple instructors, and that course content changes over time, this sort of analysis is not possible for the scale of courses examined in the current research.

The fourth limitation is that this research was conducted at a single university. In the introduction, we argued that the course structure at the University of Pittsburgh is fairly typical of psychology majors in the U.S. However, there is also wide variance in the structure of majors and the courses offered at different universities, and courses with similar titles may be taught very differently across institutions. On the other hand, even within our university, most courses are taught by multiple faculty, who do not necessarily teach the same way, so perhaps some of the variance at other universities may be captured by variance within our university.

Summary

The current research attests to the central role of research methods in the psychology major. Furthermore, it suggests that even though demographic factors and even math SAT scores predict achievement in the major, most of those factors tend not to be significant above and beyond achievement in research methods (i.e., those factors may enable the acquisition of research methods skills, but it is the research methods skills that enable later learning in psychology courses). Lastly, the current research provides a framework for how educational data can be used to test whether the sequencing of courses in a science major matters, and this approach may be useful for science departments outside of psychology.

References

American Psychological Association. (2013). *APA guidelines for the undergraduate psychology major* (Version 2.0). Washington, DC: Author. Retrieved from http://www.apa.org/ed/precollege/about/psymajor-guidelines.pdf

American Psychological Association Presidential Task Force on Educational Disparities. (2012). *Ethnic and racial disparities in education: Psychology's contributions to understanding and reducing disparities*. Retrieved from http://www.apa.org/ed/resources/racial-disparities.aspx

Barron, K. E., & Apple, K. J. (2014). Debating curricular strategies for teaching statistics and research methods: What does the current evidence suggest? *Teaching of Psychology, 41,* 187–194. http://dx.doi.org/10.1177/0098628314537967

Billings, S. B., Deming, D. J., & Rockoff, J. (2014). School segregation, educational attainment, and crime: Evidence from the end of bussing in Charlotte–Mecklenburg. *The Quarterly Journal of Economics, 129,* 435–476. http://dx.doi.org/10.1093/qje/qjt026

Carbonaro, W., & Covay, E. (2010). School sector and student achievement in the era of standards based reforms. *Sociology of Education, 83,* 160–182. http://dx.doi.org/10.1177/0038040710367934

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dunn, D. S., Brewer, C. L., Cautin, R. L., Gurung, R. A. R., Keith, K. D., McGregor, L. N., . . . Voigt, M. J. (2010). The undergraduate psychology curriculum: Call for a core. In D. F. Halpern (Ed.), *Undergraduate education in psychology: A blueprint for the future of the discipline* (pp. 47–61). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/12063-003

Fortus, D., & Krajcik, J. (2012). Curriculum coherence and learning progressions. In Barry J. Fraser, Kenneth Tobin, & Campbell J. McRobbie (Eds.), *Second international handbook of science education* (pp. 783–798). Heidelberg, Germany: Springer Netherlands. http://dx.doi.org/10.1007/978-1-4020-9041-7_52

Freng, S., Webber, D., Blatter, J., Wing, A., & Scott, W. D. (2011). The role of statistics and research methods in the academic success of psychology majors: Do performance and enrollment timing matter? *Teaching of Psychology, 38,* 83–88. http://dx.doi.org/10.1177/0098628311401591

Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. Oxford, England: Holt, Rinehart & Winston.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92,* 316–330. http://dx.doi.org/10.1037/0022-0663.92.2.316

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41,* 111–127. http://dx.doi.org/10.1207/s15326985ep4102_4

Hofstein, A., & Mamlok-Naaman, R. (2007). The laboratory in science education: The state of the art. *Chemistry Education Research and Practice, 8,* 105–107. http://dx.doi.org/10.1039/B7RP90003A

Hsin, A., & Xie, Y. (2014). Explaining Asian Americans' academic advantage over Whites. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 8416–8421. http://dx.doi.org/10.1073/pnas.1406402111

Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology, 20,* 362–404. http://dx.doi.org/10.1016/0010-0285(88)90004-7

Lawson, T. J. (1999). Assessing psychological critical thinking as a learning outcome for psychology majors. *Teaching of Psychology, 26,* 207–209.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29,* 331–359. http://dx.doi.org/10.1002/tea.3660290404

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39,* 497–521. http://dx.doi.org/10.1002/tea.10034

