THE EFFECTIVENESS OF PRE-COURSE AND CONCURRENT COURSE
INTERVENTIONS ON AT-RISK COLLEGE PHYSICS STUDENTS' MECHANICS
PERFORMANCE

BY

SARA JANINE ROSE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Secondary and Continuing Education
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:

       Professor Fouad Abd-El-Khalick, Chair
       Professor Gary Gladding
       Professor Sarah Lubienski
       Professor José Mestre

ABSTRACT

Students at risk for poor performance or withdrawal in post-secondary education, and particularly in scientific fields, are the focus of educators interested in improving retention and persistence rates in college science studies, as well as equity and diversity of their institutions' graduates and the overall workforce. This study aimed to assess the effectiveness of two intervention courses—one prior and one concurrent to the course of interest—on at-risk students' performance in an introductory college physics course. Participants were engineering freshman with prerequisite calculus credit enrolled in an introductory mechanics course—the target course—at a large Midwestern university. Students at different levels of risk were identified by logistic regression analysis on collected measures of prior education, national exam scores, university diagnostic scores, as well as demographic and socio-economic information. The study had a quasi-experimental, posttest only, non-equivalent control group design, which utilized propensity score matching to assess the differential impacts of the two approaches on at-risk student performance and persistence in the target course. Data analysis gauged the size and nature of the interventions' impact on participants' performance. By controlling for additional factors, analyses allowed for making generalizations related to the characteristics of students at risk for poor or failing performance in, or withdrawal from, college-level physics. Analyses indicated that students from both interventions performed better on the target course's assessments. The students who participated in the concurrent course instruction, which is focused on metacognitive skill development, saw twice the performance gain than the Pre-Course students when compared to their peers who received no intervention.

To my Mom & Dad

and the paths that wander a long way from the farm

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

CHAPTER 1

THE PROBLEM

Introduction

Interventions aimed at improving at-risk student achievement and retention in introductory college physics courses date back about forty years (Foster, 1972). Research into the characteristics and predictors of at-risk students has informed curricular and pedagogical choices aimed at devising effective interventions for improved performance, retention, and persistence in college science courses. Early at-risk student predictor studies showed that students who failed or withdrew from introductory college physics typically lacked adequate mathematical and algebraic skills, abstract reasoning skills, scientific reasoning skills, and/or basic scientific and physics knowledge (Adams & Garrett, 1954; Barnes, 1977; Champagne & Klopfer, 1982; Champagne, Klopfer, & Anderson, 1980; Cohen, Hillman, & Agne, 1978; Griffith, 1985; Hudson, 1986; Hudson & Liberman, 1982; Hudson & McIntire, 1977; Hudson & Rottmann, 1981; Liberman & Hudson, 1979; McCammon, Golden, & Wuensch, 1988). Basing their strategy on what was known about the characteristics of students at risk for low or failing performance in physics courses, early interventions generally consisted of a remedial course taken before the target physics course. Remedial curricular materials addressed inadequate skills, abilities, and knowledge among at-risk students (i.e., McDermott, Piternick, & Rosenquist, 1980a, 1980b, 1980c; Taylor, 1983). Content was the primary focus of these interventions, and the particulars of the instructional techniques were for the most part not evaluated.

Whether students completing these types of remedial interventions performed at or above par relative to their control-group peers remains unknown (i.e., McDermott et al., 1980c). At the same time, at-risk predictor research showed that students possessing adequate levels of the

identified knowledge and skills were equally as likely to fail or withdraw (Champagne &
Klopfer, 1982; Champagne et al., 1980; Griffith, 1985; Hudson, 1986). Researchers were led to
conclude that the distribution of knowledge and skills are similar across students who do and do
not fail, rendering the examined variables unreliable predictors of being at-risk. It seemed that
other or underlying variables contribute to risk, and the study of at-risk students evolved to focus
on more complex qualities that affect learning and performance. Analyses have shown that
student backgrounds and attributes, ranging from gender to high school and exam performance,
and from high school teaching and learning experiences to family attitudes toward science
courses contributed to their college physics performance (Champagne et al., 1980; Hazari, Tai, &
Sadler, 2007; Sadler & Tai, 1997, 2001). As summarized to this point, the initial association of
risk with deficiencies in scholastic aptitude and cognitive abilities transitioned to
acknowledgement that cultural and societal factors encompassing educational advantages and
resources directly affect the risk of relative poor performance in college physics. These factors
more often than not indicate that influences on students' performance are not due to low
cognitive ability or motivation but rather prior learning experiences or a comparative lack
therein. Overall, influences on student performance and characteristics of risk are thus
understood to be complex and diverse.

　　More recent work has focused on pedagogical strategies and considerations
characterizing a target physics course (Brahmia & Etkina, 2001; Cottle & Hart, 1996; Drane,
Smith, Light, Pinto, & Swarat; 2005; Etkina, Gibbons, Holton, & Horton, 1999; Holton &
Horton, 1996). In particular, cooperative learning methods, which were already being used in
college physics (e.g., Heller, Keith, & Anderson, 1992; Heller & Hollabaugh, 1992; Mazur,
1997), have been shown to have positive effects in other content domains (Fullilove & Treisman,

1990; Kagan, 1989). Cooperative learning has been used in hands-on workshops, problem-solving recitations, and discovery-based laboratories (Cottle & Hart, 1996; Etkina et al., 1999). When intentionally situated around existing curricula, cooperative learning can serve to scaffold students' learning processes, as well as contribute to a considerable difference in the number of contact hours per week for students.

Most interventions undertaken to address the needs of the at-risk students described here were implemented within physics departments (Brahmia & Etkina, 2001; Cottle & Hart, 1996; Etkina, et al., 1999; Holton & Horton, 1996; McDermott et al., 1980a, 1980b, 1980c; Taylor, 1983), with one notable but unsuccessful approach initiated by nonphysicists with support from the physics department (Drane et al., 2005). The content of most interventions was driven by helping students develop adequate skills and knowledge for improved performance in the target course. Extra time spent with students, especially in the context of collaborative group problem solving, whether in a pre- (or remedial) course or within the target course, was an effective strategy to increase the performance and retention of at-risk introductory college physics students (Etkina et al., 1999). A useful way to frame the characteristics that affect an intervention's effectiveness is to outline the initiating unit, curricular focus, pedagogical approach, contact hours per week or time spent studying material, and/or the relation to and integration into the target physics course.

Framework and Rationale

The curricula for remedial intervention courses, which enroll at-risk students prior to taking the target course, typically revolve around imparting mathematics skills, abstract reasoning skills, and knowledge about basic physics concepts to the at-risk students (e.g., McDermott et al., 1980a, 1980b, 1980c; Taylor, 1983). Pedagogical strategies centered on

utilizing cooperative learning and grouping to solve physics problems originated in Treisman's study of at-risk students at UC-Berkeley (Fullilove & Treisman, 1990; Treisman, 1985, 1992) and were further supported by studies on human learning and education (Bransford, Brown, & Cocking, 1999). Constructivism informed by cognitive psychology has greatly influenced instructional approaches in science education as a whole and in physics education specifically (Bransford, et al., 1999). For instance, White and Frederickson (1998) developed an interactive software tool that allows middle school physics students to keep track of the cognitive processing they engage in as they perform physics experiments. Such constructivist and cooperative perspectives most often are not explicitly referenced in more recent research pertaining to intervention development, yet the influence of these initial curricular and instructional perspectives is apparent. Etkina collaborated with several colleagues to create a course, whose units are cyclic, aimed at providing the environment, time, and support required to promote significant scientific conceptual change and deep understanding (Brahmia & Etkina, 2001; Etkina et al., 1999; Holton & Horton, 1996). Recent at-risk student intervention strategies in general college physics education draw heavily on what is known about best teaching practices in the sciences as informed by cognitive psychology, educational psychology, and curricular and pedagogical theory.

Effectiveness has been measured in terms of broad qualitative observations of whether or not the achievement of the objectives of interventions was exhibited (McDermott et al., 1980a, 1980b, 1980c; Taylor 1983) and by the impact on students' learning behaviors and perceptions of performance as reported on self-report surveys (Cottle & Hart, 1996). Retention rates in a program, major, or college level of students completing the interventions have also been used as an indicator of effectiveness for several interventions, with reports on just a few students'

4

subsequent success beyond a target course (McDermott et al., 1980a, 1980b, 1980c), as well as on rates of students persisting as engineering or science majors over a number of years (Brahmia & Etkina, 2001; Holton & Horton, 1996). Of the studies reviewed, only *one* had reported quantitative comparisons of intervention versus control groups on certain course performance measures, such as common final exam questions and average course grade (Etkina et al., 1999). The lack of quantitative studies that are comparative often is related to the difficulty of securing appropriate control groups.

Statement of the Problem

The review of previous research in this domain reveals lack of clarity or solid grounding of the theoretical framework(s) that were purportedly used to guide the design of curricula or choice of pedagogical approach in at-risk college physics student interventions. The foundation of the choices often made in terms of content and instructional methods of interventions has not been explicitly and solidly supported by theories in the fields of educational psychology, cognitive psychology, or the specialized field of at-risk student instruction. Deliberately designed instruction founded in what is currently known regarding the brain, cognition, and theories on how humans learn allow systematic practical application of learning and teaching theories. Much of what is known about how the brain functions is understood within classic cognitive psychology; education, however, is an application of learning and teaching theories with a number of mediating variables, which cannot be simulated and further understood in a clinical situation. Justification of the content and instructional methods of interventions is therefore important (as opposed to the mere study of their effectiveness) with regards to the theories that guide and frame such interventions. To date, there have not been systematic manipulations of the characteristics of interventions found to significantly affect at-risk introductory college physics

student performance. To be sure, one way in which interventions have not been fully explored relates to the type of integration with the target course. Specifically, in addition to the general lack of quantitative examinations of the effectiveness of interventions, a more systematic study of intervention integration, such as Pre-Course versus within-course interventions, has not been conducted. Research sufficient to evaluate the effectiveness and make comparisons across different types of implementations and interventions has not been attempted.

Purpose and Research Questions

The preceding review indicates the need for more thorough understanding of the effectiveness and contributing factors of interventions, as well as links to major theoretical factors underlying these interventions. Such is the purpose of the present study, which was undertaken in the context of two intervention approaches. The first was implemented as a Pre-Course to the target one. The second intervention, undertaken in the same disciplinary department, took the form of an additional intervention course offered to students concurrent to their enrollment in the target course after having completed the Pre-Course intervention. Thus, the study provides a means to assess the impacts of the two approaches. Two specific aspects of the effectiveness of the intervention were investigated: (a) at-risk student performance in the target course, and (b) student persistence in the target course. The questions that guided the present study were:

(1) What are the differential impacts on at-risk student performance for Pre-Course intervention, Pre-Course plus concurrent intervention, and no intervention?

(2) Are there differences in persistence to complete the target course between intervention and non-intervention students?

A thorough analysis was undertaken to delineate the objectives and design of the Pre-

Course and the concurrent supplemental course. This analysis aimed to shed light on the epistemological and theoretical underpinnings guiding decisions made in regard to the content and delivery of the two interventions. Outlining the curricular objectives and purposeful choices of instructional strategies for the intervention courses helped set the context for understanding the impact of each intervention in supporting the observed performance of at-risk students.

Significance

The effectiveness of intervention such as those examined in the present study provides the feedback needed to evaluate whether the hypothesized strategies to improve at-risk students' performance and retention indeed help students. Through answering the research questions, a better picture of the effectiveness and factors contributing to the success of at-risk physics students in introductory physics courses was generated. Also generated by the research will be a complete picture of the instructional approach in, and the associated effectiveness of, the supplemental courses offered to at-risk students. Strengths and shortcomings were identified in comparing the two approaches to current learning and teaching theory, which has implications for the design and implementation of current and future intervention aimed at achieving similar gains in student performance.

CHAPTER 2

LITERATURE REVIEW

Curricular and pedagogical approaches devised to impact the performance of students at risk for failure or withdrawal in introductory physics courses were originally intended to ameliorate the observed 'deficiencies' those students exhibited compared to their peers. This chapter first outlines the notable approaches with published results in college physics, as well as in other scientific disciplines at the college level. Conclusions are made about the distinguishing features of the effective approaches. Various issues related to equity and diversity are discussed, including an at-length discussion of Triesman's (1985) findings in his doctoral work and subsequent instructional approaches utilizing the observed behaviors of successful students. Second, a theoretical framework is developed for a conceptualization of curriculum and pedagogy grounded strongly in cognitive psychology and learning theory. The structure provided by this conceptualization serves as the method of analysis for the curriculum and pedagogical approach of the intervention courses under study.

Solutions Implemented in College Physics

Early research in the physics education research community was concerned about the poor passing and retention rates of at-risk students in introductory physics (Foster, 1972; Kaufman, 1977). Research and attention to general enrollment trends have shown that at-risk students tended to be racial or ethnic minority students (Etkina et al., 1999; Hazari, Tai, & Sadler, 2007; McDermott, Piternick, and Rosenquist, 1980a, 1980b, 1980c; Sadler & Tai, 1997, 2001) who entered college lacking requisite cognitive and content skills (Adams & Garrett, 1954; Barnes, 1977; Champagne & Klopfer, 1982; Champagne, Klopfer, & Anderson, 1980; McCammon, Golden, & Wuensch, 1988; Griffith, 1985; Hart & Cottle, 1993; Hazari et al., 2007;

Hudson, 1986; Hudson & McIntire, 1977; Hudson & Rottmann, 1981; Hudson & Liberman; 1982; Liberman & Hudson, 1979; Cohen, Hillman, & Agne, 1978; Wollman & Lawrenz, 1984) and were underprepared in regard to prior knowledge (Sadler & Tai, 1997, 2001). The conclusions of these studies highlight the deficit perspective carried through predictor research and further into interventions to address concerns for at-risk student performance. This perspective attributes poor performance in school to a deficiency with the learners rather than with administrative and organizational structures, or instructional practices arising from within the educational institution (Valencia, 1997). It also underemphasizes identifying and capitalizing on the scholastic and cognitive strengths diverse students pursuing post-secondary education possess. Particular concern arose because introductory physics courses often serve as gateway courses to continuing with the curricular course sequences for engineering majors. As such, failure to perform adequately in these courses had serious consequences for subsequent curricular and career avenues. Solutions to these concerns involved intervening instructional strategies designed specifically for at-risk students. Due to the strong focus on particular subgroup's association to being at high risk for failure or attrition, interventions have typically been developed and directed at minority students or women.

Based on early at-risk student characteristic research indicating a difference in measures of cognitive reasoning, content knowledge and skills aligned with and valued by the pre-existing traditional curriculum and pedagogy, focus from within physics departments was then aimed at adequately preparing these students using interventions targeting these specific dimensions. For instance, McDermott et al. (1980a, 1980b, 1980c) detailed reformed curricular objectives to prepare their students for future physics courses and maintained a focus on the content and skills shown indicative for being at risk. McDermott et al. (1980a) pursued student mastery for

9

understanding particular basic physics concepts, as well as mastery of skills such as differentiation of concepts, connecting reality to abstract representations, applying reasoning, and transferring concepts from one context to another. Their concern for fostering thought processes was manifested through working with students at a slower pace through the material. They explain, "Coverage of subject matter is reduced to allow the students time to acquire a thorough understanding of basic concepts and to become engaged in reasoning that is associated with a particular subject" (McDermott et al., 1980a, p. 139). Strategies for instruction include emphasis on laboratory work, stress on reasoning, integrated role of examinations and homework, and increase in challenge as the course progresses (McDermott et al., 1980b).

In the case of McDermott and her colleagues (1980b), the implementation of the course occurred entirely in laboratories to provide direct, hands-on experience for students whose background had not led to more fully developed reasoning skills. Students experienced physics in real-life situations, talked about their experiences and conceptions, and were encouraged to think more deeply and critically about what goes on in the world around them. The carefully designed list of underlying teaching and learning objectives addressed specific student difficulties. These included, for example, confusing two concepts that apply to the same situation (for example, conceptually distinguishing between "same place or position" and "same speed"), lack of connection between reality and abstract representations, problems with scientific reasoning, and inability to reason by analogy and to transfer understandings to a new context. Students who completed the entire physics preparatory course left with a fair understanding of the most basic physical concepts covered (mass, volume, density, and uniform velocity) and grasping more subtle concepts (elements and compounds, heat, instantaneous velocity). The authors reported that the class progressed to nearly everyone mastering the underlying target

10

skills, such as proportional reasoning, translation of a two-step story problem into algebraic equations, and interpretation of curved graphs and changing slopes. Performance from pre- to posttest on a reasoning test free of physics content indicated an increase of proportional reasoning skills transferred to new situations. The authors claim success because a number of students who completed the physics course were still in tracks leading to engineering or medical school, while others had been accepted into health science programs. No further intervention evaluation was reported.

Taylor (1983) adapted the modular materials for college freshman from McDermott et al.'s (1980a, 1980b, 1980c) curriculum with an additional focus on the development of abstract and scientific reasoning skills independent of physics content identified in at-risk student characteristics literature. The main motivation for the preparatory course was to address what were considered "deficiencies" in student's prior learning by science faculty, including " a lack of academic experience as a foundation on which the abstractions of science can be built, a weakness in mathematical and verbal skills, students' lack of confidence in their ability to solve problems, and low personal standards for academic achievement" (Taylor, 1983, p. 9). The course design was motivated by studies that indicated reasoning skill development—proportional reasoning, hypothesis building and testing, control of variables, and logical implication—were crucial for success in scientific careers, as well as science faculty experience with the challenges at-risk students faced. Since students often did not distinguish between repeating a definition and actually understanding a concept, laboratory work was utilized to give them direct experience with the concepts prior to the introduction of technical terms, generalizations, and abstractions of the core concept. The course was taught using a Personalized Self-Instruction (PSI) system, wherein students work sequentially through the materials and pace themselves against fixed

examination dates. During the four two-hour sessions per week intervention, students independently read text, performed simple experiments, and worked on exercises, all with one-on-one help available. The incorporated modules adapted from McDermott et al. (1980a, 1980b, 1980c) intentionally created cognitive dissonance for students who had confused two related but fundamentally different concepts (such as same position versus same speed). Analogical reasoning was another ability developed by the course through exposure to specific examples that allow transfer of similar procedures, mechanisms, or concepts to other situations.

Throughout curricular and pedagogical approaches, Taylor (1983) utilized the environment and materials to provide students with self-reflective opportunities and scaffolded the development of their thinking and reasoning skills. Practice with graphs, diagrams, equations, and verbal statements in different situations encouraged productive reasoning around these scientific tools. Homework focusing on the reasoning of a solution to problems prepared students for assessments that "emphasize the correct use and interpretation of data rather than memorized facts through questions that normally demand an explanation of the reasoning used to obtain an answer" (p. 10). Group discussions were occasionally used to summarize or practice a particularly difficult concept. Furthermore, extra tutoring sessions helped students to reason through homework problems, and in-class help was provided to students working collaboratively on physics problems in the form of Socratic dialogue. The latter approach provided students with cognitive scaffolding and modeling through the instructors' use of questions, which encouraged them to reason through their questions independently. As students experienced and mastered this skill for themselves, less help was offered, and most students developed the skill and confidence to solve problems on their own. The author recounted that the students went on to adopt the questioning technique modeled by tutors to scaffold reasoning while helping other students.

Quantitative evaluation of this pre-physics course's effectiveness in developing students' reasoning and preparing them sufficiently for the introductory physics course was not conducted. Most students showed improvement on a diagnostic quiz administered pre- and post-course by increasing their skill to solve simple word problems and algebraic equations. Taylor (1983) also reported anecdotally on the subsequent performance of a few students. The author observed that the students had a greater problem-solving ability and were more willing to independently attempt a solution. Standard end-of-semester course evaluations indicated improved affective outcomes related to reasoning through problems and confidence in learning difficult material. Additionally, the ability to use, interpret, and create abstract representations of reality were fostered through repeated use to understand the results or implications of a real life experiment or experience conducted in the laboratory.

In interventions like those of McDermott et al. (1980a, 1980b, 1980c) and Taylor (1983), the focus of the intervention both in regard to pedagogy (how to impart material to the students) and curriculum (what to impart) did not seem to be grounded in cognitive and educational research and theories on learning and teaching. As more about learning and teaching theories and research were incorporated into instructional practice, change in science education, including physics education, was often directed away from teacher-centered toward student-centered instruction (see Hake, 1998, for examples). More often than not, students identified at high risk for failure or attrition were perceived to need additional instruction on cognitive skills and remedial material to prepare them to keep up with peers in the unmodified target course. Interventions, thus, evolved to focus on more than curricular materials targeting difficult content areas and promoting mastery of requisite prior knowledge and science reasoning skills perceived as underdeveloped in at risk students. Pedagogy capitalizing on the advantages of collaborative

and group learning activities—highly influenced by Treisman's (1985, 1992) and Fullilove and

Treisman's (1990) work (expanded upon later)—opened a new direction for supplemental

physics courses directed toward at-risk student performance and retention issues.

