Brigham Young University
BYU ScholarsArchive

All Theses and Dissertations

2016-03-01

An Assessment of the Effects of Spiritual and Relational Teaching on Student Learning

Matthew Alan Hiatt
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Psychology Commons

BYU ScholarsArchive Citation

Hiatt, Matthew Alan, "An Assessment of the Effects of Spiritual and Relational Teaching on Student Learning" (2016). *All Theses and Dissertations*. 5698.
https://scholarsarchive.byu.edu/etd/5698

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

An Assessment of the Effects of Spiritual and Relational Teaching on Student Learning

Matthew Alan Hiatt

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Brent D. Slife, Chair
Jeffrey S. Reber
Alan L. Wilkins
Robert D. Ridge
Melissa A. Newberry

Department of Psychology

Brigham Young University

March 2016

Copyright © 2016 Matthew Alan Hiatt

All Rights Reserved

ABSTRACT

An Assessment of the Effects of Spiritual and
Relational Teaching on Student Learning

Matthew Alan Hiatt
Department of Psychology, BYU
Doctor of Philosophy

Research suggests that students are more interested than faculty in addressing spirituality in the classroom. This study tested the extent to which professors could meet student demand for greater attention to spirituality in their classes without sacrificing rigor and student learning. Previous research done at Brigham Young University (BYU) identified three areas of focus that are important to implementing spirituality into the classroom: 1) Professor Self-Disclosure, 2) Intellectual Connections, and 3) Interpersonal Connections. Research on the integration of faith and learning also supports these focus areas. Two BYU professors from different colleges were recruited for participation in this study. After teaching as they normally would for 30-minutes (the control condition), the professors were prepared to incorporate these focus areas into four experimental conditions (one for each focus area and one that combined all focus areas). A sample of 203 student participants were recruited to attend the different teaching sessions, fill out a questionnaire about their perception of the teaching quality, and take a retention examination on the material taught by the professor. An Exploratory Factor Analysis of the Teacher Rating Questionnaire showed that a three-factor pattern would be best to use in data analysis as it explained 89% of the variance in student responses. Based on the items that grouped together, those factors were labeled as General Teaching Skills, Openness and Respect, and Spirituality. A 5 (condition) X 2 (professor) factorial MANOVA—using the three extracted factors as dependent variables—indicated that there were significant main effects and an interaction. ANOVA and post hoc follow-ups revealed that professor ratings on General Teaching Skills and Spirituality greatly improved after the preparation; however, ratings also depended on the professor. A factorial ANOVA also revealed that student retention scores did not significantly differ from the control condition to the experimental conditions, but insufficient power suggests this finding should be interpreted with caution. It is concluded that applying a pedagogical preparation such as used in this study can be a useful tool in educating willing faculty to successfully implement spirituality and improve their general teaching skills.

Keywords: spirituality, relationality, higher education, integration of faith and learning, student learning, pedagogy

ACKNOWLEDGEMENTS

I would first like to thank Jeff Reber for bringing me on to his team. His mentorship has helped me to have an enjoyable experience in graduate school. I appreciate the hours he spent helping me get this dissertation to the point it is at now. His connection with Alan Wilkins in the BYU Faculty Center was essential to this dissertation. I could not have completed this dissertation without the support of Alan. He, too, has been a very caring mentor. His interest in my life has given me the desire to become a better person.

I also want to express my appreciation for the professors who participated in this study. They selflessly gave up their time to help further our understanding of how faith and learning can coincide. I asked a lot of them and they generously delivered. It was great to work with them.

Most importantly, I want to express my deep appreciation for my wife, Rebekah, who was by my side from the beginning of this journey. I always had her loving support, whether it was by bringing me dinner while I spent long hours on campus analyzing pilot data or by encouraging me to finish my preparations for the defense even though both she and the baby were miserably sick. I am eternally grateful for her.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS iv

LIST OF TABLES vii

LIST OF FIGURES viii

Introduction 1

Literature Review 4

- Integration of Faith and Learning (IFL) 4
- Teacher Immediacy 7
- Interpersonal Connections: The Foundational Focus 9
- A BYU Education 10
  - Theistic spirituality 11
- Pilot Studies 14
- Hypotheses 16

Method 18

- Brief Overview 18
- Participants 19
  - Students 19
  - Professors 20
- Measures 21
  - Reading excerpts 21
  - Teacher Rating Questionnaire 22
  - Retention Examination 25

Professor Self-Evaluation 26
Procedure 26
Procedure for students 26
Procedure for professors. 27
Procedure for manipulation check 30
Results 31
Manipulation Check 31
Data Screening 34
Demographic analyses 34
Evaluation data screening 35
Examination data screening 35
Data Reduction 35
Data screening by factor 39
Significance Tests 40
Teaching quality 40
Retention 47
Professor Self-Evaluations 49
Discussion 51
Effective Preparation 51
Difference in Professors' Perceived Experience 52
Teaching Quality Improvement 53
No Apparent Gains or Losses in Student Retention 56
Limitations and Suggestions for Future Research 58
Unequal groups 58

Lack of statistical power 58

Relationship between topics and focuses 59

Preparation and treatment fidelity 60

Other limitations 62

Additional directions for future research 63

Conclusion 65

References 66

Appendix A 73

Appendix B 78

Appendix C 81

Appendix D 84

Appendix E 85

Appendix F 86

LIST OF TABLES

Table 1 23

Table 2 32

Table 3 37

Table 4 41

Table 5 48

LIST OF FIGURES

Figure 1 43

Figure 2 45

Figure 3 46

Figure 4 48

An Assessment of the Effects of Spiritual and Relational Teaching on Student Learning

There is a current divide between what students want and what professors provide in regards to spirituality in higher education in America. A survey conducted by the University of California at Los Angeles (UCLA)'s Higher Education Research Institute (HERI) with over 100,000 student respondents from 236 diverse institutions found that up to 80% of students nationwide have an interest in spirituality, and 73% agree that it is "essential" or "very important" that colleges "encourage the personal expression of spirituality" (Astin, Astin, & Lindholm, 2011, p. 31; Lindholm, Astin, & Astin, 2005; Lindholm, Szelényi, Hurtado, & Korn, 2005). HERI also surveyed over 40,000 faculty from 421 universities and colleges and found that while four out of five faculty described themselves as "spiritual person[s]", only 30% of faculty felt that "colleges should be concerned with facilitating…students' spiritual development" (Astin et al., 2011, p. 39; Lindholm, Astin, & Astin, 2005; Lindholm, Szelényi, Hurtado, & Korn, 2005, p. 111). Another survey conducted by the Center for Survey Research at Indiana University found that 84% of professors agree with the statement "American colleges and universities welcome students of faith" (Gross & Simmons, 2008, p. 27), yet professors do little, if anything, to address faith. Indeed, 62% of students reported that their professors "never" encourage discussion of spiritual matters (Astin et al., 2011, p. 37).