Matz, R. L., Rothman, E. D., Krajcik, J. S., & Banaszak Holl, M. M. (2012). Concurrent enrollment in lecture and laboratory enhances student performance and retention. *Journal of Research in Science Teaching, 49,* 659–682. http://dx.doi.org/10.1002/tea.21016

McLean, C. P., & Miller, N. A. (2010). Changes in critical thinking skills following a course on science and pseudoscience: A quasi experimental study. *Teaching of Psychology, 37,* 85–90. http://dx.doi.org/10.1080/00986281003626714

Mill, D., Gray, T., & Mandel, D. R. (1994). Influence of research methods and statistics courses on everyday reasoning, critical abilities, and belief in unsubstantiated phenomena. *Canadian Journal of Behavioural Science, 26,* 246–258. http://dx.doi.org/10.1037/0008-400X.26.2.246

National Center for Education and Statistics. (2014). *Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2001 to 2014*. U.S. Department of Education. Retrieved from https://nces.ed.gov/ccd/pubschuniv.asp

National Research Council. (2006). *America's lab report: Investigations in high school science*. Washington, DC: National Academies Press.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

Nokes-Malach, T. J., & Mestre, J. P. (2013). Toward a model of transfer as sense-making. *Educational Psychologist, 48,* 184–207. http://dx.doi.org/10.1080/00461520.2013.807556

Norcross, J. C., Hailstorks, R., Aiken, L. S., Pfund, R. A., Stamm, K. E., & Christidis, P. (2016). Undergraduate study in psychology: Curriculum and assessment. *American Psychologist, 71,* 89–101. http://dx.doi.org/10.1037/a0040095

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments—A comprehensive review of the literature. *Teaching in Higher Education, 8,* 195–209. http://dx.doi.org/10.1080/1356251032000052447

Osborne, J. A., Simon, S. B., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25,* 1049–1079. http://dx.doi.org/10.1080/0950069032000032199

Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal, 50,* 714–754. http://dx.doi.org/10.3102/0002831213481240

Royston, P. (2004). Multiple imputation of missing values. *The Stata Journal, 4,* 227–241.

Royston, P. (2005). Multiple imputation of missing values: Update of ice. *The Stata Journal, 5,* 527–536.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734–760. http://dx.doi.org/10.1037/0033-295X.103.4.734

Sadler, P. M., & Tai, R. H. (2007). Transitions. The two high-school pillars supporting college science. *Science, 317,* 457–458. http://dx.doi.org/10.1126/science.1144214

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of U.S. mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37,* 525–559. http://dx.doi.org/10.1080/0022027042000294682

Schraagen, J. M. (1993). How experts solve a novel problem in experimental design. *Cognitive Science, 17,* 285–309. http://dx.doi.org/10.1207/s15516709cog1702_4

Schunn, C. D., & Anderson, J. R. (1998). Scientific discovery. In J. R. Anderson & C. Lebière (Eds.), *Atomic components of thought* (pp. 385–427). Mahwah, NJ: Erlbaum.

Schunn, C. D., & Anderson, J. R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science, 23,* 337–370. http://dx.doi.org/10.1207/s15516709cog2303_3

Schunn, C. D., & Anderson, J. R. (2001). Science education in universities: Explorations of what, when, and how. In K. Crowley, C. D. Schunn, & T. Okada (Eds.), *Designing for science: Implications from professional, instructional, and everyday science* (pp. 73–102). Mahwah, NJ: Erlbaum.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Stoloff, M., McCarthy, M., Keller, L., Varfolomeeva, V., Lynch, J., Makara, K., . . . Smiley, W. (2010). The undergraduate psychology major: An examination of structure and sequence. *Teaching of Psychology, 37,* 4–15. http://dx.doi.org/10.1080/00986280903426274

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44,* 331–349. http://dx.doi.org/10.1016/j.jsp.2006.04.003

VanderStoep, S. W., & Shaughnessy, J. J. (1997). Taking a course in research methods improves reasoning about real-life events. *Teaching of Psychology, 24,* 122–124. http://dx.doi.org/10.1207/s15328023top2402_8

Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology, 30,* 220–223. http://dx.doi.org/10.1207/S15328023TOP3003_04

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8,* 247–261. http://dx.doi.org/10.1037/h0074898

Received June 13, 2017
Revision received November 29, 2017
Accepted January 8, 2018 ■