In the mid 1990s, Cottle and Hart (1996) published a study about a departmental staff's

implementation of cooperative group learning in the recitation sections of a preexisting large

physics course at Florida State University. The authors incorporated research findings regarding

cooperative learning studies (Burron, James, & Ambrosio, 1993; Heller, Keith, & Anderson,

1992; Heller & Hollabaugh, 1992; Kagan, 1989). Recitation time and the tutorials were existent

in the first semester, but emphasis was placed on collaborative learning in groups of 4-6 students

with the tutorials in the second semester. The authors studied the differential effects of these

pedagogical techniques by introducing in-class cooperative learning materials and comparing the

students' perception and performance to that of less interactive recitation sections. Passing final

course grades for a handful of students in the small groups led the authors to conclude that the

students' participation in cooperative learning activities resulted in the observed achievement

levels of those students. Investigation into the survey results of the whole course indicated that

students experienced a greater commitment to learning physics (probably due to the increased

rate of attendance and active participation), greater involvement with other students, and finally,

decreased distraction (probably due to providing a distraction-free setting for a reasonable time

period). Though careful evaluation and investigation into increased achievement levels for

intervention students in the course was not conducted in this study, the authors suggested that in-

class cooperative learning is practical and effective in improving students' learning activities.

In addition to effective use of collaborative grouping, the greatest results in at-risk

student performance has been observed in more intensive pedagogical intervention by course

staff in an alternative course track or within a preexisting course. Brahmia and Etkina (2001), Etkina, Gibbons, Holton, and Horton (1999), and Holton and Horton, (1996) outlined two courses that served as optional parallel course tracks for engineering majors (Extended Analytic Physics, EAP), and for science, science teaching, and pre-health professions majors (Extended General Physics, EGP). The development of the structure of the course was based on pedagogical research implementing cooperative learning and interactive engagement, specifically that of Arons, McDermott, Thornton, and Mazur (1997), Laws (1991), Tobias (1990), Hake (1998), and Redish, Saul, and Steinberg (1997). As such, the course elements included small section size, accessibility to course staff, weekly staff meetings, weekly review sessions, notebooks, mini-labs, group problem-solving workshops, laboratories, and interactive lectures, supplemented by aid in the Math and Science Learning Center on campus. A central feature of the two courses' structure was the focus on group work around all course components, causing coherence and overall student inclusion and accountability. Essential features of the course elements included weekly spiral curricular structure; carefully outlined, nontraditional roles for the course coordinators, TAs, and lecturer; and assessment and grading procedures aligned with the goal of positive, noncompetitive, and cooperative feedback and encouragement (Brahmia & Etkina, 2001). Grade assignment was distributed over several alternative assessments in addition to a multiple-choice exam. Students were evaluated on notebook quality, attendance, in-class performance, and preparation, which were intended to provide a clearer and more accurate picture of student progress and ability (Holton & Horton, 1996). Extended Physics made use of the real-world problems developed by Heller et al. (1992), which often require students to "make approximations or assumptions, create models, search for relevant data, and appropriate theory—often from several different phenomena" (Etkina et al., 1999, p. 813).

Student reasoning and justification of their decisions on important variables were frequently the center of discussion, helping students become more aware and critical of their thought processes.

Authors evaluated the alternative courses' success by examining retention and persistence rates. Results indicated that from 1985 to 1993, the number of African Americans remaining in an engineering major from freshman to senior year increased from 8% to 52% and female retention rates increased from 30% to almost 60% (Brahmia & Etkina, 2001). Etkina et al. (1999) also indicated that the drop-out and withdrawal rates for the Extended Physics (EP) courses (both EAP and EGP) were drastically lower than those seen in the regular physics courses, with 1% EP and 19.5% regular physics African American students dropping or withdrawing, and 5.5% EP and 15% regular physics Hispanic students. Brahmia and Etkina (2001) interpreted these substantial increases in student retention as positive effects caused by these students' enrollment in EP. While comparisons on immediate physics performance were not made between the EP students and a similar control group in regular physics, using a specially designed course alone to aid in achieving higher success in physics contributed to attaining the overall goal of higher retention rates of minority students.

By using many nontraditional elements that have been shown to be successful in physics instruction (Hake, 1998), Etkina and her colleagues created a physics learning environment where at-risk students, particularly women and minorities, could successfully participate in and master introductory physics. Those engaged in physics education and concerned with the performance of minorities and at-risk students in physics at the college and university levels are clearly focused on utilizing the classroom, curriculum, and instructional methods as places for the transformation of these students' science careers. The research outlined shows that by utilizing best teaching and learning practices for achievement for all students, researchers and

16

course administrators have made progress in the improvement of minority student performance in physics, and in engineering and science as a whole.

Solutions Implemented in College Science and Other Related Disciplines

As seen in the approaches in introductory physics, research has been directed at manipulating a variety of possible factors to bring about improved performance and retention levels for at-risk students in other disciplines. Programs and workshops providing help with target courses also were deployed in departments other than those in which these courses were offered. For example, Barlow and Villarejo (2004) and Mendez (2006) described programs for biology, chemistry, and physics and for algebra and calculus, respectively, which are motivated by programs and funding not originating in the content department. The Biology Undergraduate Scholars Program (BUSP) at the University of California-Davis (UCD) is an educational intervention program funded by the Howard Hughes Medical Institute Undergraduate Biology Education Program, the Initiative for Minority Student Development, and the National Institute of General Medical Sciences. As described by Barlow and Villarejo (2004), the program provided not only academic instructional support with supplemental instruction and facilitated study groups, but also financial support services, academic and personal advising, research laboratory employment, and peer networking. The McNeill program, initiated and supported by the Learning, Excellence, Achievement, and Diversity (LEAD) Alliance at University of Colorado-Boulder (UCB), serves low-income, first-generation, disabled, and rural and inner-city high school graduates. The program is offered as alternative algebra and calculus courses taken for the same credit as those offered in the mathematics department, but are managed completely by the program itself outside of the mathematics department.

Work by Beeber and Bierman (2007) approached at-risk student preparation for biology

at a small, urban community college in a similar fashion, namely, with a foundations course directed at skill development. While not all students participating were at-risk, the majority of students had completed a general education degree (GED) rather than having obtained a high school diploma, had poor background in the sciences, or was weak in mathematics or English. The preparatory course served as an alternative for a general biology course, which is a prerequisite in many of the curricula offered by community colleges. The content goals for the foundations course included teaching basic memorization skills; demonstration of conceptual understanding; data collection in a laboratory setting; interpretation of data through charts and graphs; quantitative manipulations, logic, and reasoning; utilization of tabulated and graphical data for problem solving; and presentation of findings. The instructional strategies mentioned were claimed to facilitate content goal achievement and consisted of a two-hour weekly laboratory session, a one-hour lecture/discussion session, and project completion. An associated unpublished study indicated that 71.3% of the students who took the foundations course compared to 80.9% of those who took the general biology course passed a human anatomy and physiology course that followed, leading the authors to conclude that students clearly benefited from the various aspects of the course prior to taking anatomy and physiology courses. No comparative analysis was conducted on the effectiveness of the different delivery methods used in the course.

A few notable studies stand out for the pedagogical approaches developed by the content department that also supported activities alongside the target course. In a line of research funded by the National Science Foundation, following a National Science Board Task Committee's suggestion (National Science Board, 1987, p. 4), Born, Revelle, and Pinto (2002) attempted to expand efforts in a biology course at University of California-Davis that could potentially lead to

increased participation in the sciences by female and minority students. Workshop groups, an idea developed by Fullilove and Treisman (1990) in mathematics, were formed with the expectation of improving biology performance "by imparting experiences and skills that should increase the subjective expectancy of success, reduce the fear of doing poorly, and thereby allow students to maintain their value for biology and remain engaged and enrolled" (Born et al., 2002, p. 352). The authors thought this approach would work especially well with minority students, who might try to avoid failure at all costs due to perceived stereotype pressures. Subgroups of 'minority' African American and Hispanic American students, and "majority" European American and Asian American students who were enrolled in the first quarter of a second year biology course in 1997 or 1998 and took at least one exam participated as volunteers for the intervention groups. A historical control group was constructed from minority students enrolled in previous years. About half of the volunteers were assigned to the workshop intervention, consisting of group work on problem sets for 2 hours per week; the other half were assigned as motivational control participants who did not receive any different treatment than the nonvolunteers, yet were considered a comparison group based on the criterion of volunteerism. Groups were told research had shown that working as a mutually supportive team resulted in higher grades and better understanding of complex material. With a focus on the intervention effectiveness with minority students, the authors found that, although no difference in total points earned existed between workshop minority and historic minority students, an interaction of workshop status and exam performance over time appeared. The workshop minority students performed similarly to the historic control minority students on the first two exams, but showed a marked increase in performance of 0.7 standard deviations on the third ($p < 0.05$). The workshop grouping on problem sets was considered an effective strategy for improving all participants'

19

performance, as measured against appropriate controls.

Drane, Smith, Light, Pinto, and Swarat (2005) also studied a concurrent workshop program, originally developed by Born et al. (2002), but including physics, chemistry, and biology courses designed to improve the retention and performance of minority students at a large university. The workshop sessions for the three different courses were each held for 2 hours per week. Students in the sessions worked in groups of 5 to 7 on solving challenging, conceptually rich problems developed by course professors while being facilitated by undergraduate teaching assistants. Participation was optional for biology and chemistry students, and was partially required by physics students in that they were allowed to enroll in the workshop sessions as a replacement for the regular discussion session if they so chose. Results in the physics component indicated that workshop participants saw no improvement over nonparticipants. Biology and chemistry students, however, were much more successful. The mean final grades of the workshop participants in both biology and chemistry were significantly higher than the nonparticipants by 0.4 and 0.5 standardized points, respectively. Furthermore, retention rates for both minority and majority students improved, showing that not only were minority participants more likely to be retained than minority nonparticipants, but also that the minority participants' retention rate nearly matched that of majority participants and nonparticipants.

Summary and Critique: Implemented Solutions

Several curricular and pedagogical strategies discussed here have resulted in some improvement in the area of at-risk student performance, retention, and persistence. Over time, the solutions to improve at-risk student achievement and retention rates have taken on trends that help distinguish the different lines of research. Three important distinctions can be made about

20

the research on solutions implemented in college physics, college science, and other related disciplines. The first is whether the initiating entity is within the content department or outside of the department. The second is related to the focus of the intervention, namely, whether the focus is curricular (what content to teach), pedagogical (how to teach the content), and/or includes other supports for students (financial, mentoring, advising, tutoring). The third relates to whether the change or intervention is implemented within the target course design, alongside or additional to the target course, or as an alternative for the course. Research involving the evaluation of the relative effectiveness of interventions related to these three foci has yet to be conducted; as such, no statement can be made about direct comparison of the different approaches discussed above. Furthermore, evidence for the effectiveness of the implemented solutions is not measured consistently across the various studies.

In at least some of the studies reviewed above, researchers framed the problem of at-risk student achievement and retention and the nature of the developed intervention in ways that suggest a deficit model perspective. A deficit model perspective assumes that learners' poor performance in school stems from a deficiency with the learners rather than with the educational institution itself, and associated administrative and organizational structures, or instructional practices (Valencia, 1997). Two different deficit models (genetic determinist and cultural determinist) focus on student characteristics framed as inherent student deficiencies, typically due to genetics, language, lifestyle, values, social structure, and/or cultural practices (Solorzano & Yosso, 2001). The selection and implementation in the reviewed studies has mostly been undertaken at the student and course levels and neglects to invite critique and discussion around the social and political beliefs and policy from which the perceived problem arises. Inherent in the choice of intervention is a belief of the onus of responsibility; that is, interventions were

21

directed at the students, not at the educational system within which they navigate learning and have little power to change. While offering some effective and creative solutions for improved at-risk student performance and persistence, programs directed at the individual student and target courses withhold the responsibility for improved performance for at-risk students from the institutions of higher education. Interventions based in a deficit-model perspective perpetuate the latter mindset by failing to express interest in fundamentally changing relevant policies and practices in higher education.

Seminal Work by Treisman

Much of the research done on improving at-risk or minority student performance in the sciences and mathematics in higher education stems from the seminal work undertaken by Treisman (1985), as does the associated instructional models (Fullilove & Treisman, 1990; Treisman, 1992). This reform model was anchored in Treisman's dissertation (1985) conducted at the University of California—Berkeley (UCB) during the 1975–1976 academic year to understand the differences in performance between two first-year ethnic minority groups—one that generally did well in the UCB mathematics classes and another that did not. The original focus of the study shifted from factors predicting success in calculus by first-year students to factors explaining differences in the observed performance between the two groups of students. Special interest emerged because the strong and weak students shared one important characteristic—membership in an ethnic minority group. Treisman collected data in the form of videotaped student work both inside and outside of class, which he described as making a decision "literally to move in with the students and to videotape them at work" (Treisman, 1992, p. 365). Extensive videotaping and observation of about 20 African American students (who experienced a 60% failure rate as a group) and Chinese American students (who were the most

22

successful of the student groups in calculus) provided detailed insight into the study strategies leading to the observed performance differences.

Many misconceptions were dismissed by the analysis of this video, including ideas that minority students were unmotivated, unsupported by families, unprepared, and relatively poor on the SES scale. The most influential finding emerging from Treisman's dissertation was the discovery that the two groups used different study strategies for completing course work and preparing for exams. The Chinese American students were more likely to follow independent study with a substantial amount of study time with peers both enrolled in the course, as well as those who previously took and performed well in the course (often roommates or siblings). This group study often took the form reciprocal home visits, cooking and eating meals together, and going over homework assignments or testing each other with practice exams. Fullilove and Treisman (1990) expanded on the advantages of the group work:

> The Chinese American study groups facilitated the exchange of information between group members, and this exchange became a critical component of each student's mastery of calculus. Students checked each others' work, pointed out errors in each others' solutions, and freely offered each other insights that they had obtained—as a result of their own efforts or through conversations with TAs or professors—about how to manage difficult problems. Although mathematical insight may be derived from many sources, the most dramatic advantages of working in groups were observed in the ways group members corrected misperceptions or errors in their strategies for working problems. (pp. 466–467)

The social interactions not only provided students with opportunities to get specific feedback on their individual work, but also to help them identify what areas they did not understand

compared to their high performing friends. This led to the overall effect of the Chinese American students knowing how well they had mastered the material and where their current understanding stood in relation to their peers, as well as compared to what was expected of them on exams. Such feedback provided vital information for further study strategies. The African American students engaged in about the same amount of independent study time, but in contrast, did not involve themselves with any social study strategies. Indeed, Treisman (1992) reported that 18 out of the 20 participant African American students never studied with their classmates. These students kept their social and academic activities separate, thereby missing a vital opportunity that was concluded to be the Chinese Americans' "edge" for performing at higher levels.

Mathematics Workshop Program. Over the next ten years, Fullilove and Treisman (1990) utilized these observations to develop the Mathematics Workshop Program (MWP) at UCB. The program was designed to facilitate collaboration among students in first-year calculus, especially with African American and Hispanic American students. Minority students accepted to UCB were reported to perform at levels reaching two standard deviations above national means on such tests as the Scholastic Aptitude Test Verbal (SAT-V) and Mathematics (SAT-M) (Fullilove & Treisman, 1990). As such, these high performing students rarely arrived at UCB with beliefs that they were deficient in key skills or needed remediation in any areas. Therefore, the workshop was introduced as an honors program and recruitment is directed at students who typically held perceptions of themselves as high achievers and who also showed promise to maintain a willingness to work hard to achieve academic excellence. Students enrolled in MWP attended classroom sessions in a way that was very similar to concurrent enrollment in discussion sections accompanying lectures. In these sessions, workshop's participants worked in groups of 5 to 7 on packets of carefully constructed, unusually difficult problems. In an attempt

24

to reconstruct what was observed to be the Chinese American students' strategy of success found in Treisman's (1985) study, the MWP students were encouraged to talk at length about the problems, potential solutions, and how to execute the solutions. In these sessions, though students could work alone for as long as they liked and choose the problem order in which they wanted to work, at some point they were required to share what they had done and provide feedback on other students' work. Instructional assistance was provided to the students in the form of a graduate teaching assistant. The TA's role centered on assessing group and individual progress, asking key guiding questions or giving suggestions when a group was stuck on a solution, and stepping in with individual help when a student was particularly struggling.

Effectiveness was determined by inspecting failure rates for MWP versus non-MWP participants prior to and after the implementation of the program (note that students dropping the course were not included). In the five years before the program, the percentage of African American students receiving a grade of D-plus or lower in first-year calculus ranged from 27% to 43%. The African American student failure rates in the five years following the program ranged from 0% to 6 % for MWP and 24% to 52% for non-MWP students. The differences in failure rates were significant at the $p < 0.0001$ level. Furthermore, comparing percentages of students who achieved high performance in the target course showed that over the years, MWP African American students were twice to three times as likely to earn a B-minus or better than non-MWP African American students ($p < 0.01$). These results were sustained even when enrollment numbers increased to nearly three times than in previous years. Conclusions derived from data analysis centered on the indisputable out-performance of MWP students in comparison to their non-MWP peers, even after controlling for prior academic measures. Speculation about the program's success introduced the ideas that (1) the self-selecting students were committed to

maintaining the levels of performance they value, (2) students spent more time in more efficient ways when doing work for the course, and (3) the students were, as the program design intended, connected to a broader social network within the university that promoted and stimulated the desired social interactions as seen with Chinese American students in Treisman's (1985) study.

Influential factors. Fullilove and Treisman (1990) found that the key to the MWP's success was the type of problems included in the problem sets. The problems provided to students were selected not only based on the level of challenge they provided, but also reflected the degree of skills and conceptual understanding calculus students were expected to demonstrate on quizzes and examinations (i.e., alignment among learning objectives, teaching objectives, and assessments). The researchers viewed the problem set worksheets as the core of the program's efforts that benefited students in three ways. First, the problems promoted practice of the skills to earn a final grade of A in first-semester calculus. Second, the problems served as a foundation in mathematics that the workshop graduates could transfer into second-semester calculus and upper-division mathematics without further assistance of the program. Finally, the problems provided opportunities for students to identify areas in mathematics knowledge that they needed to strengthen for complete understanding. The feedback from other students as well as from teaching assistants provided an additional effect by encouraging and supporting students to continue the learning process until they achieved levels of mastery required to perform at high levels in the course.

Prior to Treisman's work and its dissemination, research and actions directed at decreasing at-risk student failure and attrition rates tended to be more sociopolitical in nature, with college administrative and academic offices supporting special programs. These programs centered on providing resources for students, such as financial aid, mentors, and tutors to help

them in the days before an examination. Treisman's model brought retention concerns and efforts into the hands of the course's department and faculty. Students recruited to enroll in MWP were not viewed as at-risk for failure or attrition, but rather motivated students with a history of being high achievers in their previous academic environments. Treisman (1992) said:

> We did not question that minority students could excel. We just wanted to know what kind of setting we would need to provide so that they could. We also recognized early on that we would be successful only if we depoliticized the issue of minority access. We had to link our program with other issues that the faculty cared about, such as producing quality majors, and de-emphasize the purely political characteristics of the program so that it could take hold in academic departments. (p. 369)

This perspective, which highlighted students' competency and resourcefulness—informed and supported by insights from video analyses of Treisman's (1985) study—was contrary to ideas or assumptions that underlay many early studies, namely that minority students lack motivation, need remedial help from targeted programs, and/or are best served by segregating them from other students, including high performers.

Treisman's 1985 work promoted creating learning environments where at-risk students are viewed as powerful actors in their own education, providing strong academic resources at the individual level and strong social influence at the group level. Treisman (1992) maintained that the calculus problem sets used in the MWP were what drove student collaboration, not any focus on group learning tactics. The nontraditional impetus of the program design was not focused on student preparation or remediation, but rather on fostering effective study strategies that play strong roles in producing future mathematicians or at least students who pursue graduate degrees in the field.

Interventions utilizing the MWP model. A number of different programs have since been developed following the Fullilove and Treisman (1990) model. After spending time developing the program at UCB, Treisman transported the MWP idea as the Emerging Scholars Program (ESP), as implemented at the University of Texas-Austin since 1988 (Moreno & Muller, 1999; Moreno, Muller, Asera, Wyatt, & Epperson, 1999). The model has been adopted and adapted at more than 100 institutions in the United States, more recognizably as MathExcel at the University of Kentucky at Lexington (Freeman, 1995, 1997) and at Oregon State University (Duncan & Dick, 2000), as the Academic Excellence Workshop Program at California State Polytechnic University of Pomona (Bonsangue, 1994; Bonsangue & Drew, 1995), and as the Wisconsin Emerging Scholars program at University of Wisconsin at Madison (Alexander, Burda, & Millar, 1997; Kosciuk, 1997). Though the majority of programs adapting the Treisman model remain in mathematics, programs have been created for content areas other than math, such as the Medical Scholars Program at the University of California San Francisco School of Medicine (Fullilove, Fullilove, Terris, & Lacayo, 1988), the Workshop Chemistry Project at City College of New York (Gosser et al., 1996), and BioExcel at the University of Kentucky (Cohen, 1997). The few mentioned (Born et al, 2002; Drane et al., 2005) are examples of how the Treisman Workshop model for minority and at-risk students has been implemented with introductory college science courses. Indeed, departmental implementation of the collaborative grouping has been studied in physics (Cottle & Hart, 1996) and influenced ensuing research (Etkina et al., 1999).

A Theoretical Foundation of Instructional Approaches

A dissection of the instructional strategies and the establishment of a theoretical framework around the strategies can shed light on why such methods have been successful and

how to improve on them. One manner of mapping out the distinguishing features of any learning

content and instructional strategy is to identify the principles guiding the instructor's and

instructional decisions. Any instructor inherently possesses a set of implicit and explicit

beliefs—including an epistemological stance—and is motivated from some sort of personal

theory about how teaching and learning take place. An epistemological stance is a set of beliefs

about the nature of knowledge to be learned, how that knowledge develops and is most

effectively learned, and how and why we know what we know. Though this set of beliefs is not

always deliberate or well articulated, it intrinsically provides the basis for decisions regarding

curriculum and pedagogy (C&P) and therefore can be construed from choices in content and

delivery.