This research suggests that students are more interested in the expression of spirituality in higher education than the faculty are interested in addressing it. Students come to college expecting to grow intellectually, but they also "bring their faith to campus with them" and "expect to 'grow' in their…spiritual lives while in college" (Braskamp, 2008, p. 125; Edwards, 2008). To the students' disappointment, most professors provide little opportunity to discuss spirituality in class (Astin et al., 2011). This can have serious implications on the students such

as a diminishing of their faith or spirituality (Pascarella & Terenzini, 1991) or a loss of an important connection with the faculty. One reason for this divide is that many professors fear that addressing spirituality might hinder intellectual enlargement (Wilkins & Birch, 2011). While students believe intellectual enlargement and spiritual growth can coexist, professors have their reservations.

Similar research findings and concerns are also present at Brigham Young University (BYU), a religiously affiliated (Christian) university where the present study took place. In a student survey Wilkins & Birch (2011) found that 90% of BYU students agree, at least to some extent, that "every course at BYU should be both spiritually strengthening and intellectually enlarging" (p. 2). Of those students, almost 60% feel that it is "very important" that courses are both spiritually strengthening *and* intellectually enlarging; however, only 35% believe that BYU is doing "very well" at meeting these expectations, suggesting that BYU could do better (p. 3). Improvement could come from addressing the faculty's specific concerns regarding spirituality in the classroom. Among those concerns, Wilkins and Birch (2011; A. Wilkins, personal communication February 6, 2014) note that some wonder how to start implementing spirituality into their teaching; others believe they will not be able to find time to do it. Additionally, some think that spirituality may not fit with their discipline. Perhaps one of the greatest concerns is that bringing up spirituality might jeopardize academic rigor and critical thinking and "lower expectations for performance" (Wilkins & Birch, 2011, p. 1). These anecdotal concerns are backed by Edwards (2008) who says that even in some religiously affiliated institutions "faculty rarely mention [their] personal religious or spiritual convictions in [their] scholarship or teaching" (p. 81). He notes that one of the most common reason given by religious academics

for this neglect is that many faculty think that their personal beliefs are not directly relevant to their discipline.

In the current study, I attempt to demonstrate that these faculty concerns may not be warranted by means of a quasi-experiment designed to help bridge the gap between student and faculty expectations and interests around spirituality. Speck (2005) claims that spirituality may be a genuine academic concern because many students will go into careers where their clients will raise issues of spiritual need. If students do not get insights from faculty into addressing the topic, they will be at a disadvantage in their careers. Some research suggests that one reason why faculty are less interested in spirituality is that they lack education, training, or preparation, in addressing the topic (Speck, 2005). "Faculty education in spirituality is virtually absent and absent again from the faculty's classrooms" (Speck, 2005, p. 8). My study intends to show that it is possible with pedagogical preparation and support for faculty to bring spirituality into the classroom and maintain academic rigor.

I use research (most unpublished because they are for internal use only) previously done by the Faculty Center at BYU as a foundation in addressing this student-professor divide. The Faculty Center used both student and faculty surveys and focus groups to look at what role faith could play in learning and what students consider important in both their courses and their professors for a spiritually strengthening experience (Wilkins & Birch, 2011). From this research the Faculty Center has identified three main factors, or areas of focus, that are important to implementing spirituality into learning. For purposes of this dissertation, I have labeled these factors as follows:

1. A focus on the professor's journey, personal experiences, and the ways in which spirituality is working in his or her personal and professional life. (Professor Self-Disclosure)

2. A focus on intellectual connections between spirituality and the discipline, and illuminating the meaning that those connections have in students' lives. (Intellectual Connections)
3. A focus on the professor's relationship with students. (Interpersonal Connections)

The aim of this study is to demonstrate the extent to which implementing these three areas of focus into a classroom lecture can improve students' perceptions of teaching quality while maintaining academic achievement. This study makes a unique contribution over the survey, interview, and anecdotal studies done thus far because it is the only study that utilizes a quasi-experimental design to investigate this specific topic. By so doing, it adds another empirical study to the small number of empirical studies that focus on similar, yet broader, topics.

In the literature review that follows I explain the background needed to understand this study's purpose. I will first illustrate how the three factors tie into the larger research literatures on the integration of faith and learning (IFL) and on teacher immediacy. I will then examine how the unique, faith-based, educational context of BYU fits with the areas of focus and how it can be used as a setting to address the divide between what students want and what faculty provide in regards to spirituality in the classroom. Because this study was conducted at a Christian university, I narrowed much of my focus to the literature on Christian spirituality. Notably, though, many scholars suggest that spirituality can be implemented at any college—religiously affiliated or not (e.g., Jacobsen & Jacobsen, 2008; Jacobsen & Jacobsen, 2012).

Literature Review

Integration of Faith and Learning (IFL)

For four decades, researchers have investigated the integration of faith and learning (IFL) in higher education. IFL has been defined as an interaction between professors and students

(Cooper, 1999) where both seek to make connections between faith (typically Christian) and academic disciplines (Hasker, 1992). As I will demonstrate, IFL closely relates to the three focus areas identified above and helps to illustrate how the three focus areas are interwoven with each other.

One of the primary aspects of IFL relates to the first factor: a focus on the professor's journey, personal experiences, and the ways in which spirituality is working in his or her personal and professional life. I have labeled this factor "Professor Self-Disclosure" because it involves the professor revealing parts of him- or herself that relate to his or her spiritual journey. In a Christian context, it is through this disclosure that students learn that professors have a personal relationship with God. In fact, using a sample of students from seven schools affiliated with different Christian denominations, Sherr, Huff, and Curran (2007) found that students feel that a professor's personal relationship with God is one of the most important contributors to IFL. They also found that students appreciate it when professors share personal experiences of God in which they refer to or draw on spiritual knowledge to make decisions.

This closely ties to the next factor, which I have labeled "Intellectual Connections." This factor focuses on intellectual connections between spirituality and the discipline, and illuminating the meaning that those connections have in students' lives. In IFL, many professors connect their personal experiences—including spiritual ones—to key concepts and topics of their discipline. According to Hasker (1992), when IFL is put into practice, professors do not have to create connections between faith and knowledge, but simply look for the relationships that naturally exist between the two. Professors can then spotlight the inherent connections between their faith and their discipline in the classroom (Sherr et al., 2007). Hasker (1992) also describes IFL as a practice that involves scholarly thinking that is already saturated by religious attitudes

and beliefs. For professors, this means recognizing that their faith informs their professional and personal lives in ways that relate to their discipline. As an example of how this might play out in a classroom, imagine a professor teaching about the endowment effect (the idea that people are less willing to give up items they possess and therefore attribute more value to them once they own them) in a course on Behavioral Economics. The professor—informed by his or her Christian beliefs—highlights the connection between the endowment effect and the parable of the talents found in Matthew 25:14-30 noting that the servant who was given one talent became too focused on his possession and its value became exaggerated beyond what he believed could be gained by putting it to use in the marketplace. The servant consequently hid his talent in the ground and was later reprimanded for not putting his talent to better use. The professor then invites the students to reflect on how they might be like the servant, placing too much value on what they possess and thus shortchanging their lives.

The third factor—a focus on the professor's relationship with students (what I call "Interpersonal Connections")—is a relational aspect of spirituality that has a very important role in IFL. Besides the professor's personal relationship with God, Sherr et al. (2007) found that students believe that the professor's relationship with students is vital to IFL. In defining IFL, Cooper (1999) keeps the main emphasis on an open relationship between professors and students, stating that "interaction between students and faculty members seems to lay the foundation for the integrative process" (p. 386). Prior studies found that students highly desire interaction with faculty yet studies have shown that interaction between professors and students in college classrooms can be infrequent and insincere (Feldman & Newcomb, 1969; NSSE, 2006). Wilson, Bavry, Gaff, Wood, and Dienst (1975) suggested that to increase the interaction between professors and students, professors should be more accessible (see also Frymier, 2013;

Witt, Wheeless, & Allen, 2006). However, accessibility is more than just being physically available (Cooper, 1999); it could also mean being spiritually available—being ready to open up and share faith-promoting, spiritual experiences or insights.

While the literature on IFL provides support for the three areas of focus, it lacks empirical studies on its impact on student learning. I was only able to find one quasi-experimental study (Wallace et al., 2008) that measured the influence of a spirituality education program on student learning, particularly in the nursing discipline.

Teacher Immediacy

The education literature on teacher immediacy provides more theoretical backing for the areas of focus and offers some empirical evidence for these factors having an impact on student learning.

Immediacy was introduced by social psychologist Albert Mehrabian (1969) and was originally conceptualized as communication behaviors that "enhance closeness to and nonverbal interaction with another" (p. 203). Mehrabian developed this concept based on approach-avoidance and reinforcement theories citing the immediacy principle as the idea that "people are drawn toward persons and things they like, evaluate highly, and prefer; and they avoid or move away from things they dislike, evaluate negatively, or do not prefer" (Mehrabian, 1971, p. 1). Immediacy now involves a set of both nonverbal and verbal communication behaviors that reflect liking and is referred to as the "degree of perceived physical or psychological closeness between people in a relationship" (Frymier, 2013, p. 425). This has direct application in teaching as students tend to draw toward teachers who display communication behaviors (e.g., calling students by name, smiling, use of humor and self-disclosure; for a more complete list of

immediacy behaviors see Gorham, 1988; Mehrabian, 1969, 1971) that enhance the perception of immediacy.

Immediacy clearly relates to the third factor, Interpersonal Connections. In fact, according to Witt and Wheeless (2001), "Immediacy…is predominately relational communication" (p. 328). The literature on immediacy also gives some support to the first factor, Professor Self-Disclosure. Self-disclosure is a verbal behavior that Mehrabian (1971) identifies as an important immediacy cue. It is perceived by students as an illustration of an interpersonal approach that has also been linked to positive learning outcomes (Witt & Wheeless, 2001). It is important to note that the immediacy literature suggests that Professor Self-Disclosure and Interpersonal Connections are closely related. In fact, Reber, Downs, and Peterson-Nelson (in press) use self-disclosure as the relational component in their study on the effect of relational teaching on learning outcomes. This can make it difficult to truly separate these focuses.

In a meta-analytical review involving 81 studies and 24,474 participants Witt, Wheeless, and Allen (2006) found a large, positive relationship ($r = .50$) between teacher immediacy and student learning. However, there were large differences across the type of learning measurement. The researchers found that immediacy is highly correlated with perceived and affective learning, but it has a small, yet significant, correlation with cognitive learning. Affective learning measures typically deal with assessing attitudes toward the teacher, the course, and enrolling in a similar course while cognitive learning measures are usually associated with recognition, recall, and understanding of course content (Allen, Witt, & Wheeless, 2006). Affective learning is easily measured through attitudinal surveys whereas cognitive learning is often measured through test scores or grades. Perceived learning is simply a measure of how

much the students believe they learned. The current study measures affective and cognitive learning.

Among the studies on teacher immediacy, only a few involve controlled experiments (e.g., Kelley & Gorham, 1988; Witt & Wheeless, 2001). The current study adds a controlled quasi-experiment to the broader education literature looking at student outcomes.

Interpersonal Connections: The Foundational Focus

Both the literature on IFL and on teacher immediacy suggest that the interaction—or more fundamentally, the relationship—between professor and student may actually be the factor that matters most. Sorenson (1997) conjectured that the relationship between students and their faculty mentors matters more than the actual content involved in integrating faith and learning. According to the literature on IFL, professors serve as attachment figures, or mentors, to students; they play a critical role in student development (Sorenson, 1997), and relationships that form in and continue outside of the classroom have the greatest impact on students (Wilson et al., 1975).

The literature on teacher immediacy gives further insight into how the professor-student relationship may be the most fundamental and important factor. Immediacy is largely relational (Witt & Wheeless, 2001), and given that immediacy has been shown to increase student learning (especially perceived and affective learning; e.g., Mullin, 2014; Witt et al., 2006), there is support to the essential role of the professor-student relationship. Indeed, decades of research on teacher immediacy have consistently demonstrated that "a teacher who is close to his or her students is more likely to be a good teacher than one who maintains a considerable social, emotional, and psychological distance between him/her and the pupils" (Sibii, 2010, p. 531).