Identification of epistemological stances in effective interventions. In modern learning

and teaching theories, the behaviorist, cognitive, and human (or sociocultural) perspectives

constitute three broad epistemological stances regarding knowing and learning. Behaviorist

perspectives concentrate on the types of behaviors and tasks students are able to perform

resulting in routine forms of learning aimed at factual and procedural accuracy. Another broad

category for learning theories, human perspectives, concentrates on the characteristics humans

bring to bear on the learning process and the processes of interaction of individuals with people

and tools in their environments. Both individual human and social factors emphasize the idea

that learning and development cannot be divorced from the social aspects and contexts that

mediate such occurrences. The last broad category, cognitive perspectives, subsumes learning

theories that concentrate on the processes that occur in the mind and brain to enable learning.

The latter theories concentrate on how cognition occurs and what sorts of strategies allow or

discourage the formation of memories, schemata, and understanding. Contrary to behaviorist

29

perspectives, cognitive perspectives provide a model of learning that incorporates the mind and brain mechanisms, concentrate on how cognition occurs, and predict what sorts of strategies allow or discourage the formation of meaningful learning. Of interest are how the system constructs knowledge and how the cognitive structure accommodates experiences. Cognitive theories of learning tend to focus more on the mental activities students are to engage in, whereas the human perspective theories focus more on what occurs (particularly the social interactions) within the classroom environment (Brown, 1975; Brown, Collins, & Duguid, 1989; Deutsch, 1949; Flavell, 1973; Greeno, Collins, & Resnick, 1996; Lewin, 1935; Novak, 1977; Piaget, 1926; Vygotsky, 1978; Wertsch, 1985).

Each of these perspectives has related and different sub-stances, which could be connected to learning theories. While learning theories do not prescribe the construction of learning opportunities or environments, they focus attention on those variables that are crucial in choosing and building curriculum and pedagogy that are most effective to address specific learning outcomes. Thus, there is certain target knowledge to be mastered, as well as actions and expectations for students and teachers, and implications for the structure and type of learning activities based on an understanding of the very development of the target knowledge domain. An epistemological stance and associated learning theory can be both explicitly and/or implicitly exhibited through the choice of curricular material and pedagogical approaches. Evidence for an epistemological stance and learning theory can therefore be inferred from curricular materials, teacher behaviors, student behaviors, teacher-student interactions, student-student interactions, as well as embodied in the learning materials, learning environment, and assessments. These latter aspects are more fully developed and identified in Chapter 4 toward the construction of the intervention course profiles.

Identification and instructional embodiment of underlying educational perspective. Of the work and research already discussed, the cognitive perspective has influenced the initial transition from more traditional behaviorist methods of physics instruction. Broad curricular reforms occurred as the cognitive perspective filtered into physics education, which originated in greater awareness of mental states and schema, and of the cognitive structures that enable the accommodation of learning experiences.

Of the different learning theories encompassed by cognitive perspectives, constructivism serves as a central influential theory. In various forms ranging from radical-individual to socio-cultural formulations, the defining characteristic of a constructivist perspective is the self-regulation of the elements and structures of the cognitive system: the learner consciously and actively constructs his or her own knowledge. As summarized by Fosnot (1996), "Rather than behaviors or skills as the goal of instruction, concept development and deep understanding are the foci; rather than stages [of cognitive development] being the result of maturation, they are understood as constructions of active learner reorganization" (p. 10). Constructivists value richer, fuller understanding of conceptions that often arise from complex reasoning processes accounting for understandings learned from prior experiences. A major factor contributing to the adoption of constructivism is this emphasis on using prior knowledge to facilitate and scaffold new learning. Research has shown that alternative student conceptions, when not isolated, articulated, and built upon, can render more traditional types of instruction ineffective (Bransford et al., 1999).

Constructivism requires awareness of one's own thoughts and reasoning processes, critical thinking skills, and creativity to engage in meaningful learning situations. Learners must learn to raise their own questions, examine their own reasoning, generate hypotheses for

31

alternative conceptions, and map out consequences or tests for the viability of alternatives. Reasoning about and extracting generalities, trends, and abstractions are encouraged in order to meaningfully organize experiences and information. Errors in reasoning are perceived as valuable resources for further learning when deliberately acknowledged and challenged for affirming or contradicting logical soundness. Learning essentially takes place in a social context, intertwined with thinking that occurs independently, but also within a community of learners. In constructivism, according to Fosnot (1996), it is the learners—not the teachers—who are responsible for "defending, proving, justifying, and communicating their ideas to the classroom community" (p. 30).

The instructor in constructivist learning environments, thus, directs his or her energy toward facilitating key mental states and processes to occur through the instructional materials and social interactions. Scaffolding and modeling are foundational instructional techniques used to facilitate the learner's self-regulatory episodes. While the importance and extent of scaffolding can be discussed at length (see Bickmore-Brand & Gawned, 1993), its primary purpose is to direct the learner's attention to his/her own conceptions regarding a topic, extending or challenging the implications of the conception, and offering alternative conceptions or possibilities for consideration when the conception does not adequately explain or account for experiences or observations (Fosnot, 1996). In this way, learning becomes dialogical for the learner, whether with him/herself, with someone who has greater mastery of the content or process to be learned, or alongside someone also striving to more fully understand the conception.

It was constructivist ideas and those who were willing to adopt more student-centered approaches to instruction that transformed science courses in higher education more than 30

years ago. Pioneering at-risk student interventions in physics mirrored the transition of undergraduate science education from more behaviorist means of instruction (i.e., "traditional" and focused on information transmission with little to no student participation) to those that were more student-centered and collaborative in nature.

Constructivist techniques are abundant throughout the literature. Initial implementations from the broad stance of cognitive perspectives required major epistemological shifts on the part of instructors, as well as requiring students to adjust their expectations and practices about what it means, or what it takes, for them to "learn." One common way of viewing physics is to equate it to retaining formulae, and memorizing and practicing successful and efficient problem-solving algorithms. Traditional methods of teaching physics grounded in the image of learning as absorbing knowledge include extended, passive lectures; confirmatory, highly structured laboratories; and observing instructors solving problems followed by drill-and-practice sessions to master those solutions. Constructivist cognitive perspectives transformed this image of learning into one in which student mental activities and cognitive processes involved in building knowledge are the most crucial agents in learning. With such transformed epistemology, the role of the teacher shifts from that of a source of knowledge to a facilitator in the discussion and building of knowledge and understanding. This fundamentally changes the power dynamics in a classroom and requires the teacher not to act as the authority in the classroom but as a mediator and coach. As such, early reforms in college science courses might, in retrospect, look rather minor; but they did entail a much larger shift in the underlying epistemological perspective and learning theory. Because they tend to build upon one another, reforms and transitions extended the practice of constructivism as a learning theory in each iterative practical application. Over the years, constructivism has been enacted more thoroughly in practice than in its early adoptions.

The epistemological adjustments and curricular reforms provided the building blocks on which further constructivist techniques could be incorporated. A fuller reform emerged from the resultant pedagogical considerations and modifications. Constructivist theories implied many techniques for how to create effective learning environments and teaching opportunities. More deliberate thought was accorded to the types of activities that provide scaffolding and modeling for student thinking and self-regulation processes.

Theoretical Perspectives Framing Constructivism and Collaboration

Springer, Stanne, and Donovan (1999) conducted a meta-analysis of research conducted on small group learning in science, technology, engineering, and mathematics (STEM) courses. They concluded that, "Conceptual frameworks for small-group learning are rooted in such disparate fields as philosophy of education (Dewey, 1943), cognitive psychology (Piaget, 1926; Vygotsky, 1978), social psychology (Deutsch, 1949; Lewin, 1935), and humanist and feminist pedagogy (Belenky, Clinchy, Goldberger, & Tarule, 1986)" (p. 24). As a manner of organization, they developed motivational, affective, and cognitive theories as three interrelated theoretical perspectives to account for the effects of small groups on academic achievement.

The *cognitive perspective* entailed that through group interactions, students process information more intensely and effectively. Generally grounded in the pioneering work of Piaget (1926) and Vygotsky (1978), these theories focus on cognitive growth facilitated through interpersonal interactions on open-ended tasks with many possible paths leading to multiple acceptable solutions. Mirroring constructivist views, Springer et al. (1999) summarized this viewpoint as valuing the opportunity for students to discuss, debate, and present their own perspectives, as well as hearing one another's perspectives in small-group learning. They also noted that from a cognitive perspective, learning occurs because when students discuss the

content with one another, "cognitive conflicts will arise, inadequate reasoning will be exposed, and enriched understanding will emerge" (p. 25).

The *motivational perspective* highlights the supportive nature of group dynamics through encouraging and helping group members achieve because there is value placed on the success of the entire group, not just its individual members. Thus, emphasis is placed on individual accountability, for if one person fails, the whole group fails and as such, sharing answers, completing each other's work, and competing with one another are not desirable or permissible in this social structure. Individuals take responsibility for teaching one another and regularly assessing one another's learning through the motivating incentive of group success and accomplishment.

Under the *affective perspective*, the instructor strives to create a nonthreatening learning environment where everyone, especially marginalized individuals or groups, have greater opportunities to be heard. He or she does this by maintaining a nonauthoritative role and cultivating more collaborative and democratic teaching and learning processes among the members of the learning community, within both small student groups and the larger classroom or course community.

The perspectives in practice. The previously discussed distinctions by Springer et al. (1999) set the stage for better dissecting the ways that different constructivist instructional approaches have been used in physics courses, particularly with at-risk students. Early interventions might have incorporated group activities for students, but these did not seem to deliberately utilize the principles provided by social cognitive theory in building curricular materials and choosing pedagogical approaches, as highlighted by Springer et al. (1999), to enhance at-risk student learning. For example, Taylor (1983) did not intentionally use small

35

groups as a crucial pedagogical approach though she mentioned their occasional use to discuss particularly difficult concepts.

However, capitalizing on the social aspects of learning in combination with scaffolding and modeling acknowledges the effects that others have on an individual's learning. Social constructivism gained strong ground because it was shown that humans learn not only through their own experiences, but also by interacting with others. This dynamic catalyzes opportunities for one to become more aware of his/her own conceptions and articulate and evaluate them. An intentionally structured social context can provide scaffolding and modeling from instructors as well as from other students. Cottle and Hart (1996) and Etkina and her colleagues (Brahmia & Etkina 2001; Etkina et al., 1999; Holton & Horton, 1996) used small group work as the *key* element to improved performance and, following Fullilove and Treisman (1990) and Treisman (1992), utilized small groups to improve both teaching and learning.

Collaborative grouping in minority and at-risk student interventions in the MWP/ESP Treisman Model of instruction (Fullilove & Treisman, 1990; Treisman, 1992) not only capitalized on promoting learning through the cognitive perspective but also included several essential features from both the motivational and affective perspectives. Student groups in the workshop sessions were encouraged to discuss how to manage difficult problems, point out each other's errors, correct misperceptions, share insights or learning from other sources, and generally facilitate information processing. The MWP's close attention to the role of instructors as coaches offering help through guiding questions or suggestions also highlighted the techniques inspired by the affective perspective. Similarly, the motivational perspective was woven into the expectation for students to work and solve problems as a group, helping one another and providing individual feedback on each other's work.

36

Etkina and her colleagues (Brahmia & Etkina, 2001; Etkina et al., 1999; Holton & Horton, 1996) were able to take the constructivist approach further than Taylor (1983) did in their consideration of the learning community that engaged their students, providing not only cognitive but also motivational support. Etkina and her collaborators created a sense of community through the cooperation of students and instructors. This was achieved by incorporating structured group work, continuous feedback, follow-up on exam performance, and ample course staff availability. Students working long-term in teams of three provided immediate help and discussion of individual ideas, and deterred the sense of isolation that at-risk students often experience in college-level courses. These strategies additionally addressed other at-risk student difficulties, such as low confidence, feeling undeserving of success (also known as "impostor" syndrome), and lack of community (Brahmia & Etkina, 2001). The nontraditional role of the instructor in the EP classroom was established and itself scaffolded through faculty and teaching assistant training. The instructor training was crucial to break prior conceptions and expectations about teaching and learning. Instructors were able to model and encourage hypothesizing of alternative solutions or strategies, and help students externalize their mental processes and conceptions by asking probing questions. Simply put, the instructors were trained to serve as facilitators of learning by guiding students' cognitive processes rather than providing answers.

Summary

Utilizing group work in conjunction with constructivist techniques has been recognized as an effective strategy for improving performance of at-risk students in introductory physics and other science courses, and some study and discussion have taken place around the reasons for its effectiveness. As detailed at length here, this approach originated as a strategy for at-risk student

intervention by attempting to emulate the dynamics observed to work for other successful student groups. Linking these strategies and dynamics to the prevailing epistemological stance and current learning theories—identifying and understanding the reasons they work in a framework of teaching and learning and thus providing grounding for further improvements– were not investigated in depth in prior research. Over time, however, the idea of group work filtered in from different perspectives developed within cognitive and social psychology. The effective dynamics and reasons for success of the implementations of group work and social learning environments were then tied more closely to research to provide a broader foundation for understanding human learning and what influences it (Bransford et al., 1999).

CHAPTER 3

METHODOLOGY

This study addresses a gap in the literature related to the lack of quasi-experimental, quantitative studies that provide a delineation of the effectiveness of interventions for at-risk students in college physics. The study provides a means to assess the differential impact of two approaches to improving at-risk student performance in an introductory college physics course. Each of the two approaches took the form of an "intervention course." The first intervention course is completed one semester prior to the target course (i.e., the course where improved student performance is expected); it covered the first hour exam content in the target course and was focused on conceptual understanding, reasoning, and problem-solving strategies. The second intervention course was administered concurrent to the target course and mirrored the exact content covered in the target course. This second course is administered as a supplemental two-hour problem-solving session each week, which also focuses on the same conceptual understanding, reasoning, and problem-solving strategies as the Pre-Course with emphasis on creating problem solution strategies (roadmaps) by implementing a provided problem-solving process. A more complete description of the interventions is provided in Chapter 4.

The design allowed for the quantitative isolation of the differential effects in student performance scores for the Pre-Course and Concurrent-Course (Con-Course) intervention compared to no intervention. The design also allowed for the quantitative study of the persistence effect in high-risk student completion of the target course.

Design

The study used a quasi-experimental, posttest only, nonequivalent control group design. Inclusion in the intervention courses (the Pre-Course only, or the Pre-Course plus concurrent

course), or no intervention served as the independent variable for this study. The dependent variables are measures of performance and persistence rates. In the tradition of understanding influential factors on college physics performance (Hazari, Tai, & Sadler, 2007; Sadler & Tai, 1997, 2001), additional available student information were included in propensity score matching analysis to determine variations of effectiveness across student characteristics. A Mechanics course served as the target course for improved performance and persistence of students enrolled in the intervention courses.

Participants in the study belonged to one of four groups (see Figure 1). The intervention
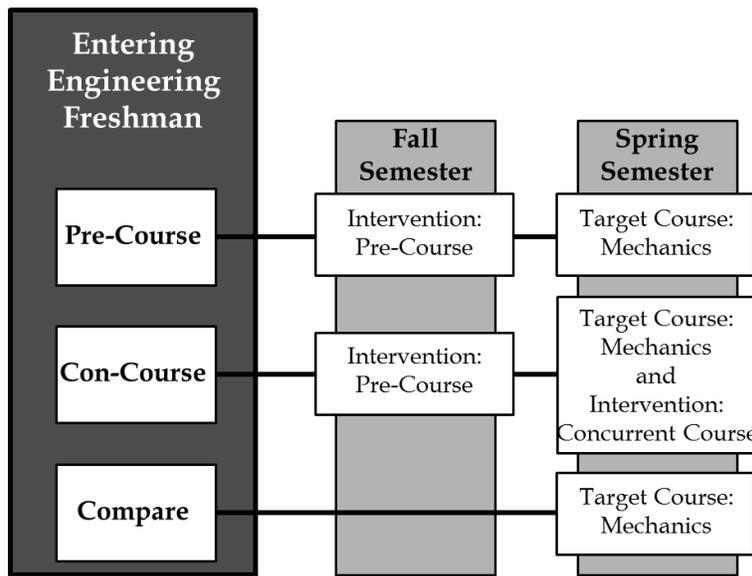
Figure 1: Chronology of Intervention and Target Courses

group comprised two subgroups: Group 1 included students enrolled in the first intervention (Pre-Course) in the fall semester and then in the target Mechanics course in the spring semester. Group 2 students also were enrolled in the first intervention (Pre-Course), then proceeded to simultaneously enroll in the target course (Mechanics course) and in the second intervention (Con-Course) during the spring semester. A third group served as the comparison group and included students who enrolled in the target course in the spring semester of their freshman year

without any pre- or concurrent intervention course (Compare). Participants did include a fourth group of students who were simultaneously enrolled in the concurrent intervention and target courses. However, the number of students in this group was too small to conduct any meaningful and significant comparisons (see Table 1). Thus, these students' data and scores were excluded from further analyses.

Prior to registering as entering freshman, students entering since Fall 2010 are required by the Physics Department to complete the Physics Diagnostic (PD) the summer prior to enrollment. Academic advisors in the College of Engineering—which houses the Physics Department—use the 13-problem PD placement exam to assist students in course selection. Students who score 7 or below on the PD are advised, but not required, to enroll in the Pre-Course. Students scoring between 8 and 9 points, after a brief evaluation of their background and national test scores, are recommended to enroll in the Pre-Course if there is indication of weakness in mathematics. Students performing at a level of 10 or higher on the PD are permitted to enroll in the intervention courses—though the majority of these students choose not to—and this decision is made at the individual level. Important to note is that the intervention courses are not required for any major curriculum for graduation, so no students are required to enroll in them; that is, enrollment is completely voluntary.

Measures

Independent variables. As previously stated, enrollment in an intervention course serves as the independent variable. This variable has three levels: Pre-Course intervention, Pre- and Con-Course intervention, and no intervention course.

Control variables. Control variables included characteristics of the participants, which have been shown in prior literature to explain variance in performance in introductory

mechanics. Participants' demographic information included gender, race/ethnicity, citizenship, state residency status, and first generation college status. Control variables also included students' high school, federal financial aid data, PD exam scores, and national exam scores. The latter comprised ACT Composite, English, Math, Reading, and Science scores, as well as AP scores for Calculus AB, Calculus BC, Physics B, and the Physics C Mechanics and Electricity and Magnetism scores. High school information included high school metro locale code, high school type (public, private, denominational), and percent of student body designated as low income for in-state students. High school background information also included years of precalculus, calculus, and physics. Financial information submitted on the Free Application for Federal Student Aid (FAFSA) submitted upon freshman enrollment included student adjusted gross income, father's yearly earnings, mother's yearly earnings, parental adjusted gross income, net worth of parents' investments (including real estate but excluding parents' home), net worth of parents' business and/or investment farms (excluding family farms or family businesses with 100 or fewer full-time or equivalent employees), number of people in the parents' household, and Pell Grant eligibility. Finally, diagnostic scores for all students were necessary for recommended placement into the Pre-Course treatment.

Outcome variables. Data routinely maintained by the Physics Department at the participant university for evaluation of progress in the intervention and target courses provided primary evidence about student performance. These included final course grade, hour and final exam scores (three midterms throughout the semester plus one final, designated as Ex1, Ex2, Ex3, and Final, respectively), and total points earned in the course. Total course points include components from the three one-hour exams, the final exam, in-class quiz scores, online homework scores, and laboratory scores. A composite exam score was computed for each

42

student [ExamComp = (Exam 1 + Exam 2 + Exam 3 + 2*Final)/5] to isolate individual

performance-based knowledge from the participation components of the course. Students who

did not complete the final exam were omitted from the analysis, and students who withdrew from

the target course were treated in a separate analysis related to persistence. If the target course

was repeated, similar performance measures were collected for the second enrollment.

Persistence has traditionally been conceptualized as a student's continued behavior

leading to a desired goal. Students enrolling in the intervention courses were presumed to have

the goal of completing the target course with a passing grade. Persistence was therefore defined

as eventually earning a passing score in the target course, after either the first or several attempts.

The effect of the intervention courses on persistence of at-risk students was studied by examining

the percentages of students who eventually earned a passing score in the target mechanics

course.

Participants

The participants in the study include 1591 engineering students admitted as freshman in

the Fall 2010 and 2011 semesters to a large Midwestern university and subsequently enrolled in

a calculus-based introductory physics course sequence. All students entering the course

sequence—Mechanics, Electricity and Magnetism, Quantum Physics, and Thermal

Physics—have successfully completed Calculus I at university level or received AP credit for

this course. All are United States citizens and had completed a Physics Diagnostic required for

advising the summer prior to university enrollment. All students in the study enrolled and earned

a letter grade in the target course, regardless of enrollment in an intervention course.

Descriptive Statistics and Initial Differences

In addition to generating a general description of students participating in the study,

Table 1

Number of Participants

Group	N	Name
Pre-Course only	508	Pre-Course
Pre-Course + Con-Course	62	Con-Course
Con-Course only	4	(not included in study)
Comparison	1017	Compare

descriptive and inferential simple statistics were used to compare the two treatment groups and the comparison group on the aforementioned control variables. A number of initial differences are observed across the three groups for several of these variables.

Means of control variables. Means for all collected control variables were calculated, as presented in the tables below. The first significant difference seen is in regard to Physics Diagnostic score, which is expected due to its use by advisors to guide students in intervention enrollment. Out of a maximum 13 points total, the PD score for the Compare group was 10.0 ± 0.07, for the Pre-Course group 6.9 ± 0.08, and for the Con-Course group was 6.4 ± 0.25.

Across the demographic variables (Table 2), significantly more women and first-generation students and fewer Asian students enrolled in the Pre-Course. Significantly more out-of-state residents and black students and fewer white students enrolled in the Con-Course intervention. Overall, women comprise only around 20% of the engineering freshman student body. Furthermore, the entry rate of 1.9% Black students is much lower than the 14.5 % of state residents and 12.6% of United States (US) citizens (United States Census Bureau, 2012). The

Table 2

Student Characteristics – Percentages

Variable	Mean			Difference of Means			
	Compare	Pre	Con	ΔPre	p	ΔCon	p
State resident	98.2	95.6	83.8	-2.6	**	-14.4	**
Female	16.0	25.8	19.4	9.8	***	3.4	
First generation	16.9	21.3	17.7	4.4	*	0.8	
Pell Grant recipient	11.7	13.6	11.3	1.9		-0.4	
Asian	22.2	17.5	24.2	-4.7	*	2	
Black	1.2	2.6	9.7	1.4	.	8.5	*
Hispanic	6.4	6.9	8.1	0.5		1.7	
Multi-race	2.9	3.3	1.6	0.4		-1.3	
White	66.3	68.9	56.5	2.6		-9.8	*

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

same statement can be made regarding the Hispanic student entry rate of 6.6% compared to the 15.8% state residency and 16.7% US citizen rates.

Scores for scholastic aptitude and prior knowledge show significant difference among the groups. Table 3 shows that the intervention groups scored significantly lower on the ACT English and Math tests, both of which are known to influence physics learning and thus correlate to college physics achievement. ACT Reading and Science scores indicate that, for the percentage of students who participated in these optional tests (87.4%, 93.7%, and 90.0% of the comparison, Pre-Course, and Con-Course students, respectively), the intervention students do not possess the same science reasoning and comprehension skills as their comparison peers. The comparison group performed at the 95th, 90th, and 97th percentiles compared to national averages for the Composite, English, and Math scores (ACT, 2013). Accordingly, the Pre-course

Table 3

American College Test (ACT) Scores and Number of Students Reporting Scores

Variable	Mean			Difference of Means			
	Compare	Pre	Con	ΔPre	p	ΔCon	p
Composite	30.4	29.7	28.8	-0.7	***	-1.6	***
English	29.7	29.3	28.9	-0.4	*	-0.8	*
Math	32.1	30.9	30.3	-1.2	***	-1.8	***
Reading	29.2	28.9	27.3	-0.3		-1.9	**
Science	29.7	28.9	27.7)	-0.8	***	-2.0	***

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

group performed at the 94th, 90th, and 96th percentiles; the Con-course group at the 92th, 90th, and 95th percentiles.

Table 4 reports Advanced Placement (AP) scores for the different groups (maximum score is five). Again, the intervention group students earned significantly lower scores than the

Table 4

Advanced Placement (AP) Scores

Variable	Mean			Difference of Means			
	Compare	Pre	Con	ΔPre	p	ΔCon	p
CalcAB	4.15 (65.6%)	3.74 (57.7%)	3.56 (62.9%)	-0.41	***	-0.59	**
CalcBC	4.34 (31.6%)	3.85 (18.3%)	3.82 (17.7%)	-0.49	***	-0.43	
PhysicsB	3.69 (20.3%)	3.24 (11.4%)	2.75 (6.5%)	-0.45	**	-0.94	**
Physics C–Mechanics	3.32 (19.2%)	2.75 (4.7%)	2.33 (9.7%)	-0.57	**	-0.99	**
Physics C–E&M	3.30 (12.6%)	2.33 (2%)	3.00 (3.2%)	-0.97	**	-0.30	**

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

comparison students. The percentages of students reporting AP scores (in parentheses in Table 4) indicates that a large percentage of the students in the study did not or were not able to engage in courses promoting rigorous or advanced study of introductory physics content prior to entering as a freshman at the institution in the study. Combining this information with the different number of years of high school course enrollment (Table 5), the intervention students have significantly less experience with and prior knowledge about physics content.

Table 5

High School Education (Years)

	Mean			Difference of Means			
Variable	Compare	Pre	Con	ΔPre	p	ΔCon	p
Precalculus	0.96	0.96	0.95	0.00		-0.01	
Calculus	1.08	1.05	1.09	-0.03	**	0.01	
Physics	1.49	1.23	1.20	-0.26	***	-0.29	***

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

Summary statistics for household financial information reported on the FAFSA (for those students and families who filed) shows that students in the Con-Course intervention originate from households with significantly lower parental yearly incomes (Table 6). Students who enrolled in the interventions also originate from households with significantly lower investment worth, which serves as proxy for accumulated wealth. Predictor research, as reviewed in Chapter 2, has indicated that family wealth correlates positively with college physics performance and as such, students who come from less-advantaged households in this regard are at higher risk. Information about the characteristics of the in-state high school is summarized in Table 7. Students who attended suburban high schools did not pursue enrollment in the Con-Course in the

Table 6

Reported Financial Information

	Mean			Difference of Means			
Variable	Compare	Pre	Con	ΔPre	p	ΔCon	p
Student's income	$1,470	$1,059	$985	-$411		-$485	.
Mother's income	$33,541	$38,464	$35,009	$4,923	.	$1,468	
Father's income	$88,168	$84,644	$56,312	-$3,524		-$31,856	***
Parents' income	$120,967	$116,833	$95,707	-$4,134		-$25,260	**
Parent business worth	$15,670	$5,198	$8,500	-$10,472	**	-$7,170	
Parent investment worth	$107,697	$89,051	$50,916	-$18,646	.	-$56,781	**

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

same numbers as students whose high schools were located in towns. Also, those whose high

school's student body was higher in percentage of low-income families sought intervention.

The initial differences across these means indicate a clear need for careful statistical

control to accurately estimate the effect for the treatment groups' performance in the target

Table 7

High School Characteristics

	Mean			Difference of Means			
Variable	Compare	Pre	Con	ΔPre	p	ΔCon	p
Locale: City	25.0	22.6	26.9	-2.4		1.9	
Locale: Suburb	60.1	57.1	5.8	-3.0		-54.3	***
Locale: Town	7.4	11.4	57.7	4.0	**	50.3	***
Locale: Rural	7.4	8.9	9.6	1.5		2.2	
Religious Affiliated	9.7	14.1	9.6	4.4	**	-0.1	
Independent	3.1	1.5	0.0	-1.6	*	-3.1	***
Public	87.2	84.4	90.4	-2.8		3.2	
% Low income in HS	23.1	27.5	32.8	4.4	***	9.7	***

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

course. Only the variables with a significant difference between the Compare and Pre-Course groups and therefore important to control are included in the next steps of analysis. Also, in the interest of maintaining accurate representation of the student population, control variables with missing values for the majority of the students are omitted from further analysis. This list includes: ACT Reading and Science scores, all AP scores, and all high school information

Correlations. To isolate the control variables contributing to any variance in the outcome variables, correlations were evaluated between each of the control variables and the outcome variables (Table 8). The correlation between the Composite exam score and the Total course is r = 0.904 ($p < .001$). The composite exam score includes only performance-based components of the target course, while the total course points earned include components in which students are evaluated on participation. The additional information in the evaluation of total course points gives indication of interaction with the participation components of the course rather than with the performance and knowledge components. Total course points was omitted as an outcome variable from further analysis due to its high correlation with the Composite exam score and the fact that the latter suffices in assessing the effectiveness of the interventions on student performance.

Significant correlations exist in most of the control variables with the performance measures. Parental net business worth was not correlated while investment worth was, confirming that wealth as a proxy for socioeconomic status is related to performance at the college level. The significance across the variables, however, needs to be considered with discernment for magnitude. For example, race/ethnicity is not highly correlated to performance for this particular population of students. Also, scholastic aptitude and physics knowledge scores are not as highly correlated as expected, confirming prior conclusions in literature that influences

49

Table 8

Correlations of Select Control Variables with Performance Measures

Variable	Composite Exam Score	p	Total Course Points	p
Composite exam score	--		0.904	***
ACT-Composite	0.355	***	0.251	***
ACT-English	0.242	***	0.188	***
ACT-Math	0.376	***	0.287	***
Physics Diagnostic Score	0.395	***	0.307	***
Female	-0.131	***	-0.052	*
First generation	-0.094	***	-0.083	***
Asian	-0.085	***	-0.067	**
Black	-0.085	***	-0.065	**
Hispanic	-0.048	*	-0.035	.
White	0.113	***	0.089	***
Years – HS physics	0.144	***	0.098	***
Father's income	0.086	**	0.083	**
Parents' income	0.121	***	0.115	***
Parent business worth	0.033		0.030	
Parent investment worth	0.130	***	0.109	***

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

outside aptitude and prior knowledge can have significant effect on final performance in a college course.

Covariances. Covariances were calculated between each of the control variables (shown in Table 9) to determine which could be eliminated from further analysis, being highly correlated with another variable. For instance, Father's Income was a measure that correlated highly and significantly with each Parent's Income and as such was omitted. Another similar variable was the Composite ACT score, as it correlated highly with the English and Math ACT scores.

Table 9

Covariances Between Select Control Variables

Measure	1	2	p	3	p	4	p	5	p	6	p	7	p	8	p	9	p	10	p	11	p	12	p	13	p	14	p
1. ACT-Composite	--	.76	***	.65	***	.30	***	.02		-.15	***	.03		.14	***	.18	***	.11	***	.00		-.13	***	-.21	***	.16	***
2. ACT-English		--		.34	***	.20	***	.11	***	-.11	***	-.03		.08	**	.13	***	.10	**	-.02		-.10	***	-.15	***	.13	***
3. ACT-Math				--		.34	***	-.07	*	-.12	***	.04		.12	***	.13	***	.06	*	.08	**	-.09	**	-.12	***	.03	
4. Physics Diagnostic						--		-.12	***	-.07	*	.26	***	.05	.	.03		.08	***	.05	*	-.05	.	-.08	*	.00	
5. Female								--		-.05	*	-.06	*	.05	*	.06	*	.01		.06	*	.03		-.03		-.07	*
6. First generation										--		-.04		-.23	***	-.26	***	-.13	***	-.02		.00		.17	***	-.05	*
7. Years HS physics												--		.01		.01		.04		.06	*	-.04	.	.03		-.07	*
8. Father income														--		.80	***	.35	***	-.15	***	.00		-.06	*	.16	***
9. Parent income																--		.44	***	-.15	***	-.01		-.08	*	.16	***
10. Parent investment worth																		--		-.04	.	-.06	*	-.02		.06	***
11. Asian																				--		--		--		--	
12. Black																						--		--		--	
13. Hispanic																								--		--	
14. White																										--	

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$, . $p < 0.20$)

A general picture of the student population has been developed in this section, along with clarifying the necessity to control for initial differences in isolating the treatment effects of the Pre-Course and Con-Course interventions. Additional analyses identified a subset of control variables, which would suffice to control for initial differences among the participant student groups. These control variables, which were included in further analyses, were: (a) Physics Diagnostic score, (b) ACT-English, (c) ACT-Math, (d) years HS physics, (e) Female, (f) First generation, (g) Parent income, (h) Parent investment worth, (i) Asian, (j) Black, (k) Hispanic, and (l) White. Due to these initial differences as well as a non-randomized design, propensity score matching and subsequent analysis on overall performance were used to help isolate the effects of the intervention courses.

Qualitative Analysis Procedure: Generating Profiles for the Intervention Courses

The two intervention courses (Pre- and Con-) differed not only in terms of sequence in relation to the target course, but also in terms of content, objectives, and duration. A thorough description and theoretical contextualization of each of the intervention courses' design and goals was essential and allowed for an understanding of the course content, instructional strategies, and intended outcomes. Doing so enabled a qualitative comparison between the investigated interventions, as well as between these and other interventions directed at improving the performance, persistence, and retention of at-risk students in college physics courses. This endeavor entailed, first, generating a general course description for the interventions, and second, characterizing the theoretical framework(s) underlying the interventions. Toward this end, intervention course materials, including both online and in-class instructional materials such as the weekly prelectures, homework assignments, and discussion problem packets were collated and analyzed. The Curriculum and Pedagogy (C&P) Analysis Protocol presented in Appendix A was used to analyze these course materials. Next, individual semi-structured follow-up interviews were conducted with the course designers/instructors. The individual semi-structured interviews were guided by the Interview Protocol (Appendix B) and aimed to shed light on the principles and theoretical underpinnings considered in the design and application of the curriculum and pedagogy of the interventions, including any references to prior research. The instructor interviews were recorded, transcribed, coded according to the C&P dimensions (outlined at length in Chapter 4) as well as for emergent themes regarding application and embodiment of the instructors' articulated personal instructional theories. These interviews provided an outline for each course's teaching philosophy and how such philosophy was translated into practice through various course components. As such, the emergent themes arose

from reference to details about each component of the courses and how they related to the intended outcomes founded in the teaching philosophies of each course. Analyses of interview data were triangulated with observations and information collected about various course components, which were guided by the C&P analysis dimensions.

Statistical Analysis Procedure

General strategy. Due to initial differences in participant groups' control variables (the Physics Diagnostic score in particular; see Table 2), it was expected that these groups would not perform at the same level. It could be anticipated that comparison students would perform higher than both intervention groups, and that the Pre-Course intervention group would perform at higher levels than the concurrent intervention group. Propensity score matching and subsequent analysis on overall performance were used to help mediate the effects of the initial differences. Propensity score analysis also provided insight into the differential characteristics between the comparison and treatment groups, and helped identify significantly contributing control variables.

An algorithm outlined by Dehejia and Wahba (2002) and exemplified by Murnane and Willett (2011) was used to conduct a controlled comparison of composite exam scores. First, a multiple linear regression model was used to identify which of the control variables contributed significantly to the variance in the composite exam score. The significant variables were then used to create a logistic regression model to explain variance in the choice of enrolling in the treatment courses. This estimate (propensity score ρ) for the likelihood of a student to enroll in the Pre-Course was then created for each participant student. All available variables suspected to affect performance were included in the model, but insignificant contributors to the propensity of a student to enroll in the intervention course were dropped from the model. The estimated

propensity score reduced the dimensionality of the covariates contributing to systematic bias between the two samples down to a single scalar quantity. The refinement of this model is of interest to physics education researchers striving to identify characteristics contributing to a student's choice to enroll in a particular intervention course.

Regression onto target course performance. Those control variables identified in the initial statistical analyses as significantly correlated to the performance outcome variables ($p < 0.20$) were entered into a multiple linear regression onto composite exam score. Setting a rather generous significance level for inclusion in the regression model ($p < 0.20$) was intended to serve as a liberal criterion for variables to qualify for inclusion (Clark & Cundiff, 2011). A second regression was used to ensure that the variables included in later steps of the analyses (i.e., development of a logistic regression model for specific subsets of students) are indeed significant contributors.

Regression onto Pre-Course enrollment. Once significantly contributing control factors were identified, a binary logistic regression with these variables onto treatment enrollment (Intervention = 1, Comparison = 0) was performed, including only students in each particular treatment and the comparison group students. Specifically, the generalized linear model is given by the form

$$\rho = [\exp(\beta_0 + \beta_1 \mathbf{X})] / [1 + \exp(\beta_0 + \beta_1{}^\mathrm{T}\mathbf{X})].$$

where ρ is the probability (propensity or likelihood) of enrolling in the intervention course, $\mathbf{X}$ is the set of control variables contributing to the propensity for enrollment, and β_0 and β_1 are parameter matrices estimated by the statistical program. Hence, the logistic regression fits a logistic function to the relationship between ρ (intervention enrollment) and $\mathbf{X}$ (set of control variables). In linear regression, the model

54

$$\text{logit}(\rho) = \log_e[\rho/(1-\rho)] = \beta_0 + \beta_1{}^T\mathbf{X}.$$

was used to estimate the parameters matrices β_0 and β_1. This step allowed the creation of a composite, reduced-bias variable based on control variables. Since this variable was reduced to a single dimension—a single propensity score rather than a set of control variables—it made possible the comparison of effect differences for values of propensity for intervention enrollment.

Propensity score analysis. Matching groups of students with similar propensity scores allowed for an accurate comparison of students with similar significant characteristics. That is, matching enabled a legitimate comparison between, for example, those students with high propensity estimates who chose to enroll and those who chose not to enroll, given indication for risk according to the significant control variables. Matching also allowed for investigating effect across different student blocks to identify students who benefited most from intervention courses.

Creating propensity scores. Once the control variables were taken into account through the final binary logistic regression, an estimated propensity score $(0 < \rho < 1)$ was calculated for each student.

Stratifying blocks. Once calculated, there are three methods for matching students on propensity scores. These methods are: exact matching, closest neighbor matching, and blocking. Choice of matching method depends on the number of participants in the groups being compared. In this study, block matching was used. Bias due to the non-randomized study design and initial differences between student groups were reduced by blocking propensity scores for students in the treatment and comparison groups with similar scores, then comparing the course performance for these matched students. The frequency distribution and average propensity score

55

values were closely examined for each block to ensure equality.

Checking equality across variables in regression equation. For the method of matching propensity scores to adequately allow for a reduced-bias comparison for treatment effects, the values for the control variables should be matched within each block. To ensure full matching within blocks, the variables contributing significantly to variance were calculated and compared. Values of the control variables will be examined to ensure equality for the blocked students.

Estimating treatment effect. Once calculated, composite exam score (ExamComp) was used to determine the performance differences among the intervention and comparison groups. Due to inherent initial differences in the control variables— the comparison students outperforming the treatment students from the start on the diagnostic exam and exhibiting significant differences on several other variables—it was expected that these groups would not perform at the same level, with the comparison students performing highest.

Persistence. Studying the subsequent paths with regard to their likelihood to re-enroll and ultimately complete the target course for Drop/Withdrawal/Fail group (DWF) students provided an avenue for investigating differential effects on students at higher risk than others. As persistence is a critical issue in performance and long-term retention of at-risk students in college physics, this matter was investigated to some degree in this study.

Persistence can be estimated on a broad level by comparing the subsequent performance and passing rates of initially failing students among the intervention and target courses. Those students were isolated, and their performance means and passing rates were compared for significant differences.

Another comparison was made regarding failure and drop rates for each group. This was of interest because of the importance of students achieving a C or higher in the target course for

making adequate progress in all engineering major curricula. Students included in the DWF group fell in one of two categories: those who completed the course but did not achieve a C or higher for their final grade, and those who dropped or withdrew from the course after putting full effort into completing course work up to the point they decided to drop or withdraw. Some of these students complete none, a few, or all of the regular exams of the course, but not the final exam. Once the Composite exam scores had been compared to their comparison counterparts within each propensity score bin, the students were further divided into Pass and DWF. The average Composite Exam scores for DWF intervention students were compared to the averages for DWF comparison students, with statistical significance determined by a difference of means *t*-test. This approach was used to give some indication whether the interventions had any effect on mitigating the degree of risk for DWF for students who had equal propensities and control scores otherwise. Flat rates of number of Pass and DWF students could also easily be generated with this step, allowing the comparison of the ratios of students with similar propensities and other scores and thus a simple way to gauge the effect of the interventions across the different propensities.

Summary

Using rigorous statistical analysis procedures, the differential effect as conceptualized in the research questions allowed a reduced-bias comparison of the impact of the two interventions on at-risk students' performance in the target course. The post-test only aspect of the design presented a limitation with regard to controlling for prior knowledge. This was overcome by using a collection of control variables related to students' prior knowledge and academic abilities. Including these variables in the statistical analysis also allowed for the identification of the significant variables that characterized those students identifying themselves as at-risk in the

present study, as indicated by their choice to enroll in an intervention course. The self-selection effects inherent in the study were reduced by directly accounting for students' voluntary decision to enroll in an intervention course through the proposed propensity score. A student's likelihood to choose to pursue help was hypothesized and calculated by including academic and family background factors shown to be characteristic of students for risk of failure or withdrawal from introductory physics.

The careful statistical treatment and generation of the effects on performance therefore provided valid grounds for their interpretation within the study's context. The steps in establishing context in which to interpret results have also been outlined. The C&P Analysis Protocol and the interviews with the course designers/instructors allowed the identification of prominent and salient features of the two intervention courses.

CHAPTER 4

RESULTS

This chapter first describes the intervention courses and outlines their theoretical underpinnings. The descriptions shed light on curricular and pedagogical aspects of the courses. The second section presents the results of the statistical analyses and associated claims about the relative impacts of the two intervention courses.

Intervention Course Profiles

The overlay of the previously developed theoretical framework (see Chapter 2) served to formalize the characteristics and components of the intervention courses in a larger context of teaching and learning theory. These analyses provided a basis for qualitative comparisons of the two courses, as well as comparisons with other interventions described in the literature review. Caution was exercised when making conclusions about courses because these were based mostly on self-report interview data provided by course instructors as compared to data related to observed course effects.

Aspects of curriculum and pedagogy analysis. An instructor's epistemological stance affects his or her approach to course design, curriculum, and pedagogy (see Chapter 2). The following sections explore the courses' instructional philosophy and approach; objectives and curriculum; learning environment; expected student engagement behaviors; and anticipated instructor and student interactions as derived from interview data and further analyses of associated course artifacts, including instructional materials, assignments, and assessments.