In addition to this support, many philosophers, theologians, and psychologists recognize the fundamental role of relationship to spirituality. Martin Buber's (1958) philosophy of the I-Thou may be most helpful in supporting this claim. Buber explains the "self" in terms of relationships and proposes that the primary way in which we relate to nature, other people, and to God is through the I-Thou relationship. The I-Thou relationship involves a spiritual engagement with the other (Thou) that is facilitated by grace, in which the Thou has no discrete bounds and therefore is not an object. It is our relation with the eternal Thou that provides this grace and is fundamental to all other relations. As Buber (1958) described it, God is the "ground and meaning of our existence" (p. 135), making possible all "spheres in which the world of relation arises" including "our life with men" (p. 6). In other words, "all relations imply the primary relation of persons to God" (Slife & Reber, 2009, p. 75).

Buber's theory is a very theistic approach to relationality that has been supported by other scholars. James W. Fowler (1981), for example, who was best known for his work on faith, or spiritual, development, states that "faith is always relational" (p. 16). Faith is a way of seeing oneself "*in relation to* [emphasis added] others against a background of shared meaning and purpose" (Fowler, 1981, p. 4). These ideas taken together show that in the context of teaching, a professor who sees him- or herself *in relation to* his or her students is demonstrating the foundation of spirituality.

A BYU Education

Certainly, talk of "faith and learning" is not foreign at BYU. In fact, one of the four major education goals outlined in the BYU Mission Statement mentions that the Christian faith should be taught: "All students at BYU should be taught the truths of the gospel of Jesus Christ" (BYU Mission Statement). Even the first part of the aims of a BYU education states, "BYU

seeks to develop students of faith" (Aims of a BYU Education). BYU labels this the "Spiritually Strengthening" aim and declares "that every…teacher in this institution would keep his subject matter bathed in the light and color of the restored gospel" (Kimball, 1967, p. 11). This statement supports the second area of focus, Intellectual Connections, and indicates that professors and students should be integrating faith and learning.

In further support of Intellectual Connections, in the aims it is also stated, "A spiritually strengthening education warms and enlightens students by the bright fire of their teachers' faith while enlarging their minds with knowledge" (Aims of a BYU Education). This also alludes to the importance of the first factor, Professor Self-Disclosure. The "teachers' faith" acts as a "bright fire" to the students, suggesting that the professors must demonstrate a level of spirituality and a personal relationship with God. In fact, while mastering their academic discipline, faculty are encouraged to remain "mastered by the higher claims of discipleship to the Savior" (Aims of a BYU Education).

Relationships with God are an important part of the greater connection with members of the university community. "Certainly all relationships within the BYU community should reflect devout love of God and a loving, genuine concern for the welfare of our neighbor" (Aims of a BYU Education). This last point backs the third area of focus, Interpersonal Connections, and, as previously discussed, this relational component of spirituality is the foundational constituent of the three focuses. Notably, it has a unique form in a theistic context like BYU.

Theistic spirituality. The unique context of a BYU education fits well with the three factors proposed in the current study. It also links faith (as in the integration of faith and learning) to spirituality. As mentioned above, part of being spiritual at BYU means that the professor integrates his or her faith into the subject matter. While faith and spirituality are

closely connected at BYU, the broader literature on spirituality does not always make that connection. In fact, most definitions of spirituality do not even mention faith or religion at all.

Although the literature on spirituality is so expansive, the construct is often not as clearly defined as it is at BYU. Though there are many views of spirituality, my intention is not to pick out all of the definitional differences, nor is my intention to discuss all the issues that result from the definitional ambiguity and disagreement (see Speck, 2005). Rather, my goal in this section is to briefly review the two main paradigmatic approaches to worldviews about spirituality—namely naturalistic and theistic—and the assumptions they carry with them. By contrasting these two worldviews, I hope to illustrate how spirituality at BYU may be different than spirituality that might be seen at other institutions. I will thus make clear and explicit the worldview that BYU takes in defining spirituality, so as to better understand the larger context of the role that the current study plays in a BYU education.

One main worldview about spirituality is naturalism, what many consider the fundamental worldview of traditional science, or "science's central dogma" (Leahey, 1991, p. 379). An ontological naturalism assumes that all that exists is the natural world (Speck, 2005) and that "the goal of science is to discover the natural laws or principles that govern all of nature, including human nature" (Slife & Gantt, 1999, p. 1456). So, psychologists with this worldview try to understand and explain certain phenomena by studying only natural events and processes (Slife & Reber, 2009). They assume that "supernatural events and processes are not relevant, or do not functionally exist" (Slife & Reber, 2009, p. 64; see also Slife & Whoolery, 2006).

Spirituality in a naturalistic worldview would certainly not involve a God. In fact, there would be no need for belief in the existence of, nor need for, such a higher being. Instead, spirituality is confined to the natural order, where human individuals are the focal point of

authority (Speck, 2005). The general literature on spirituality in higher education supports such a definition. In fact, in the broad literature, spirituality often refers to how individuals create meaning in their lives. A statement by Fried (2001) encompasses such a common naturalistic worldview of spirituality: "We need to discuss spirituality not because God exists, but because we exist and we need to create meaning for ourselves" (p. 277). Although finding meaning in life may be a worthwhile activity, only a naturalistic form of spirituality suggests that "natural" forces—with no thought of God nor religion—are in charge of dictating what that experience is like. Obviously, because part of IFL entails a professor having a personal relationship with God and because this view is saturated in religious attitudes (Hasker, 1992), it would make little sense to practice IFL in an attempt to address naturalistic spirituality in the classroom. It almost goes without mentioning that such a worldview does not fit the mission and aims of a BYU education.

Theism is the opposing worldview to naturalism (Slife & Reber, 2009) and carries with it a set of different metaphysical assumptions (Reber & Slife, 2013; Slife, Reber, & Lefevor, 2012). Slife and Reber (2009) assert that while "naturalism is concerned only with natural events and processes…theism is concerned with supernatural events and processes" (p. 65). They point out that theism assumes that God is involved in the phenomena of existence and that his activity is a necessary condition for a complete understanding of the world. God's activity also plays a necessary role in individuals' lives.

Spirituality in a theistic worldview is commonly found within religion. BYU is founded, supported, and guided by The Church of Jesus Christ of Latter-day Saints (hereto referred to as the Church), a Christian religion (also known as LDS or Mormon) that holds the theistic worldview. Theistic religious spirituality is upheld at this institution, not just by the Church, but also by the faculty and the students through the fulfillment of the university's mission and aims

as explained above. At BYU students are to be taught by professors who strive to follow Christ in their personal lives, who teach their subject matter in the light of the gospel, and who rely on God to connect with others. Students, too, seek to make spiritual connections in their learning. As the Aims of a BYU Education state, "A shared desire to 'seek learning, even by study and also by faith' (D&C 88:118) knits BYU into a unique educational community."

Pilot Studies

Having laid out the context of a unique BYU education and after briefly summarizing the competing worldviews of spirituality, I now come back to the key issue that this dissertation aims to address: the divide between what students want and what faculty provide in regards to spirituality in the classroom. We not only see a gap between student and faculty expectations, interests, and implementation around teaching spirituality in universities and colleges nationwide, we see some disparity at BYU as well (see Wilkins & Birch, 2011). BYU faculty may have a stronger desire to provide spiritually strengthening experiences than faculty at other universities (especially secular ones), but many of them wonder how it can be done (Wilkins & Birch, 2011). Previous research I conducted at BYU has given some answers.

Building on the research done on IFL, teacher immediacy, and relational teaching, two colleagues and I conducted two pilot studies to empirically investigate how applying the three areas of focus would impact students' perceptions of the quality of teaching and their retention of the material taught. These studies were the first of their kind as they attempted to experiment on what was already known through anecdotes, surveys, and interviews. In the first pilot experiment, six Teaching Assistants (TAs) were given material to teach for 30 minutes to a small group of students in a classroom setting. Before they taught, three of the TAs were randomly selected to receive special preparation on how to implement the three areas of focus into their

teaching (the experimental group). The other three TAs constituted the control group and met with a faculty member to discuss basic pedagogical practices (e.g., using PowerPoint versus the whiteboard, asking questions). After each TA taught, the student participants evaluated the teaching quality and then took a retention quiz one week later. We found that there was no significant difference in retention scores between students who were in the experimental group versus those in the control group. The results also indicated that the TAs in the experimental group, on average, were rated higher in teaching quality than the TAs in the control group. This suggested the possibility that preparation on the three areas of focus might improve perceptions of teaching quality without having the deleterious effects on rigor and performance that faculty fear. However, the results were tentative, due to what we concluded to be a rushed and poorly executed preparation session, the TAs' lack of expertise with the teaching material, and small numbers of participants in some of the teaching sessions.

After analyzing the data collected in the first pilot study, we tweaked the method to be more robust, adding a within-subjects element to the already existing between-subjects design. This way we could obtain a baseline measure so that we could evaluate the impact of the preparation by tracking TA change over time. Six TAs taught for 30 minutes on the same material that was used in the first pilot. Afterward, we randomly selected three TAs to receive preparation on how to implement the three areas of focus into their teaching. One week after the preparation session, all six TAs taught the same material a second time to a different group of students. Like in the first pilot, students evaluated the teaching quality after each teaching session and then took a retention quiz one week later. In the second pilot study we found that there was a more pronounced improvement among trained TAs on teaching quality (particularly on spiritual and relational qualities), and there was no detectable loss of learning as a result.

Again, this supported the hypothesis that preparation on the three focuses might improve perceptions of teaching quality without having the damaging effects on rigor and performance that faculty fear. Once I saw this effect could be produced with the TAs from the pilot studies, I wanted to go to the next level and see if actual professors could produce these same kinds of results and do so within different disciplines and with material from their own expertise.

Hypotheses

Previous research indicates that students feel it is "essential" for spirituality to be addressed in the classroom (Astin et al., 2011, p. 31; Lindholm, Astin, & Astin, 2005; Lindholm, Szelényi, Hurtado, & Korn, 2005), but often professors have some reservations about implementing spirituality into their teaching, particularly concerning rigor and quality of performance on learning outcomes (Wilkins & Birch, 2011). Research at BYU has led to the identification of three factors, or areas of focus, that professors can incorporate into their practice to address spirituality and to improve students' educational experience: 1) Professor Self-Disclosure, 2) Intellectual Connections, and 3) Interpersonal Connections. Two pilot studies conducted at BYU showed that teaching assistants had successfully implemented spirituality into their teaching without hurting student performance. The quasi-experiment conducted for purposes of this dissertation was designed to assess the extent to which professors from different disciplines might effectively implement teaching that incorporates each of the three focuses identified. I had four hypotheses in this study, with accompanying rationale, which I explain below.

1. *Teaching that incorporates the areas of focus (separately and combined) will produce higher ratings of teaching quality than teaching that does not incorporate the areas of focus.*

The pilot studies demonstrated that there are some positive outcomes when teaching assistants incorporate the three focuses into their teaching, particularly in regards to perceived higher teaching quality. With improvements to the method of the pilot studies, I thought this result would hold for professors.

2. *Teaching that incorporates the areas of focus (separately and combined) will produce higher retention scores than teaching that does not incorporate the areas of focus.*

Previous studies counter the fear some professors have about the deleterious effects of spirituality on academic rigor and performance. Astin et al. (2011) found that spiritual emphasis enhances academic performance. The literature on immediacy also points to improvement in cognitive learning (Witt et al., 2006). Additionally, the pilot studies demonstrated that incorporation of the three focuses results in no loss in learning performance. I used a two-tailed test to assess significant differences in either improvement or worsening of retention, though I expected to see improvement given previous research findings.

Implicit in these first two hypotheses is the idea that I will see effects on teaching quality and retention with each professor. Since I saw increases in teaching quality with no harm on retention scores with each TA in the pilot study, I expected to see similar results with each professor in the current study.

3. *The combined incorporation of all three areas of focus in the same teaching session will result in an even greater improvement of teaching quality and retention compared to the incorporation of each area of focus separately.*

Reber et al. (in press) found that a combination of three teaching pedagogies (explanatory, meaning-based, relational) was rated most highly and yielded the highest essay test average. I believed a similar benefit would result from combining the three focus areas in the

classroom. In fact, the pilot studies tested whether the collective incorporation of the three focuses had an impact on teaching quality and retention and found support for the former. I thought incorporating all three focuses would have a synergistic effect and that only focusing on one of the focuses would weaken the impact on teaching quality and retention.

As mentioned in Hypotheses 1 and 2, I believed that even focusing on just one factor would result in better outcomes than a control condition. While I believed that the combined incorporation of all three focuses would have the greatest impact on teaching quality and retention, I wanted to see which of the three focuses contributed the most. This led me to my next hypothesis:

4. *When the factors are implemented separately, the third area of focus (Interpersonal Connections) will have a significantly greater effect on teaching quality and retention scores than each of the other two areas of focus.*

Research on teacher immediacy indicates that student ratings and perceptions of teaching quality improve with a relational focus (Moore, Masterson, Christophel, & Shea, 1996; Witt et al., 2006). In fact, studies on teacher immediacy and IFL suggest that a focus on the relationship between professors and students may make the most difference (Sorenson, 1997, Wilson et al., 1975; Witt et al., 2006). The present study postulated that the relational aspect of spirituality may be the most foundational component of the three focus areas. In doing so, this study gives theoretical justification and a test of the theistic hypothesis for relationality.