Objectives and curriculum. The curriculum entails choice of course content. This includes explicitly stated goals and objectives in the course description, as well as implicitly required ones. Course content also often assumes certain prerequisite knowledge and skills (such

as vector addition and decomposition) necessary for mastery of the concepts (such as Newton's Second Law) that are not explicitly stated in the course description. Learning objectives are defined as the knowledge and skills to be assessed. Cognitive skills include information processing skills beyond content (e.g., proportional reasoning, symbolic representation, problem solving). Data collected to provide evidence included the course syllabus, course description, online assignments, and in-class assignments.

Instructional materials and assignments. The learning materials developed for students provide crucial information about knowledge targeted in a course, how the instructor intends for this knowledge to be learned, and the degree of consistency between learning objectives and instructional approach.

Assessments. The explicit summative assessments designed by the instructor convey a reinforcing message about the intended instructional approach and learning outcomes as motivated by an instructor's epistemological stance. These assessments indicate how a teacher evidences and evaluates student achievement of the target knowledge and capabilities.

Learning environment. Interviews with course designers/instructors, as well as investigation of the classroom provided information about the placement and location of student(s) and teacher in relation to each other and the resources made available to aid learning. Objectives and outcomes for students' online learning experiences were also inferred from the developer/instructor interviews.

Student engagement behaviors. This dimension expands on the intended design of the course components with regard to what students are responsible for, what they need to learn, and how they are expected to facilitate and/or monitor their own learning processes. General instructions and details of expected student engagement behaviors not outlined at length in the

course description were asked of the instructors during interviews and inferred from course materials.

Instructional philosophy and approach. Conceptions of instructor contribution to learning in and out of the classroom, as well as how and what actions are taken to facilitate the learning process can be gleaned from the nature and extent of teacher participation in the learning environment, including their role in challenging and aiding student learning. Toward examining these dimensions, the electronically-provided instructional materials were investigated since a major portion of the students' learning experiences in the Pre-Course are enacted through media and activities accessed asynchronously and independently online. Having expanded on the dimensions for analysis for each course's curriculum and pedagogy, the next two sections present the findings of the analysis for each of the intervention courses.

Pre-Course intervention. The Pre-Course, named "Thinking About Physics," is divided into two eight-week components. The material covered during the first eight weeks is the same material covered on the first exam in the target course. The difference is that students in the Pre-Course spend twice as much time on these materials compared to the target course. Other teaching objectives for the course center on conceptual and problem solving skills, including algebraic and trigonometric skills, as well as abstract reasoning skills, analysis and mathematical descriptions of physical situations, and understanding the meaning of solutions. These skills are assessed through the Physics Diagnostic, which indicates that students enrolling in this course have not developed fully the necessary physics and math knowledge for the target course.

Objectives and curriculum. The main objective of the course is to prepare students for the target and subsequent physics courses by teaching the problem solving skills and physical reasoning emphasized in these courses. This intervention course covers and integrates the topics

and concepts of one- and two-dimensional kinematics, the definition and types of forces (including simple external forces, such as an applied force, tension, gravity, spring force, and Universal Law of Gravitation), Newton's Laws, and centripetal acceleration. Understanding trigonometry, geometry, and vector decomposition and addition, as well as efficiently and effectively applying these ideas to solve word problems are key for achieving mastery of these initial topics in the course. As such, homework assignments and discussion materials are designed to challenge students and help them develop these skills. Doing so minimizes distraction from learning the physics concepts entailed in the explicit content of the target course.

The Pre-Course features a midterm exam, which identifies areas of needed focus. The topics are repeated in the second eight weeks of the course with emphasis on more fully developing reasoning processes, specifically solution strategy-building and problem solving skills. Course activities are intended to advance students to proficiently turn word problems into mathematical expressions through concept application, solve the math with applied trigonometry and algebra, and understand and verbalize the physical significance of their results.

Instructional materials and assignments. The Pre-Course follows a weekly cycle of online assignments and in-class activities. In order, these are a web-based prelecture, a web-based checkpoint, an interactive classroom lecture, a web-based homework assignment, a discussion section, and finally a short web-based quiz.

The Pre-Course utilizes secure online multimedia to present the prelecture physics content. Animated slides with voice-over by the course developer are designed to direct student attention as the concepts and examples are presented. Equations, diagrams, and graphs are frequently used. The script for the voice-over serves to verbalize not only that which the students are able to see but also to articulate the conceptual complexities or difficulties in a similar way a

62

teacher would in a traditional lecture. Students must answer summative assessment questions at a level that indicates conceptual comprehension in order to advance through the presentation to completion. Presenting the key concepts independently prior to attending the lecture replaces assignments to read a physics textbook out of class, highlights important information to students, and provides students with feedback on their initial understanding of the material. Students typically spend about 30 to 45 minutes each week with this online activity. The students then complete a short, individually scored, online formative assessment called a checkpoint. The instructor examines answers to gather information about student comprehension and common difficulties to be specifically addressed in the classroom lecture (as such, lectures change from semester to semester depending on student responses). Students answer conceptual and problem-solving questions verbally and also via an electronic classroom response system (CRS).

Following the prelecture, checkpoint, and lectures, students engage in a weekly web-based homework assignment consisting of 5 to 7 problems. The problems test understanding of the week's concepts and associated skills necessary to exhibit mastery of the concepts and their application to solving word problems. An unlimited number of attempts are permitted and full credit is awarded for each individual problem once the correct solution has been submitted. Students solve straightforward traditional pencil-and-paper problems and others accompanied by an extended help sequence of questions including conceptual, strategic, and numerical web-based assistance for the problem. Completion and correctness rates for this type of homework range at above ninety percent for all students.

Next, the students work in collaborative groups on qualitative and quantitative problems in the weekly two-hour discussion section. Typically, a teaching assistant presents a brief review of relevant concepts and common difficulties. Students then proceed to discuss the problem

packet for the week. The problems are conceptually cumulative in nature and designed to elicit common misunderstandings and difficulties in the concepts and their application. The design of the short online weekly quiz permits students to access the quiz after discussion and amend answers until the deadline when final answers are saved and scored.

Assessments. The quizzes are given as an online multiple-choice assessment. The problems include not only quantitative but also qualitative questions. In the style of midterm exam questions, the focus is on application and complex reasoning with the physics concepts. The students have a two-day period in which to complete each quiz.

The midterm exam, administered half-way through the semester, is exactly the same format and context of the exams in the target course. It is an hour and a half in length and is approximately twenty-five questions long. The midterm gives students an opportunity to receive high-stakes feedback on their progress. The final exam is given in the same format as the midterm. A multiple-choice format allows ease of grading and consistency in awarding points and in administering assessments to a large number of students. The multiple choice questions can have 2-, -3, and 5-option questions, weighted at 2, 3, and 6 points, respectively, all of which could be quantitative or qualitative in nature. Partial credit can be earned on the six-point questions by marking the correct answer in the selection of up to three options (that is, no points are earned if one of their selections is not the correct option). Thus, these assessments divert somewhat from typical equal-weight, right/wrong, fixed-number-of-distractors, multiple-choice items. The exams assess concept attainment, conceptual understanding, contextual concept application, and complex conceptual reasoning, as well as reasoning applied toward problem solving (i.e., constructs related to intended outcomes of intervention outlined in the curriculum and objectives above).

The final grade in the Pre-Course is based upon a sum of points earned out of 1000 maximum points. Table 10 summarizes the percentage each course component contributes to the final score. Partial credit of 80% of the original point value is given for any correct answers submitted in the week after homework assignments are due, allowing students to pursue help on particularly difficult problems within a limited time frame. Discussion participation points serve to deter tardiness and encourage engagement, and typically students lose points only if they are late or present significant resistance to working in their group. Full credit is given for completing each checkpoint regardless of correctness. Full credit also is given for answering 75% of the

Table 10

Component Contributions to Pre-Course Final Grade

Course Component	Number of Assignments	Final Grade Percentage
Web-based Prelecture	8	4%
Web-based Homework	8	15%
Web-based Checkpoints	11 (12)	5.5%
Lecture Participation (CSR)	11 (14)	5.5%
Discussion Participation	10 (13)	20%
Web-based Quiz	8	15%
Midterm Exam	1	15%
Final Exam	1	20%

CRS questions. Bonus points can be earned for correct CRS questions during lecture. These points are added to the online quiz score, which not only provides rewards for understanding, but also provides incentive for participating fully in lectures and discussions focused on CRS questions. No excused grades are accepted for any of these course components. Letter grades are assigned based on the number of points out of the 1000 possible points. The letter grade lines

were described in the interview as "generous," thus allowing students a low-risk grading environment in which to learn the material. Most students earn an A or B for the two-hours of course credit.

Learning environment. Students and instructors meet face-to-face in lecture once a week and in the discussion section. Students interact verbally by articulating their ideas to one another for CRS questions during lecture. Additionally, students work closely together to develop strategy and help each other solve the problems in the discussion packet each week. The rest of the course is conducted online through a web browser, limiting interactions among the students and with the instructor. The introduction of the concepts in the prelecture and the extent to which students work on their own, however, enhances the quality of the in-class interactions.

Lecture provides a place for the instructor to initiate conversations with the students, either during the regular presentation of the material that has been prepared or in discussions of CRS questions.

Student engagement behaviors. Students are made aware of the different components and requirements to complete credits and earn points at the outset of the course. For the main components of the course (prelecture, checkpoint, homework, online quiz), students determine the time and place where they get to access a web browser and complete the assignments. The length of time they spend is determined by the degree of effort and thought invested into an activity by the time they complete it. Students are not required to complete the online activities all at once (i.e., they may quit in the middle and access them again at another time), providing space for students to procure additional forms of help as they see fit or progress at a pace that accommodates their mental processing. They may look back at the prelecture and also do homework with other people. The extended-help homework problems allow students to traverse

through a deliberately structured problem-solving process.

Learning, application, and review of the material can be presumed to be an isolated task for the first part of the learning process in interaction with the prelecture and checkpoint, after which students attend the lecture. Lecture gives students an opportunity to interact with each other to articulate their ideas when presented with questions to answer via the electronic CRS. Further collaborative interactions are required while working on the problems in the discussion session each week.

Instructional philosophy and approach. The foundational element of the instructional approach for the Pre-Course is the essential requirement of students' active engagement in their own learning process. In general, the instructor's responsibilities, regardless of the material being taught, include: presenting a coherent picture/understanding of the material covered; providing examples for concept application; identifying difficulties that are pre-existent and often typical to the particular concepts being covered; assigning meaningful work that relates to prior knowledge and that serves to expand understanding; assessing understanding and providing feedback in the service of further learning; and providing space and time for students' questions in lecture, in collaborative learning, or during office hours.

The types of help provided in the course in regard to interpersonal interactions stem from the articulations exemplified by the instructor in the prelecture and lecture. Conceptual application to solve word problems necessitates a coherent understanding with physical meaning of the developed equations, which can be articulated to the instructor and other students. As such, the instructor, teaching assistants, and students are encouraged to engage in a Socratic dialogue. Expert reasoning processes and conceptual understanding are modeled by instructors and advanced students. Quick and rote answers are not given. Instead responses to student

questions entail having students articulate and explore their current understanding of the topic.

A broad glance across the course components shows that students generally determine when, where, and how long to engage with the online materials (prelectures, checkpoint, homework, and quiz). This approach provides students with individual flexibility in regard to the pace and timing of learning the material within the time frame set by the instructor, and within a structured online learning environment. The prelecture and checkpoint components ask students to answer conceptually rich multiple choice questions in order to complete the assignments for credit. In addition, the checkpoint requires students to articulate their conceptions to the extent they come to as a result of the prelecture. Actual student responses to checkpoint questions are presented anonymously and addressed in lecture, such that students receive thoughtful instructor feedback as to the incompleteness or incoherence of their conceptions, and about any misconceptions they have internalized. Lecture and discussion are structured as interactive learning environments in which students are presumed to bring along their own knowledge base and understanding about the material. As such, they are asked to explain their ideas in their own words and answer questions, which elicit either misunderstandings or incoherent or incomplete conceptions about the physics topics being covered, or any difficulties in associated reasoning process. Thus, it is assumed that progress in learning results from student receiving feedback from the instructor or collaborative group members, and from reflecting on their own understandings.

Summary. The design and chronology of course components in the Pre-Course embody an instructional philosophy and approach founded in the necessity for students to actively engage in their own learning. Conceptual understanding and its articulation are emphasized throughout the course, particularly in interactions and assessments. Students are afforded individualization

in their learning, as the web-based instructional materials can be accessed (and re-accessed) as much as the students deem necessary.

Concurrent course intervention. The Con-Course assists students with development and practice of concept application and problem solving strategy. The course is offered for one hour of credit, the students are required to attend a 2-hour discussion section per week, and there is a short weekly homework assignment. Due to the design of the course as a weekly supplement to the target course, its profile is considerably less involved than that for the Pre-Course.

Objectives and curriculum. The content of the Con-Course complements the content covered in the target course, working at the same pace and covering the same topics each week. The course is designed to help students translate the concepts learned in the target course into problem-solving strategies and techniques. Special attention is given to the process developing logical, coherent, and consistent problem-solving strategies. In particular, the conceptual understanding of physical and symbolic relationships is extended in what is essentially the students' second discussion section for the target course.

The focus of in-class time is the application of a logical problem solving process to create a solution map for concept-problems. The five-step analysis process explicitly outlines and serves to structure student thinking, and includes literary analysis, conceptual analysis, strategic analysis, quantitative analysis, and reflective analysis. The literary analysis breaks down the aspects to focus on for problem interpretation and reading comprehension of the problem text. Students are asked to create a mental structure for the problem by identifying given quantities, deducing implications about the given and unknown quantities, and creating a sketch of the physical situation of the problem. This step helps to develop an understanding of the physical parameters of the problem, as well as utilizing multiple representations (literary, symbolic, and

pictorial) of the situation. The next step of the analysis is conceptual, where students identify the conceptual class of problem being solved (e.g., two-dimensional kinematics or Newton's Second Law with static friction). Next is the strategic analysis, wherein students link together the given parameters of the problem with their physical representation, assign symbols of the knowns and unknowns for the problem, and develop equations relating these aspects based on the relevant concepts identified in the previous step. At this point in the problem-solving process, all of the actions needed in to find the final answer have been outlined and the quantitative analysis to follow is pure execution of their deliberately planned strategy (solution map). Mindfully engaging in these prior mental activities provides the basis for the last step, which is reflective analysis. Students go over each of their previously developed steps to evaluate for effectiveness and completeness, that is, to meaningfully reflect on the abstractions of the cognitive processes they just engaged to solve for the final answer. This step emphasizes that obtaining a numerical answer is not the last step or final destination in a problem solving process where deep and long-term conceptual understanding is the ultimate desired result.

Instructional materials and assignments. The content is delivered in problem packets that students work on in groups of four in a discussion section similar to the ones in the Pre-Course. Students are presented with concept-rich problems constructed by the course instructor or taken from practice midterm exams for the target course. The problems are not identical to those seen in the regular target-course discussion section but are similar to the type of mastery expected on the target course exams. These problems are divided into two separate types, which are tackled by students during the two-hour sessions. The first requires students to focus exclusively on the development of the problem solution map, which comprise the first three steps of the problem solving process described above. The instructor reports that students work in

70

close collaboration, and are assisted by teaching assistants, during this process. The same collaborative approach is used to address the second type of problems, where students are allowed to work through the problem solution fully, which sometimes entails the completion of a problem for which a solution map was already created in the first type of problems in the packet. A homework assignment is given in the last fifteen minutes of class, consisting of problems of similar difficulty and length as those covered during the discussion session. Students are allowed time in class to work on the homework in the spirit of minimizing outside work for a simple one-hour credit class, yet a certain degree of involvement is required, and as such, the course grade has a slight performance component. Students are allowed to turn the assignment in on the same day or at the beginning of the following week's class.

Assessments. No summative assessments exist in the Con-Course. Points for the final course grade are earned through: attendance, sticking to the outlined problem solving analysis steps, and compliant engagement with the teaching assistant and other students during the discussion section (50%); completing the strategy-development-problems satisfactorily during discussion (30%); and homework grades (20%). As students work through the strategy-development-only problems each week, the teaching assistant monitors their progress and ensures their adherence to the problem analysis procedure. Students earn credit by showing and explaining their work to the teaching assistant before moving forward in the problem packet. Credit is earned on the homework assignment largely through completion of the first three steps of the suggested problem analysis procedure. The majority of the students, thus, earn an A in the course.

Learning environment and student engagement behaviors. The type of help provided by the instructor centers around keeping the students focused on the problem solving analysis

procedure and scaffolding them within each of those steps. The analysis steps are presented in the course information package and discussed at length with examples of desired and undesired behavior and with justifications and implications for each step. As such, the students are not asked to do something without reason or foresight. These steps are then written on the board each week as a reminder. A teaching assistant frequents groups as they work through the problems on their own, helping to highlight guiding concepts relevant to the problems and also structure the development of problem solving strategies by reminding the students of the particular pieces of each step of the analysis process. The work is highly collaborative and students are expected to be vocal contributors throughout the analysis process. The teaching assistants are discouraged from being overly directive toward problem solutions, but instead to respond in ways that lead students through the analysis process such that the students themselves develop the solution strategies. In the spirit of Socratic dialogue, this help is in the form of hints or questions that direct student thinking toward key realizations and relationships necessary to solve the problems.

Instructional philosophy and approach. From the outset, the course is represented to be about problem solving. An inherent epistemology is imparted regarding what it means to "do physics." This idea entails not simply understanding the concepts on a deep level and as they relate to the problems to be solved, but also to plan for and execute the cognitive processes necessary to achieve mastery of the physics. The instructional approach adopted is akin to the structured, step-by-step development of expert cognitive activities by providing the definition of the relevant cognitive structures and reasons for their utility in a specified procedural method, followed by engaging students with the repeated practice of these processes. As such, attention is initiated and sustained in relation to the outlined problem-solving process as it scaffolds students

72

to develop a consistent, coherent, and deliberate approach to solving physics problems. By outlining this analysis for the students, the instructor provides a clear delineation of desired actions and aspects to concentrate on in an attempt to prevent or deter students from making a premature pitfall or jump to superficial equation manipulation.

Summary of course profiles. The preceding analysis provided a fuller picture of the content and instructional strategies used in the two intervention courses. In addition to a difference in chronology with respect to the target course, the scope of the interventions differ in regard to curriculum. While the Pre-Course encapsulates parts of the curriculum in the target course, including the physics content and relevant pre-requisite skills and abilities, the Con-Course intervention's focuses on the metacognitive component of the problem-solving process. Both utilize an interactive, collaborative approach to learning and assessing conceptually rich material.

Statistical Results: Relative Impacts of the Intervention Courses

Regression onto target course performance. Variables included in this analysis were Physics Diagnostic score, ACT-English, ACT-Math, Sex (Female), First Generation, Race (Asian, Black, Hispanic, White), Years of HS Physics, Parent Income, and Parent Investment Worth. Composite exam score was used to determine the performance differences among the intervention and comparison groups. A stepwise multiple linear regression was conducted in the forward manner, including each variable correlating significantly to composite exam performance, beginning with Physics Diagnostic score (Table 11).

The stepwise multiple linear regression of the control variables onto Composite Exam Performance in the target course shows that academic aptitude and prior education, as represented by the Physics Diagnostic score, both ACT Math and English scores, and years of

Table 11

Summary Table for Stepwise Multiple Linear Regressions onto Course Performance

	Model A			Model B			Model C			Model D		
	Estimate	*p*	Error	Estimate	*p*	Error	Estimate	*p*	Error	Estimate	*p*	Error
Intercept	53.39	***	1.46	19.05	***	3.25	11.70	***	3.50	10.9	**	3.46
Pre-Course	2.24	**	0.77	2.23	**	0.74	1.95	**	0.74	1.93	**	0.73
Con-Course	5.21	**	1.60	5.73	***	1.54	5.51	***	1.52	5.83	***	1.51
Physics Diagnostic	2.24	***	0.14	1.78	***	0.14	1.70	***	0.14	1.72	***	0.14
ACT-Math				1.21	***	0.10	1.08	***	0.11	1.20	***	0.11
ACT-English							0.42	***	0.08	0.34	***	0.08
Asian										-4.31	***	0.69
R^2	0.1642			0.2308			0.2448			0.2628		
$F(df_1, df_2)$	$F(3,1583) = 103.6$			$F(4,1582) = 118.6$			$F(5,1581) = 102.5$			$F(6,1580) = 93.87$		

	Model E			Model F			Model G			Model H		
	Estimate	*p*	Error	Estimate	*p*	Error	Estimate	*p*	Error	Estimate	*p*	Error
Intercept	7.44	*	3.62	8.63	*	3.86	8.98	*	3.83	9.57	*	3.90
Pre-Course	1.80	*	0.74	2.05	*	0.80	2.20	**	0.80	2.20	**	0.80
Con-Course	5.17	***	1.54	5.82	***	1.66	5.86	***	1.64	5.81	***	1.64
Physics Diagnostic	1.61	***	0.14	1.69	***	0.16	1.63	***	0.16	1.63	***	0.16
ACT-Math	1.21	***	0.11	1.11	***	0.12	1.08	***	0.12	1.07	***	0.12
ACT-English	0.40	***	0.08	0.43	***	.09	0.49	***	0.09	0.48	***	0.09
Asian	-4.76	***	0.71	-4.56	***	0.78	-4.32	***	0.77	-4.35	***	0.77
Years HS physics	1.82	***	0.55	1.62	**	0.59	1.63	**	0.59	1.62	**	0.59
Par invest worth				4.46E-6	**	1.52E-6	4.53E-6	**	1.50E-6	4.38	**	1.52E-6
Female							-3.60	***	0.77	-3.63	***	0.77
First Gen										-0.62		0.76
R^2	0.2731			0.2860			0.2984			0.2988		
$F(df_1, df_2)$	$F(7,1504) = 80.73$			$F(8,1242) = 62.2$			$F(9,1241) = 58.66$			$F(10,1240) = 52.84$		

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$)

high school physics courses contribute significantly to the variance observed in performance. Other control variables contributing include students' status as Asian, Female, or First Generation in higher education, as well as the amount of each student's accumulated wealth as represented by parents' investment worth.