Method

Brief Overview

The two pilot studies indicated that TAs could successfully be prepared to implement spiritual and relational components in their teaching. While this is compelling, the pilot studies

lacked generalizability and applicability to full-time professors who teach as a career. The pilot experiments also did not address all of the hypotheses because there was only one experimental condition: a combined incorporation of all three areas of focus in the same teaching session. This dissertation seeks for greater applicability to faculty (by the use of faculty participants) and seeks to address all four hypotheses by breaking the experimental condition down into four conditions: one combined condition and three separate conditions (one condition per area of focus).

Participants

Students. I recruited a convenience sample of 203 student participants from undergraduate psychology courses at BYU through class announcements and SONA, a web-based research participant management system used by the Department of Psychology. Of those recruited, four did not complete the second half of the study which contained a retention examination and demographic questions so their data was only included in the analysis of the teacher evaluations. The sample consisted of 97 males (48.7%) and 101 females (50.8%) with one student who did not indicate gender. The average age of the participants was 21.5 (*SD* = 4.60), within a range from 17 to 53 years old. Almost half of the participants (49.7%) were between 21 and 25 years old and 45.5% were between 18 and 20 years old. The majority of participants identified themselves as White/European American (87.9%), with the next largest groups identifying as Hispanic/Latino (3.5%) and Asian/Pacific Islander (3.5%). Other ethnicities represented include Native American/American Indian (1%), Black/African-American (.5%), and Other (1%). Five percent did not identify an ethnicity and 2% identified as Mixed. Each college class was represented: freshmen (35.7%), sophomores (27.5%), juniors (23.4%), seniors (12.9%), and one graduate student (.5%). All student participants received course credit.

Forty-one percent of the student participants (*n* = 83) selected a teaching session randomly assigned to Professor 1 while 59% (*n* = 120) of student participants selected a teaching session randomly assigned to Professor 2. While the number of seats made available in each condition was limited to help make groups equal, I did not have as many student participants as expected as evidenced by the different proportions of students in each condition.

Professors. I recruited two BYU faculty members to participate in this study—one from the College of Life Sciences (specifically, the Exercise Science department; Professor 1) and one from the College of Fine Arts and Communications (specifically, the School of Communications; Professor 2). The selection criteria for the professors were that they must: 1) hold continuing faculty status (roughly equivalent to tenure at other universities), 2) have a record of average student ratings (particularly on "intellectually enlarging" and "spiritually strengthening" items), and 3) teach a course that has content that students from a variety of majors can easily understand (preferably a General Education course). I used these selection criteria because recently-tenured faculty are more likely to be willing and available to participate in this study. Tenure also reduces fear about the effects of student ratings on promotion. Since the professors no longer have to worry about receiving tenure, their schedules become more open and they are less stressed about reaching that tenure point. Also, I assumed that faculty with average student ratings have potential to improve their teaching by emphasizing the three focus areas. Average scores on the spiritually strengthening dimension may also indicate that their teaching is less likely to already incorporate the areas of focus, making it easier to establish a baseline, control condition. Because of their average ratings, using these faculty members can help convince others that preparation on the three focuses may improve educational outcomes. In addition, implicit in my hypotheses is the idea that I will see effects on retention scores and

teaching quality across both professors, despite their academic discipline. Using faculty from different disciplines and comparing their results can counter the argument some faculty make—at least with the two disciplines represented in this study—that spiritual teaching does not fit with the subject matter of their discipline.

One of the research personnel (Alan Wilkins) and I identified faculty members at BYU who fit the criteria and emailed them an invitation to have a conversation about what they do in their classes regarding faith and learning. We met with willing candidates and gauged their current viewpoint and integration of faith and learning in their teaching. We then explained the background and methods of this research study. After the interviews I selected two candidates based off who would be an appropriate fit for this study. Fit was measured by whether they met my established criteria for participation in the study (some of the faculty members only taught advanced courses meant for students in their particular department and were therefore excluded from consideration) as well as their availability, interest, and willingness to participate.

Both participants were married, Caucasian men. At the time of participation Professor 1 was 38 years old and had a Doctor of Philosophy degree. Professor 2 was 42 years old and had a Juris Doctor degree.

Measures

Reading excerpts. Each faculty member provided relevant reading excerpts on topics of their choice from an area of their own expertise. These topics were ones that could easily be taught to students from a variety of majors. Professor 1 provided a reading excerpt from the Linear Kinematics chapter of McGinnis' (2004) exercise science textbook. Topics covered in the excerpt included definitions, equations, and examples of position, displacement, speed, velocity, and acceleration with discussions on the importance of those concepts. Professor 2

provided a reading excerpt from his own writing on the First Amendment. Topics covered in this excerpt included sources of law and the role of precedent, the United States Constitution, a summary of the history of the First Amendment, the meaning of the First Amendment today, and values served by the Speech Clause.

I worked closely with the two faculty members and used independent raters (n = 12) to ensure that the reading materials were comparable in difficulty, amount of content, and length. The Linear Kinematics excerpt was 6.5 pages long and the First Amendment excerpt was 5.5 pages long, but there was no difference in the time to read the material as determined by having the independent raters read both excerpts and record the time it took them to finish each one (Linear Kinematics: M = 15.83 minutes, SD = 2.21; First Amendment: M = 17.35 minutes, SD = 2.99; $t(8.1) = -1.067$, $p = .167$). In order to further equate the reading material, I asked them how enjoyable, exciting, inspiring, boring, arousing, emotionally-moving, entertaining, interesting, difficult to comprehend, uplifting, applicable, clear, and informative the materials were on a five-point scale (*not at all* to *very much*). Means and standard deviations for these items are reported in Table 1. Analyses showed that the reading materials did not differ on these items.

Teacher Rating Questionnaire. The Teacher Rating Questionnaire (see Appendix A) measures teaching quality through students' perceptions of their experience and the professor's capabilities. This evaluation is comprised of items adapted from four sources: 1) Feldman and Prohaska's (1979) work on students' expectations of the instructor; 2) Silva and colleague's (2008) research on online evaluations (i.e., RateMyProfessors.com) of psychology professors; 3) RateMyProfessors.com; and 4) additional items I developed to assess other aspects of teaching more particular to the areas of focus.

Table 1

Comparison of Reading Excerpts on 13 Items

	Linear Kinematics		First Amendment		
Item	Mean	*SD*	Mean	*SD*	Mean Difference
Enjoyable	2.75	1.09	3.50	1.12	0.75
Exciting	2.00	1.00	2.50	1.26	0.50
Inspiring	1.50	0.71	2.83	1.21	1.33
Boring	3.00	1.22	2.17	1.34	-0.83
Arousing	1.63	0.86	2.83	1.46	1.20
Emotionally-moving	1.25	0.66	2.50	0.96	1.25
Entertaining	2.38	1.11	2.17	1.07	-0.21
Interesting	2.88	1.05	4.00	1.41	1.12
Difficult to comprehend	2.13	0.78	2.17	1.07	0.04
Uplifting	1.13	0.33	2.50	1.12	1.37
Applicable	3.50	1.00	3.83	1.34	0.33
Clear	3.50	1.22	3.67	0.75	0.17
Informative	3.75	1.09	4.50	0.76	0.75

Note. I conducted *t*-tests for each item using the means between the two reading materials. Given that multiple comparisons were analyzed, which might inflate Type I error, the alpha cutoff was subjected to a Bonferroni correction ($p < .05/13$ comparisons = critical cutoff of $p < .004$). Results indicate no differences.

Sample items from Feldman and Prohaska (1979) assess lesson difficulty, interest, and effectiveness as well as professor competence, intelligence, and likability. Items from Silva et al. (2008) evaluate professor characteristics such as organization, enthusiasm, knowledge, and ability to stimulate interest. On RateMyProfessors.com, students rate professor helpfulness and clarity. The website also asks students to rate how easy students feel it is to succeed in the professor's class. Items from these three sources were selected because they are useful in gauging students' perceptions of various qualities about their professor and the lesson. These three sources were also used in previous research that used similar methods to test students' expectations of their professors (Ridge & Reber, 2011). These researchers showed through Exploratory Factor Analysis that a measure which combines these three sources has good

reliability across three factors: α = .93 (Teacher Pedagogical Skill), .86 (Teacher Personal Attributes), and .80 (Lecture Quality). Additional items I added to this evaluation (i.e., items 17-33, 54-59) measured aspects such as professor spirituality, working relationship with students, care for learning, and overall quality. These items were generated with research personnel and selected for inclusion based off their perceived and validated (via the pilot studies) match with the three areas of focus.

The Teacher Rating Questionnaire is composed of 59 Likert-type items on three subscales. The first subscale asks students to rate how much they agree or disagree with statements about the professor and the lesson. The second subscale has students rate the professor on the extent to which they agree that he possesses certain qualities that exhibit corresponding behaviors. Lastly, students rate the overall lesson and professor from *poor* to *excellent*. Six items (i.e., items 2, 3, 5, 10, 14, and 22) were reverse scored. Reliability tests on three factors of this questionnaire—as determined by an Exploratory Factor Analysis—indicate good reliability: α = .97 (General Teaching Skills), .94 (Spirituality), and .89 (Openness and Respect), which is consistent with the pilot studies.

In the two pilot studies the Teacher Rating Questionnaire originally used five-point scales on each subscale (first subscale: *disagree* to *agree*; second subscale: *never* to *always*; third subscale: *poor* to *excellent*). For the current study, I developed a new rating scale in order to reduce potential ceiling effects and increase the variability. Reducing ceiling effects and increasing variability in a questionnaire helps to increase statistical validity (Morling, 2014).

In order to reduce the ceiling effects, I changed the scale from a five-point scale with two positively worded options to a nine-point scale with six positively worded options. Positively packed or unbalanced scales (i.e., more positive options than negative options in the scale) can

result in lower mean scores (Dunham & Davidson, 1991), thus reducing the chance of ceiling effects. Dawes (2008) found that a ten-point scale (compared to five- and seven-point scales) resulted in slightly lower mean scores.

In order to increase the variability in responses, I added additional instructions to inform respondents that professors may not necessarily merit a high or low rating in every statement. I also used three scale labels in hopes of limiting satisficing. Satisficing is the act of accepting an available option as satisfactory, or doing the minimum requirement to achieve a particular result. In relation to this study, satisficing could occur when a respondent becomes fatigued from expending cognitive effort and thus gives the professor the same rating on all items. As the respondent becomes fatigued, he or she expends the least amount of effort required to finish the questionnaire, marking "agree", for example, on all items. Satisficing is generally thought to be related to task difficulty, respondent ability, and respondent motivation (Krosnick, 1991).

Changing the scale can reduce the cognitive load on the respondents, help prevent fatigue when answering survey questions, and help reduce satisficing. By limiting satisficing I also limit acquiescence. Although I cannot tell whether ceiling effects from the pilot studies were due to satisficing or acquiescence, I believed that the new scale would reduce cognitive load, mitigate ceiling effects, and increase respondents' ability to make accurate decisions while still allowing for variability. In fact, as the results will indicate, I did find that compared to the results of the pilot studies, the new scale used in the current study did reduce ceiling effects and increased variability in responses.

Retention Examination. A retention examination for each professor's chosen topic (see Appendices B & C) was used to measure students' retention of the material taught. In order to ensure equivalency across tests, each exam included the same number of multiple-choice

questions (nine) all of which had similar proportions of easy, medium, and hard levels of difficulty as determined by professors (some questions came from a test bank). Each exam also included one essay question of moderate difficulty level also determined by professors. I worked closely with the two faculty members to ensure that the tests were comparable in difficulty, amount of content, and length. I also had a group of independent raters (n = 12) read the materials and then take the exams to help ensure that the two topics were comparable in difficulty. After administration of the exams, student scores showed that indeed questions of easy, medium, and hard levels of difficulty were similarly proportionate and means and variances were also similar. The mean for the Linear Kinematics exam was 7.74 (SD = 2.17) and the mean for the First Amendment exam was 7.55 (SD = 2.15). The exams also had a similar ratio of question types (e.g., definition, conceptual, application).

Professor Self-Evaluation. For exploratory purposes, I asked the professors eight open-ended questions about their thoughts and experiences regarding their participation (see Appendix D). These questions mainly focused on the professors' experiences with the three areas of focus.