Multiple linear regression Model H ($R^2 = 0.2984$, $F(10, 1240) = 43.59$, $p < 0.001$), whose parameters are outlined in Table 12, was developed as a result of the stepwise regression. According to Model H, students in the Pre-Course gained a 2.20 ± 0.80 point advantage and those in the Con-Course gained a 5.86 ± 1.64 point advantage in controlled comparison to the non-intervention students. For the 1591 participants, the Composite Exam score average was 74.1 with a standard deviation (SD) of 12.7 points, yielding effect sizes of 0.17 SD and 0.46 SD for the Pre-Course and Con-Course, respectively. The regression onto the target course composite exam performance thus indicates that the Pre-Course plus Con-Course intervention caused a larger observed performance gain over the comparison students than the Pre-Course intervention alone. The variables First generation, Parents' income, and Black were insignificant contributors to the model and were omitted from further analysis.

Having isolated the significant variables for the overall population, a model including the remaining variables was generated for the Pre-Course plus Comparison groups, which provided another estimate of the treatment effect on the participants in the Pre-Course group. This subsequent regression (see Table 12, Model I) was conducted with the relevant control variables, namely, controlling for inclusion in the Pre-Course only. Model I ($R^2 = 0.3020$, $F(8,1192) = 64.5$, $p = < 0.001$), indicates that the Pre-Course students gained an overall 2.28 ± 0.80 point advantage compared to non-intervention students, which aligns with the previous first broad estimate of 2.20 ± 0.80 from Model H. A similar analysis was conducted for the Con-Course that

resulted in Model J (R^2= 0.3118, $F(8,830)$ = 47.0, p = < 0.001). Model J shows that the Con-Course students gained 6.09 ± 1.60 points compared the non-intervention students, which also aligns with the previous first estimate of 5.86 ± 1.64 from Model H.

Table 12

Summary Table for Multiple Linear Regressions onto Course Performance

	Model H			Model I			Model J		
	Estimate	p	Error	Estimate	p	Error	Estimate	p	Error
Intercept	8.98	*	3.83	8.12	*	3.91	10.7	*	4.43
Pre-Course	2.20	**	0.80	2.28	**	0.80	--	--	--
Con-Course	5.86	***	1.64	--	--	--	6.10	***	1.60
Physics Diagnostic	1.63	***	0.16	1.64	***	0.16	1.74	***	0.17
ACT-Math	1.08	***	0.12	1.09	***	0.12	0.97	***	0.14
ACT-English	0.49	***	0.09	0.49	***	0.09	0.49	***	0.10
Asian	-4.32	***	0.77	-4.31	***	0.79	-3.22	***	0.88
Years HS physics	1.63	**	0.59	1.77	**	0.59	1.77	**	0.65
Par invest worth	4.53E-6	**	1.50E-6	4.58E-6	**	1.53E-6	5.66E-6	**	1.73E-6
Female	-3.60	***	0.77	-3.56	***	0.79	-3.13	**	0.95
R^2	0.2984			0.3020			0.3118		
$F(df_1,df_2)$	$F(9,1241)$ = 58.66			$F(8,1192)$ = 64.48			$F(8,830)$ = 47.0		

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$)

Regression onto Pre-Course enrollment. A binomial logistic regression was conducted using the Generalized Linear Model fit regression tool in the statistical analysis package. The model predicts propensity score based on the included control variables; that is, the model predicts the enrollment into the Pre-Course intervention for a given set of control variables. The first model generated (R^2 = 0.3225, $F(7, 1193)$ = 81.1, $p < 0.0001$), not outlined here, indicated

that although Asian, ACT-M, and Parents' Net Investment Worth contribute significantly to the

variance seen in Composite Exam performance, they do not in the propensity for intervention

course enrollment. A similar stepwise regression was conducted to confirm the significance of

each variable to explaining the amount of variance seen in students' choice in enrolling in the

Pre-Course intervention. Model 4 ($R^2 = 0.3151$, $F(4, 1449) = 166.7$, $p < 0.0001$) (Table 13) is an

attempt for parsimony. Interaction terms between ACT-English and Female and between ACT-

English and Physics Diagnostic were not included because the loss of degrees of freedom were

Table 13

Summary of Stepwise Multiple Linear Regressions onto Pre-Course Enrollment

	Model 1			Model 2			Model 3			Model 4		
	Estimate	p	Error	Estimate	p	Error	Estimate	p	Error	Estimate	p	Error
Intercept	5.03	***	0.31	5.63	***	0.35	5.52	***	0.36	4.62	***	0.64
Physics Diagnostic	-0.68	***	0.04	-0.65	***	0.04	-0.65	***	0.04	-0.66	***	0.04
Years HS physics				-0.60	***	0.14	-0.61	***	0.14	-0.58	***	0.14
Female							0.38	*	0.17	0.33	.	0.18
ACT-English										0.03	.	0.02
R^2	0.3070			0.3110			0.3136			0.3151		

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$)

not justified in exchange for the little additional variance these interaction terms accounted for in

the model included only Physics Diagnostic score, ACT-English score, Female, and years of

high school physics.

Propensity score analysis. The propensity score analysis consisted of four steps,

described individually here.

Creation of scores. To calculate the predicted propensity once the analysis on the original data was complete, the statistical program used (R) evaluated the equation

$$\rho = [\exp(\beta_0 + \beta_1 \mathbf{X})] / [1 + \exp(\beta_0 + \beta_1{}^T\mathbf{X})]$$

for each individual student in the course. That is, it predicted the likelihood the student would enroll in the intervention course considering the student's given set of control variables.

Block stratification. Stratifying in blocks was sufficient to create equivalent groups given the large number of participating students. The students were matched by dividing them into five propensity score blocks, as recommended by Rosenbaum and Rubin (1983), as this typically reduces bias enough for adequate comparison. The frequency distributions for the two groups (Table 14) indicate a large number of participants with high propensities belonging in the Pre-Course treatment group.

Conceptually, the five different blocks can be thought of as five different levels of risk for failing the course, with the highest propensity score corresponding to the most at-risk students. The frequency distribution of students per block is shown in Figure 2 along with the average propensity score per block, indicating that the majority of the students with low propensities did not enroll in the Pre-Course intervention. Alternatively, a number of students with high propensities did not choose to take the intervention course, thus providing a matched comparison group for the Pre-Course students. Students in the third propensity score block earned Physics Diagnostic scores that put them in the area of being cleared to take the Pre-Course but not recommended (Figure 3). The number of comparison students in this block can be inferred to be high due to this difference. Further analysis below helps to flush out differences between these two groups. First, the average propensity score for each block is significantly different from the neighboring blocks. Second, propensity scores within each block are

78

Table 14

Frequency and Average Value of Propensity score and Physics Diagnostic score per block

		Compare		Pre-Course			Physics Diagnostic		
	Block	N	Mean	N	Mean	p	Comp	Pre	p
1	$0 < \rho < 0.2$	663	0.084	28	0.153	***	11.2	9.5	***
2	$0.2 < \rho < 0.4$	133	0.292	101	0.302		8.8	8.7	
3	$0.4 < \rho < 0.6$	93	0.472	124	0.505	***	7.6	7.4	*
4	$0.6 < \rho < 0.8$	47	0.693	120	0.696		6.2	6.4	*
5	$0.8 < \rho < 1$	39	0.881	106	0.878		4.3	4.6	

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$)
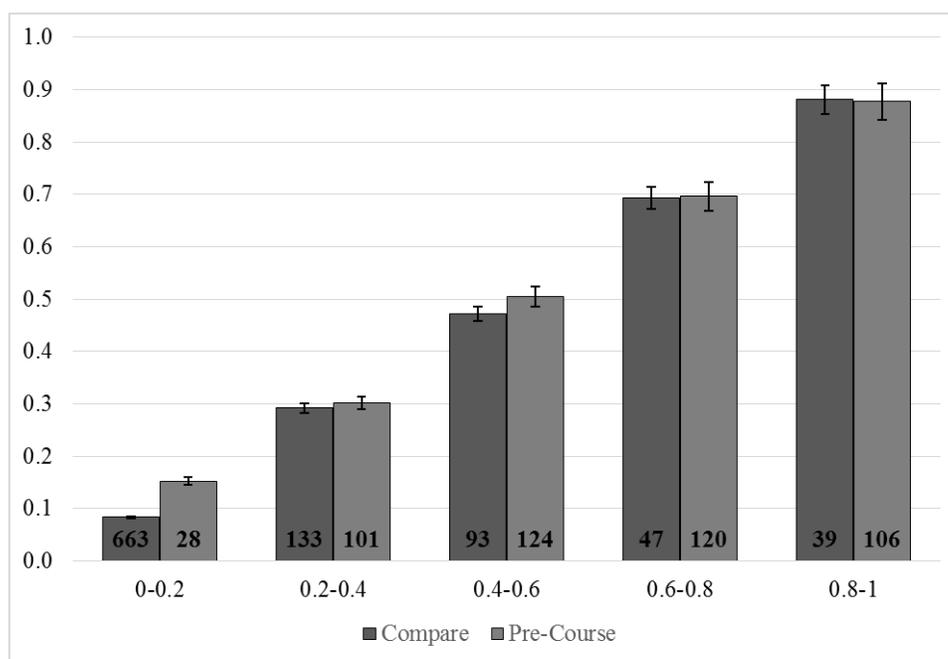
Figure 2: Propensity score and number of students per block
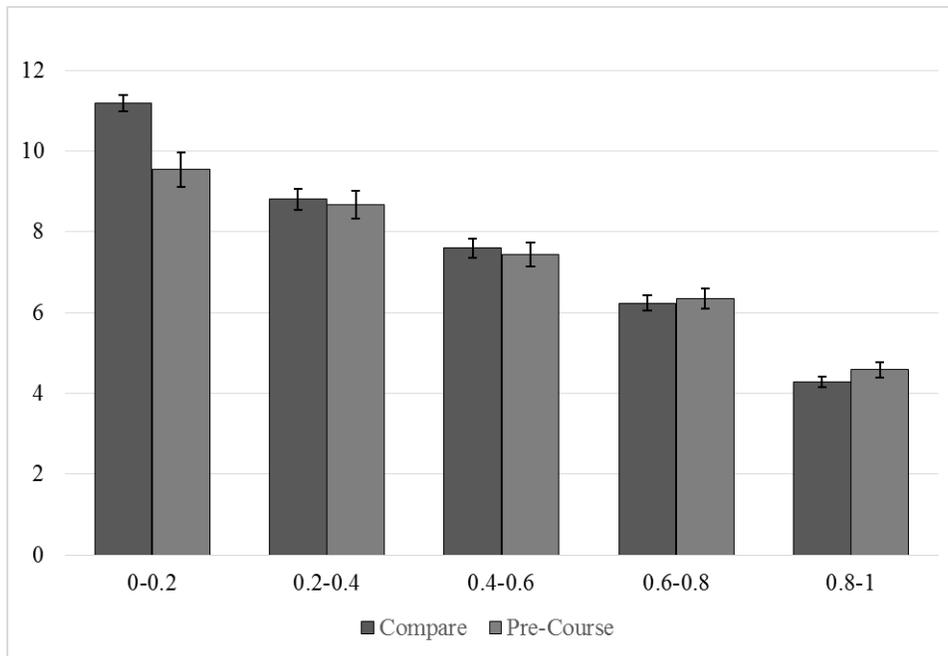
Figure 3: Physics Diagnostic score per block

essentially statistically equal except for the first $(0 < \rho < 0.2)$ and third block $(0.4 < \rho < 0.6)$, as shown by Figure 2. The consideration for moving forward with the difference in the first block was that students of low risk and the effect of the Pre-Course on their performance were not of interest for the study. That is, the study focused instead on students (who perceive themselves) at high risk for failure or withdrawal from the target course and who, thus, have a high likelihood for enrolling in the Pre-Course. Additionally, the propensity scores in the third block do not match exactly but the values are significantly different compared to the two neighboring blocks.

Check equality across variables in regression equation. To investigate matching within-blocks on the variables contributing significantly to variance, the means for each block for Physics Diagnostic, years of high school physics, ACT-English and composition of females were calculated and compared. This information is reported in Tables 14 and 15 and represented visually in Figures 3-7. In further analysis, the first block continued to indicate significant

differences in the control variables for the small number of students who enrolled in the Pre-Course treatment. However, the focus was not placed on these differences in exchange for the satisfactory comparison in the two blocks of highest average propensity scores. The average value for these control variables in each block for the two groups failed to match in the first block and this was expected due to the propensity scores for the groups in this block also not matching well. The Pre-Course students in this block, although scoring significantly lower on the Physics Diagnostic, had taken about a quarter of a year more physics in high school than their non-intervention peers. Although insignificantly different than their peers, the Pre-Course students in this block also scored lower on the ACT-E and ACT-M tests. Due to the observed differences between the two groups in the first block, it was omitted from further analysis.

While the values of Physics Diagnostic scores matched well within the remaining four blocks, differences occurred within these blocks for the other variables included in the propensity score logistical regression. Specifically, the average number of completed years of physics was greater for the comparison group by 0.12 years ($p < .1$) for the fourth block; the average ACT-E score was greater for the comparison group by 0.8 points ($p < .1$) in the second block and greater for the intervention group by 1.4 points ($p < .1$) in the fifth block; and, the average composition

Table 15

Average Control Variable Values by Propensity Score Block

	Block	Years HS Physics			Female			ACT-English		
		Comp	Pre	p	Comp	Pre	p	Comp	Pre	p
1	$0 < \rho < 0.2$	1.59	1.88	***	0.12	0.11		29.9	28.7	*
2	$0.2 < \rho < 0.4$	1.38	1.31		0.32	0.17	***	30.0	29.2	*
3	$0.4 < \rho < 0.6$	1.26	1.26		0.17	0.22	*	28.8	29.2	
4	$0.6 < \rho < 0.8$	1.23	1.11	*	0.17	0.23	*	29.8	29.8	
5	$0.8 < \rho < 1$	1.12	1.08		0.26	0.49	***	27.8	29.2	*

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$)
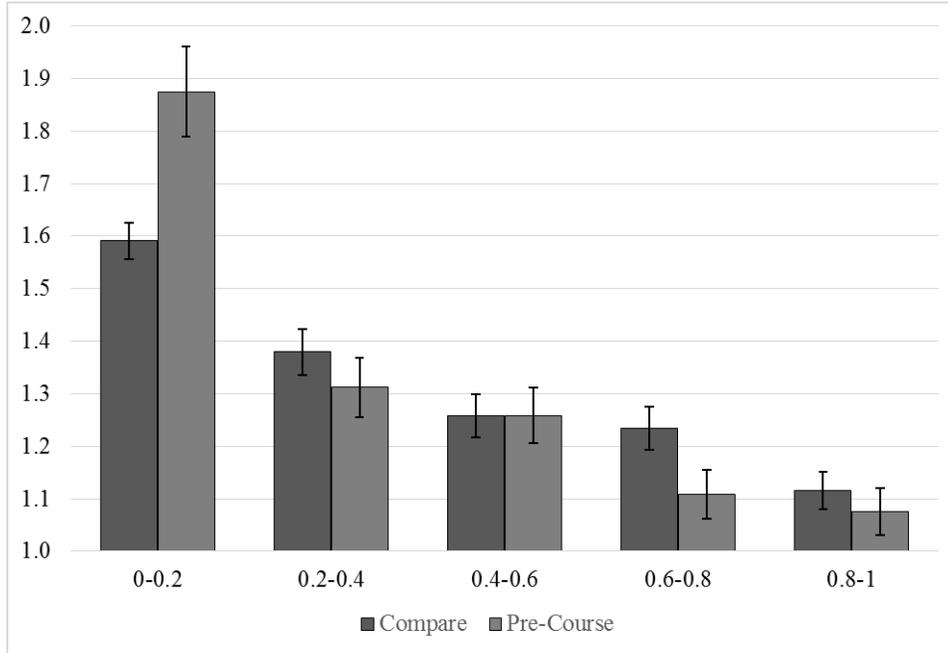
Figure 4: Years High School Physics per block
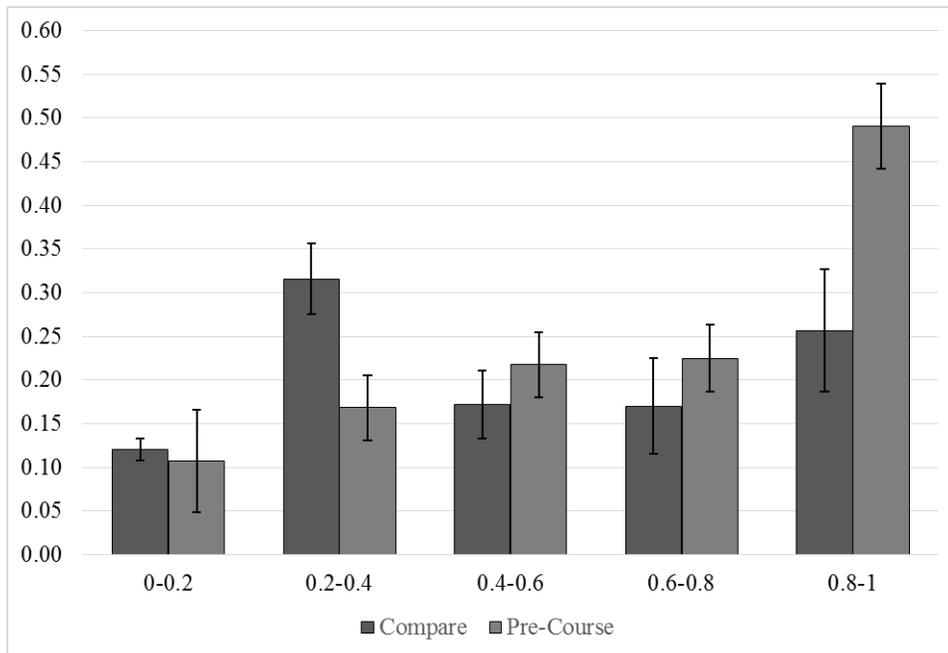

Figure 5: Percentage of Females per block
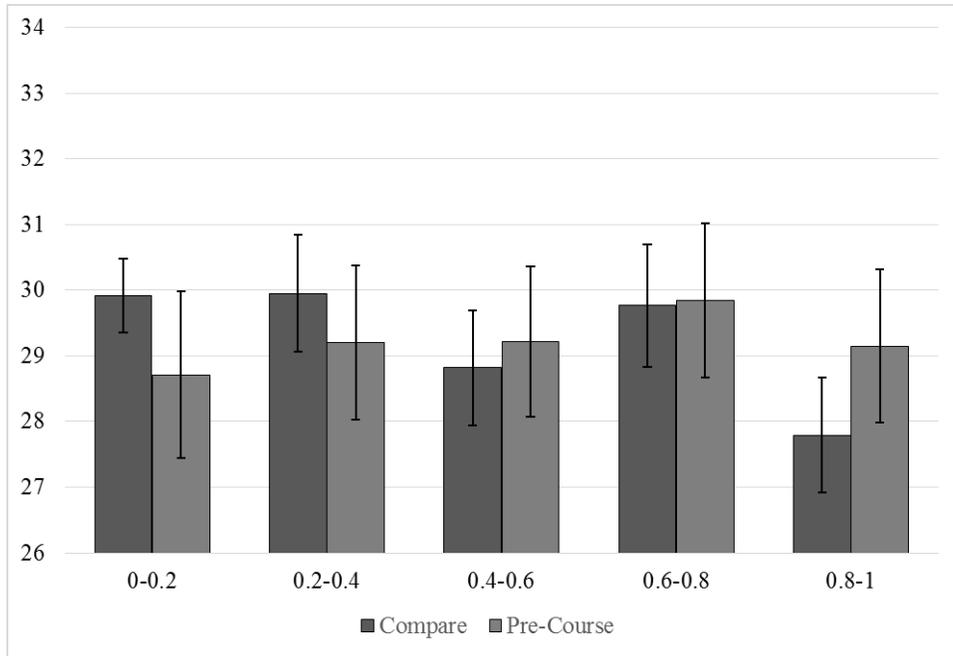
Figure 6: ACT-English Score per block
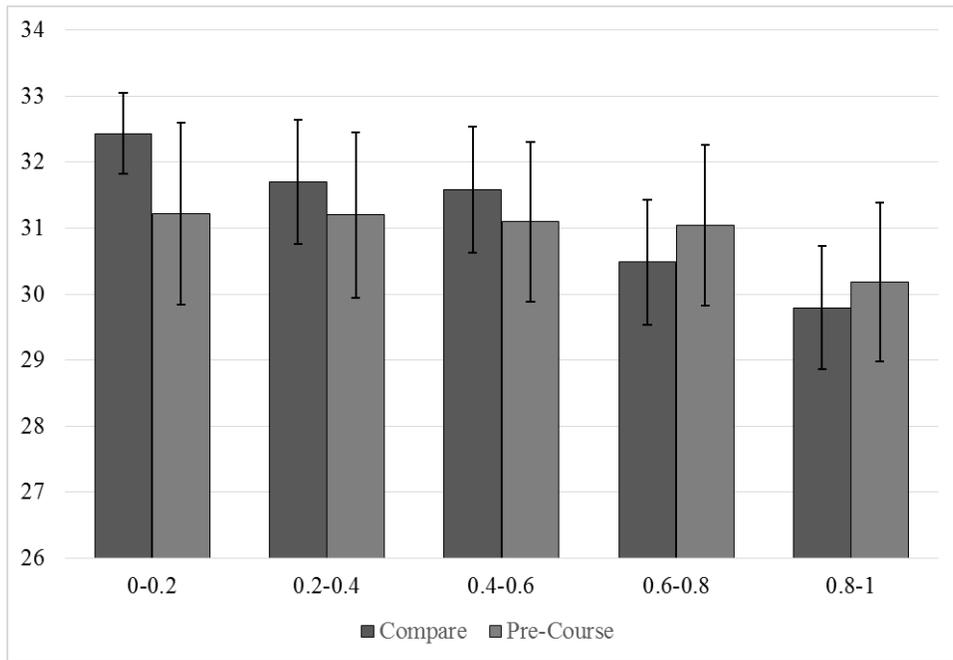
Figure 7: ACT-Math per block

of females was greater for the comparison group by 14.8% ($p < .001$) in the second block and greater for the intervention group by 23.5% ($p < .001$) in the fifth block. Judgment was made that best match in the blocks had been achieved as the control variables collected in the study allowed. Further comments regarding limitations of this model and its subsequent interpretations and implications are expanded upon in Chapter 5.

Estimate treatment effect. Average composite exam scores for the treatment and comparison groups were calculated for each block (Table 16 and Figure 8). A difference of means *t*-test was used to determine which of these differences were significantly equal for each block. An overall effect on performance was calculated by weighting the difference in each block by the number of treatment students in the block, summing these values up, and then dividing by the total number of treatment students. Omitting the first block due to initial differences, the final treatment weighted average of the effect for students enrolled in the Pre-Course treatment is 2.03 points (out of 100). This compares well to the treatment effect of 2.20 ± 0.80 points that resulted from the multiple linear regression Model H and corresponds to an effect size of 0.16 SD.

Table 16

Composite Exam Score by Propensity Score Block

	Block	Compare	Pre-Course	Diff	p
1	$0 < \rho < 0.2$	78.4	73.9	-4.49	*
2	$0.2 < \rho < 0.4$	73.5	73.3	-0.23	
3	$0.4 < \rho < 0.6$	70.0	72.5	2.49	
4	$0.6 < \rho < 0.8$	68.2	70.4	2.14	
5	$0.8 < \rho < 1$	62.7	66.2	3.51	

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$)
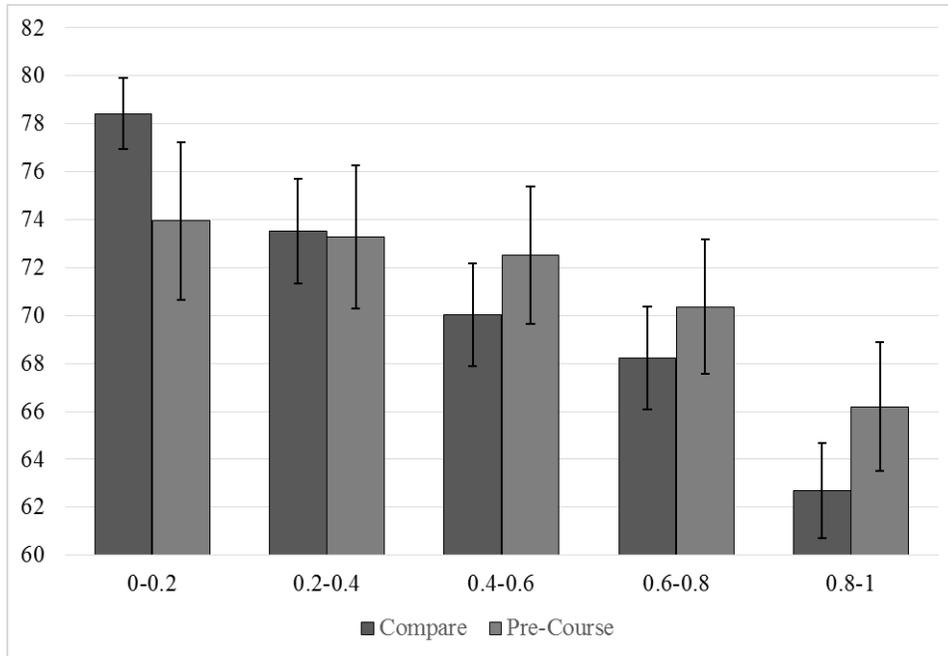
Figure 8: Average Composite Exam Score per Propensity Score Block

A final multiple linear regression onto Composite Exam Score using the calculated propensity scores was performed (Table 17). Using the propensity score as the control variable, the Pre-Course students earned a 1.52 ± 0.82 boost compared to their non-intervention peers, yielding an effect size of 0.12 SD. A low amount of variance in target course performance is explained by this model (about 13%); this issue is addressed further in the next chapter.

Table 17

Summary for Multiple Linear Regression onto Target Course Performance Using Propensity Score

	Estimate	p	Error
Intercept	79.6	***	0.47
Propensity Score	-18.1	***	1.36
Pre-Course	1.52	.	0.82
R^2		0.138	
F(df_1,df_2)		F(2,1451) = 117.1	

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$)

Concurrent intervention effect. An attempt of this process for calculating and matching propensity scores was conducted similarly for the Con-Course intervention students; however, the low number of participants in the Con-Course (62 students) prevented the development of an adequate model. Without propensity score analysis available, the initial multiple linear regression model (Model H) serves to provide an estimate for the effect on performance for this intervention, with an effect on performance of $+5.86 \pm 1.64$ points out of 100 ($p < .001$).

Persistence. Students receiving non-passing grades (defined as earning a letter grade of C- or lower in the first enrollment of the target course) in the comparison and intervention groups were gathered. Composite exam scores from subsequent enrollment were correlated with the pre-existing data and averages were calculated for each group for the first and second enrollment in the target course. The number of students fitting these criteria was small compared to the scale of students included in the propensity score analysis.

Aggregate results indicate that the Pre-Course students experienced a significantly smaller passing rate in the target course in reference to the Comparison group (Table 18) and Con-Course students experienced no difference in passing rate. These results are not surprising because the Pre-Course group was initially identified for being at risk for poor performance in the target course.

Table 18

Overall Passing Rates (Percentages)

Group	N	Passing Rate	p
Compare	1017	82.2	
Pre-Course	508	75.2	**
Con-Course	62	82.2	

(significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.10$)

To delineate the effects across the different degrees of risk, an analysis was performed within the blocks for the Pre-Course students (Table 19 and Figure 9). As expected, the passing rate decreased as propensity for enrolling in the intervention course and thus being at risk for poor performance increased. Dividing the students by propensity score block diminishes the power to observe significant differences. The observed effects, however, suggest that students in the Pre-Course intervention were no more likely to fail than their comparison peers. This is contrary to the observed effect in the overall passing rates, which, if interpreted incorrectly, could imply that the Pre-Course may even be detrimental to students compared to no intervention.

Table 19

Passing Rates by Propensity Score Block (Percentages)

|	Block	Compare	Pre-Course	Diff	p
1	$0 < \rho < 0.2$	84.8	89.3	4.52	
2	$0.2 < \rho < 0.4$	85.7	80.2	-5.52	
3	$0.4 < \rho < 0.6$	76.3	78.2	1.88	
4	$0.6 < \rho < 0.8$	68.1	75.0	6.91	
5	$0.8 < \rho < 1$	64.1	62.3	-1.84	
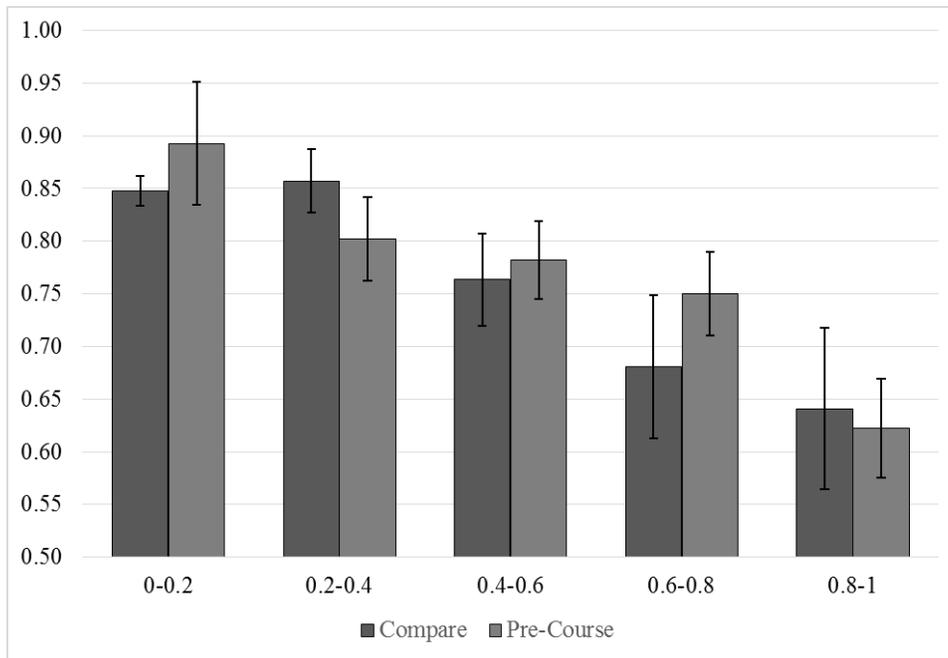
Figure 9: Average Passing Rate per Propensity Score Block

Summary of statistical analysis. In conclusion, the effect on performance in the target course, controlling for a host of variables including self-selection bias, was calculated for the Pre-Course and Con-Course students relative to the non-intervention comparison group. These results indicated that the Pre-Course improved performance in at-risk students in the study by 2.20 ± 0.80 points out of 100, corresponding to 0.17 standard deviations. The Con-Course students' performance was improved by 5.86 ± 1.64 points, corresponding to 0.46 standard deviations. Subsequent propensity score analysis provided a view of the impact on performance across degrees of risk of the Pre-Course students. This analysis confirmed the initial results with a 2.03-point performance boost that corresponds to a 0.16 SD effect size. The discussion of the statistical findings in context of the curriculum and pedagogy of the two intervention courses is expanded in a larger context in the next chapter.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

This study examined the impact of two supplemental intervention courses (Pre-Course and Con-Course) that are intended to improve the performance of at-risk students in an introductory physics course. These two interventions differed not only in chronological relation to the course but, upon analysis, also in regard to scope of curricular content. The two major aspects of this study that contributed to the results to be discussed were the curricular and pedagogical analyses of the two courses and the statistical analyses of relevant performance measures. While statistical quantification of the effects allowed for an objective comparison, the study also allowed contextualizing the results of this comparison with regard to differential curricular and pedagogical decisions in the two courses, which might account for the observed impacts. In the following sections, a comparison of the intervention courses' salient features followed by a summary of the quantitative results allows drawing a conclusion about the effectiveness of each course and a discussion of the context created in the comparison. Theoretical and methodological issues are also addressed, as well as deduced implications and suggestions for further research.

A Comparison of the Intervention Courses' Salient Features

As evident in Table 20, contrasting the content and pedagogical dimensions of the two intervention courses highlighted a number of commonalities and differences. The physics concepts in the two intervention courses overlap to a large extent as they essentially are selected to address the same content covered in the target course. Similarly, the assessment methods align in that the activities and aspects for evaluation are similar in the Pre-Course and the target course. Nonetheless, the Pre-Course features a summative assessment component that is not

89

Table 20

Comparison of the Curricular and Pedagogical Dimensions of the Intervention Courses

Course aspect	Pre-Course	Con-Course
Philosophy	Emphasis on active student engagement in learning processing	Emphasis on metacognition embedded in collaborative group work
	Teacher as collaborative facilitator and model of expertise	Teacher as collaborative and metacognitive facilitator
Objectives and curriculum	Problem solving skills and physical reasoning delivered through Kinematics and Newton's Laws	Problem solving skills delivered through the concepts covered in the target course
Instructional materials	Web-based prelecture, checkpoint, homework, and quiz; lecture notes; concept-rich problem packets	Concept-rich problem packets
Assessments	Formative: embedded prelecture questions, checkpoint questions, and weekly homework	Formative: weekly homework, weekly participation
	Summative: weekly quizzes, midterm exam, final exam	
Learning environment	Extensive web-based virtual learning environment; interactive classroom lecture with electronic classroom response system (1 hr/week); collaborative discussion section (2 hrs/week)	Collaborative discussion section (2 hrs/week)
Student engagement behaviors	Web-based instructional materials; collaborative interactions in discussion	Collaborative interactions in discussion

included in the Con-Course. While both courses feature student collaboration, the Pre-Course

makes more use of web-based tools to augment the learning environment.

90

Additionally, the philosophy underlying the curriculum and pedagogy in the Pre-Course and Con-Course is very similar and based on the active engagement of students in their own learning around conceptually rich knowledge. Instruction focuses on physics as a process, rather than as a set of facts or isolated knowledge to be learned. The framing of the learning environment and processes in both intervention courses serve students well because they are aligned with the instructional approach in the target course.

Thus, the Pre-Course differs from the Con-Course in terms of the inclusion of a summative assessment component and reliance on web-based tools as part of the learning environment. The one additional and salient difference is that the two-hour discussion section of the Con-Course goes beyond providing additional, conceptually rich practice problems by including a substantial metacognitive component of structured reasoning. The Con-Course's five-step problem-solving reasoning process uses specific and prescribed steps that foster an awareness directed toward deliberate and masterful concept application. Essentially, comparing the two courses from within the developed theoretical framework, the emphasis on scaffolding the development and application of metacognitive processes proves to be the major practical and philosophical difference in addition to the clear chronological difference identified at the outset of the study.

A Comparison of the Relative Impacts of the Two Intervention Courses

Multiple linear regression including statistically significant control variables showed a 2.20 ± 0.80 point (0.17 SD) advantage for Pre-Course students and a 5.86 ± 1.64 point (0.46 SD) advantage for Con-Course students in the target course composite exam score relative to the non-intervention comparison students. The reader is reminded that Con-Course students had enrolled in both intervention courses. A hypothesis could be presented that if the Con-Course had no

substantial effect, there would be no difference in performance gain for each group compared to the target course students' performance. A *t*-test of this difference in performance for the two intervention groups ($N_{PC} = 412$, $N_{CC} = 54$) yielded a gain of ΔExComp of 3.51 points ($t = 1.89$; *p* = 0.0594). While not statistically significant, an inference could be made that this additional effect seen in the Con-Course intervention group beyond the base effect of the Pre-Course is due to the Con-Course curriculum and pedagogy. However, simply receiving additional instruction might also have contributed to the observed difference. Overall, the results of the study suggest that both Pre-Course and Con-Course interventions had a positive impact on participant student performance in the target course. Another hypothesis could be made that the Pre-Course intervention laid the foundation for the Con-Course intervention to work. The Con-Course students received two interventions, not just one, as would be ideal in proper quasi-experimental comparison of interventions. Since single treatment for each group was not established, no clear statement can be made as to whether the Con-Course intervention would work independent of the Pre-Course intervention.

Propensity score analysis provided a more detailed examination of the Pre-Course students' performance gain. Inspecting the average composite exam scores within each of the five propensity score blocks delineated to what degree students of different levels of perceived risk were affected by the Pre-Course intervention. No statistically significant differences were evident for the composite exam score within each block outside of the first, which was discarded due to disparate control variable differences in this block. The three groups of students with highest propensity scores (and thus of highest perceived risk) consistently earned a higher composite exam score compared to their non-intervention peers. The treatment weighted average effect across these three particular groups was 2.68 points (out of 100). The group of highest

propensity achieved a gain of 3.51 points (0.28 SD)—which corresponds to about half a letter grade, compared to their peers in the comparison group, indicating that the Pre-Course intervention was most effective for this block of students.

With regard to persistence to target course completion, the passing rate of Con-Course intervention students (0.82) was not different than the non-intervention students (0.82). The Pre-Course passing rate (0.75) was investigated further within propensity scores to control for degree of risk across the different blocks. While this analysis revealed no significant boost in passing rate within any one block as with the trend seen with performance, it is suggestive that Pre-Course students pass at higher rates relative to their non-intervention peers of comparable risk. Combined with the performance results, the study suggests that the Con-Course allowed students not only to earn higher scores compared to their peers on assessments in the target course but also to gain in overall performance to allow them to maintain pace with their peers, at least in passing the target course.

Finally, inconsistent with the research literature, the present results showed that the variables of First generation, Parent income, and Race (Black) were not significant contributors to composite exam performance for this particular population of students. Multiple linear regression indicated that variables that contributed in significant negative ways to the observed performance in the target course were Gender (Female), and the races Asian and Black, indicating these groups experienced a disadvantage compared to their peers in the context of this particular study. Furthermore, only four variables out of the considerable number included were identified to contribute significantly to the propensity score created for the Pre-Course students. ACT-English and Female contributed positively, while the Physics Diagnostic and years of high school physics contributed negatively. This indicates that women and those with high ACT-

93

English scores were of greater likelihood to enroll in the Pre-Course intervention. These variables rose as influential due to the homogeneity across the values of the traditional predictor variables related to scholastic aptitude, cognitive ability, and prior educational experiences collected for the study. The cumulative effect of the included significant variables in accounting for the observed variance in performance is indicated by the R^2 value of 0.2984. That is, about 30% of the variance observed in performance can be accounted for by the variables included in the regression Model H. The remaining 70% is unaccounted for, suggesting two major conclusions.

The first major conclusion is that the collected measures from the referenced constructs do not serve well to reliably predict enrollment in the intervention course or target course performance for this population of students, which can best be described as a homogeneous group of high-performing individuals with no presumed inadequacies with regard to scholastic aptitude and cognitive ability. The measures included in the present study to describe the population in ways that the research or policy community finds useful (e.g., achievement measures and scores, including high school performance, national exam scores, Physics Diagnostic score, etc.) do not serve as useful proxies or paint a clear picture for all the factors considered when enrolling in an intervention course.

The other major conclusion is that variables representing constructs not included in the collected data contribute to students' observed intervention enrollment and physics performance. Students in both intervention courses in this study were self-identified as at risk due to the enrollment in these courses not being required by the course curricula of their majors. The degree to which this self-selection affected their performance was not isolated. The persistence of gender being correlated to performance and choice in pursuing intervention despite adequacies in

the traditional predictor variables is suggestive that affective factors such as self-confidence and perception of self-efficacy contribute to the observed results. Gutiérrez (2008) suggests that a focus is needed in achievement gap literature on those measures that are easily quantifiable and a neglect of the complexities that contribute to student identity and student agency. Underlying skills and knowledge that might contribute to overall performance and persistence exist among some students and not others. Due to this emphasis placed on national standardized exams and performance in high school, students may not even be aware or trust the diagnostic exam in assessing their risk due to their previous academic success, which is particularly high due to the admission standards at the participant university. Additionally, this study does not reveal which other factors students consider when choosing enrollment in an intervention course intended for at-risk students.

Limitations of the Study

Methodological Issues. First, a propensity score analysis for Con-Course students was not possible because the number of students in that intervention group was much too small. As such, the study was unable to isolate and produce a fuller understanding of the intervention effects on students of varying degrees of perceived risk in the Con-Course as was done for the Pre-Course. An issue presenting challenge in the interpretation of results is the amount of variance explained by the statistical models both in the multiple linear regression and for the propensity score analysis. In performing a complete and thorough regression, selection of a model of best fit included making discretionary decisions about the simplicity of the model. The differences seen across the blocks on the control variables in the propensity score model bring about critique of how matched these groups were on relevant variables with regard to evaluating the difference in performance. On this issue, Dehejia and Wahba (2002) noted that if the average

95

values of the covariates within each stratum are not statistically equal, the stratum may be too coarsely defined or, worse, the propensity score may be poorly estimated, such that interaction or higher order terms for the covariates should be included in the multiple linear regression creating propensity scores. However, thorough stepwise development of stepwise regression models ensured the best fit of the available control variables. As discussed above, inclusion of variables representing constructs affecting student risk for this population would improve the control on variations such that performance effects could be better isolated. Another consideration is the loss in statistical power due to the division of the Pre-Course students into different blocks. This resulted in an inability to make conclusive statements about the impact of the intervention courses on any one propensity score block. This loss of power continued through to the analysis of passing rates.

Generalizability of the Findings. The present findings are limited in their generalizability to different student populations at other institutions because of the characteristics of students perceived as "at risk" in this study. Participants in this study are high performing compared to observed national averages, thus, limiting claims that can be made about the possible effectiveness of the present intervention at other institutions with different admission standards.

Reliability and Validity of the Physics Diagnostic. The Physics Diagnostic is a 16-item web browser-based assessment completed on campus by students in the early summer prior to academic advising for the fall. The intended purpose of the assessment is to identify students who, having earned a low score in the assessed physics and mathematics knowledge and reasoning, would benefit from the Pre-Course's focus on how to approach solving physics problems. The first three items ask students for information regarding their prior physics

96

experience: "How many years of high school physics did you take?" "If you took any physics in high school, how would you rate the quality of the class?" "Do you think you will have trouble passing [the target course]?" The 13 other diagnostic items cover the conceptual topics of kinematics and uniform circular motion, as well as mathematical and proportional reasoning, algebra, and trigonometry with vectors.

Statistical comparisons performed on the original data showed that a relatively low correlation of the Physics Diagnostic score with the Composite Exam score (r = 0.395, $p < .001$) and the Total Course Points (r = 0.307, $p < .001$). Additionally, low covariances were observed among the Physics Diagnostic and all of the ACT scores, as well as years of high school physics completed. This finding first indicates that the Physics Diagnostic assesses knowledge and skills beyond those on the ACT exams. Second, it indicates that prior knowledge cannot be accurately measured by years of high school physics.

Implications and Future Directions

Future research branches in two directions stemming from the conclusions and observed limitations of the study. First is the collection and inclusion of control variables representing constructs that affect performance on a second-order level past the traditional predictor variables of amount of prior education, scholastic aptitude, and cognitive ability. The quality and characteristics of prior education has been studied in predictor variable literature. For example, Hazari, Tai, and Sadler (2007) collected information regarding the curriculum and pedagogy of high school physics courses as well as family attitudes toward science to include as control variables in multiple linear regressions to predict performance in university physics. As mentioned, Gutiérrez (2008) posits that measures of student identity and agency contribute to performance at the post-secondary level. This statement is supported by literature surrounding

the topic of stereotype threat. Initially identified by Katz, Roberts, and Robinson (1965), the theory of stereotype threat begins with the assumption that one must accommodate into their self-definition and remain accountable to a sense of school success and ability to achieve academically in order for such realities to manifest (Steele, 1997). Stereotype threat can thus be thought of as the discomfort students feel when they are aware of negative stereotypes associated with their identity or social group. Their cognitive function and intellectual performance can be affected by the fear that they could behave in such a way as to confirm the stereotype—in the eyes of others, in their own eyes, or both at the same time. Regardless of actual belief, knowledge of the negative stereotype can threaten aspects of self-concept, including self-esteem, self-confidence, and self-presentation. The construct of stereotype threat is often quantified as evaluation apprehension, self-efficacy, and anxiety that disrupts and interferes with performance and especially on high-risk assessments as well as a degree of disidentification from achievement in the domain (Aronson, Quinn, & Spencer, 1998). Surveys and student interviews can be conducted to collect this affective information to determine its influence on performance, likelihood to be at risk, and likelihood to pursue intervention for risk.

Second is design and implementation of thorough study isolating the effects of the Pre-Course curriculum and pedagogy from that of the Con-Course. The results of the study leave unclear whether the Con-Course had a statistically significant effect in addition to the Pre-Course intervention and also whether the Con-Course curriculum and pedagogy was dependent upon the foundation of knowledge and skills established in the Pre-Course.

REFERENCES

ACT (2013). National Ranks for Test Scores and Composite Score. Retrieved from http://www.actstudent.org/scores/norms1.html

Adams, S., & Garrett, H. L. (1954, March). Scholastic background as related to success in college physics. *Journal of Educational Research, 47*, 545–549.

Alexander, B. B., Burda, A. C., & Millar, S. B. (1997). A community approach to learning calculus: Fostering success for underrepresented ethnic minorities in an Emerging Scholars Program. *Journal of Women and Minorities in Science and Engineering, 3*(3), 145–159.

Ameduri, R. A. (1974). An analysis of student grades in a university physical science course for pre-nurses and their high school science and mathematics backgrounds. *School Science and Mathematics, 74*(March), 251–254.

Arons, A. B., McDermott, L. C., Thornton, R. K., & Mazur, E. (1997). *Conference on the Introductory Physics Course*. Wilson, J. M. (Ed). New York: Wiley.

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swin & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83-103). San Diego: Academic Press.

Banks, C. A. M., & Banks, J. A. (1995). Equity pedagogy: An essential component of multicultural education. *Theory into Practice, 34*(3), 152–158.

Barlow, A. E. L., & Villarejo, M. (2004). Making a difference for minorities: Evaluation of an educational enrichment program. *Journal of Research in Science Teaching, 41*(9), 861–881.

Barnes, G. (1977). Scores on a Piaget-type questionnaire versus semester grades for lower-division college physics students. *American Journal of Physics, 45*(9), 841–847.

Beeber, C., & Bierman, C. A. (2007). Building successes out of at-risk students: The role of a biology foundations course. *The American Biology Teacher, 69*(4), 48–53.

Belenky, M., Clinchy, B., Goldberger, N., & Tarule, J. (1986). *Women's ways of knowing.* New York: Basic Books.

Bickmore-Brand, J. & Gawned, S. (1993). Scaffolding for improved mathematical understanding. In J. Bickmore-Brand (Ed.), *Language in mathematics* (pp. 43–58). Carlton South, Victoria: Australian Reading Association.

Bonsangue, M. V. (1994). An efficacy study of the calculus workshop model. In E. Dubinsky, A. Schoenfeld, & J. Kaput (Eds.), *Issues in mathematics education (vol. 4): Research in collegiate mathematics education I* (pp. 117–137). Providence, RI: American Math Society.

Bonsangue, M. V., & Drew D. E. (1995). Increasing minority students' success in calculus. *New Directions for Teaching and Learning, 61*, 23–33.

Born, Wendi K., Revelle, William, & Pinto, Lawrence H. (2002). Improving Biology Performance with Workshop Groups. *Journal of Science Education and Technology, 11*(4), 347–365.

Brahmia, S., & Etkina, E. (2001). Switching students on to science: An Innovative course design for physics students. *Journal of College Science Teaching, 31*(3), 183–187.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school.* Washington, D.C.: National Academy.

Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H.W. Reese (Ed.), *Advances in Child development and Behavior (Vol. 10).* New York: Academic Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Burron, B., James, M. L., & Ambrosio, A. L. (1993). The effects of cooperative learning in a physical science course for elementary/middle level preservice teachers. *Journal of Research in Science Teaching, 20*, 697–707.

Campbell, Denes, & Morrison (Eds.) (2000). *Access Denied: Race, Ethnicity, and the Scientific Enterprise.* NY: Oxford University.

Carifio, J. (2006). Standard scores based on the median and inter-quartile range. *Journal of Mathematics and Statistics, 2*(1), 367–367.

Champagne, A. B. & Klopfer, L. E. (1982). A causal model of students' achievement in a college physics course. *Journal of Research in Science Teaching, 9*(4), 299–309.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. (1980). Factors influencing the learning of classical mechanics, *American Journal of Physics, 48*(12), 1074–1079.

Clark, M. H., & Cundiff, N. L. (2011). Assessing the effectiveness of a college freshman seminar using propensity score adjustments. *Research in Higher Education, 52*, 616–639.

Cohen, H. D., Hillman, D. F., & Agne, R. M. (1978). Cognitive level and college physics achievement. *American Journal of Physics, 46*(10), 1026–1029.

Cohen, W. S. (1997). *BioExcel at the University of Kentucky.* (Unpublished project description). University of Kentucky, KY.

Cottle, P. D. & Hart, G. E. (1996). Cooperative learning in the tutorials of a large lecture physics class. *Research in Science Education, 26*(2), 219–231.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84*(1), 151–161.

Deutsch, M. (1949). A theory of cooperation and competition. *Human Relations, 2*, 129-152.

Dewey, J. (1943). *The school and society* (Rev. ed.). Chicago: University of Chicago Press.

Drane, D., Smith, H. D., Light, G., Pinto, L., & Swarat, S. (2005). The Gateway Science Workshop Program: Enhancing student performance and retention in the sciences through peer-facilitated discussion. *Journal of Science Education and Technology, 14*(3), 337–352.

Duncan, H., & Dick, T. (2000). Collaborative workshops and student achievement performance in introductory college mathematics courses: A study of a Treisman Model Math Excel program. *School Science and Mathematics, 100*(7), 365–373.

Etkina, E., Gibbons, K., Holton, B. L., & Horton, G. K. (1999). Lessons learned: A case study of an integrated way of teaching introductory physics to at-risk students at Rutgers University. *American Journal of Physics; 67*(9), 810–818.

Flavell, J. H. (1973). Metacognitive aspects of problem-solving. In L.B. Resnick (Ed.), *The Nature of Intelligence*. Hillsdale, NJ: Erlbaum.

Fosnot, C. T. (1996). Constructivism: A psychological theory of learning. In C. T. Fosnot (Ed.), *Constructivism: Theory, perspectives, and practice* (pp. 8–33). New York: Teachers College Press.

Foster, H. (1972). The problem of identifying the potential black (or other minority) physics student. *The Physics Teacher, 10*, 330–346.

Freeman, M. (1995). *MathExcel: A special opportunity in calculus.* (Unpublished report). University of Kentucky, KY.

Freeman, M. (1997). Collaborative math and science workshops in Kentucky and Appalachia. *Mathematics and Education Research Forum, 9*, 5–8.

Fullilove, M., Fullilove, R. E., Terris, M., & Lacayo, N. (1988). Is "Black achievement" an oxymoron? *Thought and Action, 4*, 5–20.

Fullilove, R. E., & Treisman, P. U. (1990). Mathematics achievement among African American undergraduates at the University of California, Berkeley: An evaluation of the Mathematics Workshop Program. *Journal of Negro Education, 59*(3), 463–478.

Gay, G. (2000). *Culturally responsive teaching.* New York: Teachers College Press.

Gay, G. (2002). Preparing for culturally relevant teaching. *Journal of Teacher Education, 53*(2), 106–116.

Gay, G. (2004). Beyond Brown: Promoting equality through multicultural education. *Journal of Curriculum and Supervision, 19*(3), 193–216.

Gosser, D. K., Roth, V., Gafney, L., Kampmeier, J. A., Strozak, V. S., Varma-Nelson, P., Radel, S., & Weiner, M. (1996). Workshop chemistry: Overcoming the barriers to student success. *The Chemical Educator, 1*(1). doi:10.1333/s00897960002a

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and Learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 15-46). New York: Simon & Schuster Macmillan.

Griffith, W. T. (1985). Factors affecting performance in introductory physics courses. *American Journal of Physics, 53*(9), 839–842.

Gutiérrez, R. (2008). A "gap gazing" fetish in mathematics education? Problematizing research on the achievement gap. *Journal for Research in Mathematics Education, 39*(4), 357-364.

Hake, R. R. (1998). Interactive engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. A*merican Journal of Physics, 66*, 64–74.

Hart, G. E., & Cottle, P. D. (1993). Academic backgrounds and achievement in college physics. *The Physics Teacher, 31*(8), 470–475.

Hazari, Z., Sadler, P. M., & Tai, R. H. (2007) Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors. *Science Education, 91*(6), 847–876.

Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping, Part 2: designing problems and structuring groups. *American Journal of Physics, 60*(7), 637–644.

Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping, Part 1: Group versus individual problem solving. *American Journal of Physics, 60*(7), 627–636.

Hollins, E. R. (1996). *Culture in school learning: Revealing the deep meaning*. NJ: Erlbaum.

Holton, B. E., & Horton, G. K. (1996). The Rutgers Physics Learning Center. *The Physics Teacher, 34*, 138–143.

Hudson, H. T. (1986). A comparison of cognitive skills between completes and dropouts in a college physics course. *Journal of Research in Science Teaching, 23*(1), 41–50.

Hudson, H. T., & Liberman, D. (1982). The combined effect of mathematics skills and formal operational reasoning on student performance in the general physics course. *American Journal of Physics, 50*(12), 1117–1119.

Hudson, H. T., & McIntire, W. R. (1977). Correlation between mathematical skills and success in physics. *American Journal of Physics, 45*(5), 470–471.

Hudson, H. T., & Rottmann, Ray M. (1981). Correlation between performance in physics and prior mathematics knowledge. *Journal of Research in Science Teaching, 18*(4), 291–294.

Kagan, S. (1989). *Cooperative learning resources for teachers.* San Juan Capistrano, CA.

Katz, I., Roberts, S. O., & Robinson, J. M. (1965). Effects of task difficulty, race of administrator, and instructions on digit-symbol performance of Negroes. *Journal of Personality and Social Psychology, 2*, 53-59.

Kaufman, H. (1977). Black engineering students: Why the poor retention rate? *New Engineer, 7*, 33–38.

Kosciuk, S. (1997). *Impact of the Wisconsin Emerging Scholars first-semester calculus program on grades and retention from Fall 1993–1996.* (Unpublished manuscript). University of Wisconsin-Madison, WI.

Ladson-Billings, G. (1992). Reading between the lines and beyond the pages: A culturally relevant approach to literacy teaching. *Theory into Practice, 31*(4), 312–320.

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice, 34*(3), 159–165.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465–491.

Laws, P. W. (1991). Calculus-based physics without lectures. *Physics Today, 44*(12), 24–31.

Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.

Liberman, D., & Hudson, H. T. (1979). Correlation between logical abilities and success in physics. *American Journal of Physics, 47*(9), 784–786.

Mazur, E. (1997). *Peer Instruction.* Upper Saddle River, NJ: Prentice Hall.

McCammon, S., Golden, J., & Wuensch, K. L. (1988). Predicting course performance in freshman and sophomore physics courses: Women are more predictable than men. *Journal of Research in Science Teaching, 25*(6), 501–510.

McDermott, L. C., Piternick, L. K. & Rosenquist, M. L. (1980c). Helping minority students succeed in science III: Requirements for the operation of an academic program in physics and biology. *Journal of College Science Teaching, 9*, 261–265.

McDermott, L. C., Piternick, L. K., & Rosenquist, M. L. (1980b). Helping minority students succeed in science: II. Implementation of a curriculum in Physics and Biology. *Journal of College Science Teaching; 9*(4), 201–205.

McDermott, L. C.; Piternick, L. K. & Rosenquist, M. L. (1980a). Helping minority students succeed in science: I. Development of a Curriculum in Physics and Biology. *Journal of College Science Teaching, 9*(3), 135–140.

Mendez, C. G. (2006). Teaching calculus and other challenging courses to nontraditional and at-risk students at a research university. *College Teaching, 54*(4), 291–297.

Moreno, S. E., & Muller, C. (1999). Success and diversity: The transition through first-year calculus in the university. *American Journal of Education, 108*(1), 30–57.

Moreno, S. E., Muller, C., Asera, R., Wyatt, L., & Epperson, J. (1999). Supporting minority mathematics achievement: The Emerging Scholars program at the University of Texas at Austin. *Journal of Women and Minorities in Science and Engineering, 5*(1), 53–66.

Murnane, R. J., & Willett, J. B. (2011). Dealing with bias in treatment effects estimated from non-experimental data. In *Method matters: Improving causal inference in educational and social science research*. New York: Oxford University.

National Science Board. (1987). *Task committee on undergraduate science and engineering education*. Washington, DC: National Science Board.

Nelson, P., Radel, S., & Weiner, M. (1996). Workshop chemistry: Overcoming the barriers to student success. *The Chemical Educator, 1*, 2–3.

Novak, J. D. (1977). *A theory of education*. Ithaca NY: Cornell University Press.

Ory, J. C., & Braskamp, L. A. (1988). Involvement and growth of students in three academic programs. *Research in Higher Education, 28*(2), 116–129.

Piaget, J. (1926). T*he language and thought of the child.* New York: Harcourt Brace.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. *American Journal of Physics, 65*, 45–54.

Riley, D., & Lionel, C. (2009). From persistence to resistance: Pedagogies of liberation for inclusive science and engineering. *International Journal of Gender, Science, and Technology, 1*(1), 37–60.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of propensity score on observational studies for causal effects. *Biometrika, 70*(1), 41–55.

Sadler, P. M., & Tai, R. H. (1997). The role of high-school physics in preparing students for college physics. *The Physics Teacher, 35*(May), 282–285.

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education, 85*, 111–136.

Solorzano, D. G., & Yosso, T. J. (2001). From racial stereotyping and deficit discourse. *Multicultural Education, 9*(1), 2–8.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 21–51.

Steele, C. M. (1997). Threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613-628.

Steele, C. M. & Aronson, J. (1995). Stereotype Threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.

Tai, R. H., Sadler, P. M., & Loehr, J. F. (2005). Factors influencing success in introductory college chemistry. *Journal of Research in Science Teaching 42*(9), 987–1012.

Taylor, B. A. P. (1983). Teaching scientific reasoning to underprepared students. *New Directions for Teaching and Learning, 15*, 7–17.

Terenzini, P. T., Cabrera, A. F., & Bernal, E. M. (2001). *Swimming against the tide: The poor in American higher education.* New York: College Entrance Examination Board.

Terenzini, P. T., Springer, L., Yaeger, P. M., Pascarella, E. T., & Nora, A. (1996). First-generation college students: Characteristics, experiences, and cognitive development. *Research in Higher Education, 37*(1), 1–22.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*, 89–125.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition (2nd ed.).* Chicago: University of Chicago Press.

Tinto, V., Goodsell Love, A. & Russo, P. (1993). Building community. *Liberal Education, 79*(4), 16–21.

Tobias, S. (1990). *They're not dumb, they're different: Stalking the second tier*. Tucson, AZ: Research Corporation.

Treisman, U. P. (1985). *A study of the mathematics achievement of Black students at the University of California, Berkeley.* (Unpublished doctoral dissertation). University of California, Berkeley.

Treisman, U. P. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal, 23*(5), 362–372.

Valencia, R. (1997). *The evolution of deficit thinking.* London: Falmer.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction, 16*(1), 3-118, 1998.

Wollman, W., & Lawrenz, F. (1984). Identifying potential "dropouts" from college physics classes. *Journal of Research in Science Teaching, 21(*4), 385–390.

United States Census Bureau (2012). Annual estimates of the Population for the United States, Regions, States, and Puerto Rico: April 1, 2010 to July 1, 2012 Population Estimates., Population Division.

APPENDIX A

GUIDED INTERVIEW QUESTIONS WITH COURSE DESIGNER AND

INSTRUCTORS

Done for each intervention course separately.

(1) Objectives and Curriculum

– What knowledge and skills do you presume they come in with that you're able to capitalize

on?

– In what ways is this integrated into the course?

(2) Instructional Materials and Assignments

– What sorts of materials are given to the student to engage in learning of the content?

– How are the teaching objectives delivered and fostered through the different materials and

assignments?

(3) Assessments

– What sort of evidence is evaluated to determine if students are progressing in the content and

skills?

– What is your justification for the scoring procedure for the course and how grades are

assigned?

– How do you determine how frequently assessments are administered?

– How are the methods of assessment consistent with your beliefs of the nature of what is to be

learned and how it is to be learned?

(4) Instructional philosophy and approach

– What are the key elements of your instructional philosophy that drive the development of the

course?

– How are these elements conveyed through your choice of topics and content?

– How do the main components of the course and content delivery align with your instructional philosophy?

– What are the motivations and factors for the particular choices and decisions you made in each of these instructional strategies?

– What is the reason for the particular ordering of the course components?

– How important is it that prior conceptions and knowledge be identified and addressed in the course? How is it accessed and utilized in the course?

– When are teaching assistants present and what is their intended role?

– Could you outline the teaching objectives?

– What do you see as your primary role as the instructor?

– What sorts of activities do you participate in as the instructor: to teach the material? To aid student learning?

(5) Student Engagement Behaviors

– What are the expectations for students' participation in the different course components?

– How are students intended to interact with and in the different course components?

(6) Interactions (student-student and student-teacher)

– Do you concern instruction with skills to articulate to one another when you put students in collaboration?

– What sort of articulation skills do students need to utilize?

(7) Learning Environment

– How are the discussion sections set up?

– How much do the students collaborate and how much do they work on their own in the

different components of the course?

– What is the role of who is present at the different times students are in a classroom? In what ways does this person(s) interact with the students to expedite the learning process?

APPENDIX B

CURRICULUM AND PEDAGOGY ANALYSIS PROTOCOL

(1) Objectives and Curriculum:

– Outline topics covered in the course.

– Outline learning objectives as specified in the course description and materials.

– Outline cognitive and content skills as specified in the course description and materials.

– Investigate and inventory content knowledge and required cognitive and content skills to

engage in and complete assignments and tasks in different components of the course

(2) Instructional Materials and Assignments:

– Outline type, frequency, and context of homework assignments

– Outline type, frequency, and context of in-class materials and tasks

– Outline type, frequency, and context of formative assessments outlined in course description

(3) Assessments:

– Outline type, frequency, and context of summative quizzes

– Outline type, frequency, and context of examinations

– Investigate and outline of the evidence required to exhibit mastery of knowledge and skills

– Summarize the explicitly stated methods of evaluation to score assessments, including

awarding of partial credit

– Outline scoring procedure for course and how grades are assigned

(4) Learning Environment:

– Describe the placement and location of student(s) to one another

– Describe the placement and location of students and teacher to one another

– Describe at length the resources made available to aid learning

(5) Student Engagement Behaviors:

– Describe at length what students do to learn

– Describe at length what students are required to put forth effort to learn

(6) Interactions (student-student and student-teacher):

– Describe who initiates interactions

– Types of help provided

– What type of information is communicated between parties

– Structure of interactions

– Communication methods used

(7) Instructional Philosophy and Approach:

– Nature of knowledge and how that knowledge is most effectively learned (brief & broad)

– Outline course principles & goals for long/short term goals for performance & retention

– Describe at length the content delivery methods

– Type, frequency, & context of aiding student learning

– Investigate consistency between online and in-class instructional materials