Procedure

Procedure for students. Each student participant signed up for one teaching session through SONA. While signing up, the students were told that the study was being conducted to explore teaching practices that improve students' educational experience. They were told they would be taught by a teacher and fill out questionnaires about their experience. In the study description for recruitment, the students were asked to arrive promptly and were made aware of the basic outline of what would happen (see Appendix E). They did not know which faculty member would teach their session nor which of the randomly assigned conditions they would be a part of. The number of seats available in each condition were limited to 35 to help make the

groups equal, but as mentioned previously, I did not have as many student participants as expected resulting in a different proportions of students in each condition.

At the start of each study condition, the students arrived at a classroom on the BYU campus prior to the faculty member and were again told why they were participating and what they would be doing. After signing the consent forms, they read the provided material on the topic that the faculty member would teach. Each student was given the same amount of time to read the excerpts (approximately 20 minutes). The excerpts were collected before the faculty member entered the room to teach.

The students then participated in a 30 minute lecture taught by the professor. After the professor taught and left the room the student participants rated him on the quality of teaching using the Teacher Rating Questionnaire (see Appendix A). Exactly one week after the session students received an email linking them to the Retention Examination (see Appendices B & C) and demographic questions online.

Procedure for professors.

Control condition. In order to show that professors can be prepared to incorporate the three areas of focus, I established a baseline (pre-preparation control) teaching condition before the professor went through a pedagogical preparation session. In this condition each professor arrived at the designated classroom and taught a group of students for 30 minutes where no intentional spiritual or relational elements were emphasized. Prior to teaching the control condition they were instructed to treat the teaching session as if they were giving a lecture as part of a job interview at a state university. This was to reduce the chance of incorporating the three focuses into their teaching. They were informed that for each teaching session they would be teaching a group of students from various majors.

Preparation. About one week after teaching in the control condition, the faculty members attended a 60-minute preparation session where they were educated on how to incorporate the three areas of focus into their teaching. Specifically, they were prepared to teach in four experimental conditions: 1) a class where they would implement only the first area of focus (Professor Self-Disclosure), 2) a class where they would implement only the second focus (Intellectual Connections), 3) a class where they would implement only the third focus (Interpersonal Connections), and 4) a class where they would combine the three focuses simultaneously. Two fellow research investigators (Jeffrey Reber and Alan Wilkins) and I led the preparation session. We met with one professor at a time making sure the preparation was consistent across professors but also tailored to their teaching material.

Before the preparation session, the professors received a Preparation Handout (see Appendix F) that was designed to help them think of ways in which they could create a community of learning (i.e., show the students that they care about them and about learning the material) and make spiritual connections to particular topics in the material. The handout also prompted them to think of personal examples or stories that would show how they have experienced a particular concept (related to the material) and that would reveal something important about their values, weaknesses, and/or spiritual commitments. They were instructed to write their ideas down on the handout.

It is important to note that as the ones leading the preparation session we did our own preparations as well. Our preparations began with the pilot studies where we met together often via email and in person to come up with ideas on how we could best prepare the TAs. We reviewed a video that highlighted some of the previous research about faith and learning at BYU (found on http://faithandlearning.weebly.com), studied the teaching materials, created

preparation outlines and agendas, and practiced administering our allotted parts of the preparation. In the pilot studies, we also came prepared to share our own examples and ideas if needed so that we could help the TAs think of ways in which they could implement the areas of focus. After each pilot study we discussed how we could improve the preparation and we fine-tuned the aspects that we felt needed work. One of the aspects of the preparation that we changed from the pilot studies was to spend more time doing the preparation—one hour instead of 30 minutes. We also placed higher emphasis on having the professors come prepared with and generate their own ideas. We spent more time having them share and write down their own examples so that they could feel more prepared to teach each experimental condition. Moreover, we watched the video recordings of the control conditions with the purpose of helping the professors think of specific things they should or should not do in each experimental condition.

To begin the preparation session itself, the professors first watched the faith and learning video. They then discussed with us—the preparers—how to use their ideas from the handout to teach the four different experimental conditions. During this discussion, the professors wrote ideas in their lecture notes (one set of notes for each experimental condition) of specific things they could do in each experimental condition. The key focus of the preparation was to discuss the principles of each focus area and then have the professors apply the principles to their teaching style and material so it would be genuine. This discussion also helped the professors know how to teach according to one specific area of focus while not teaching according to the other areas of focus, so as to make each teaching session—each experimental condition—true to its purpose. While they were necessarily aware of the different areas of focus, the professors were not aware of the hypotheses. Before the conclusion of the preparation session the

professors were given an assignment to keep working on their ideas for integrating the focus areas into their teaching.

Experimental conditions. One week after the preparation, each faculty member taught two groups of students—covering two experimental conditions. Then, about one week after that, they taught two more groups of students—rounding off the last two experimental conditions. The order of single-focus teaching conditions was counterbalanced between the two professors to correct for any practice effects. I had the professors teach the single focus conditions first so that they could have a better understanding of what I was expecting them to do in the more difficult combined condition. Including the control group, a total of five groups, or five teaching sessions, were taught by each faculty member (ten total teaching sessions between the two faculty members). Each teaching session lasted 30 minutes. If the professors had to teach two sessions in one day I allowed for 30 minutes between teaching sessions to give them enough time to switch gears and get ready to teach a class with a different focus. When the professors finished teaching the last experimental (combined) condition, they spent about 10 minutes answering the Evaluation for Professors (see Appendix D).

Procedure for manipulation check. All participants were aware that the teaching sessions were videotaped. This was done so that a check on the pedagogical uniqueness of each session could be performed. After the teaching sessions concluded, two trained independent raters watched each video and recorded the time (in seconds) each professor spent demonstrating each of the areas of focus. Discrepancies between the raters were resolved by a trained independent arbitrator. I then calculated the percent of time each professor spent in each condition on each focus.

This served two main purposes. First, it acted as a manipulation check. I wanted to ensure that the professors adhered to their preparation. In other words, I wanted to make sure they changed their teaching significantly enough from the control condition and they did not teach according to the other conditions. For instance, I wanted to make sure that the professors did not implement Intellectual Connections when they were supposed to concentrate on Interpersonal Connections. Second, this gave me a base rate for each of the professors. How much time each professor spent implementing the focuses in the control condition—unbeknownst to them—gave me an idea of what I should expect to see in the experimental conditions. For example, if Professor 1 naturally used Interpersonal Connections in the control condition, I should expect to see some use of Interpersonal Connections seep in to the Professor Self-Disclosure and Intellectual Connections conditions.

Results

Manipulation Check

Table 2 indicates the percent of time each professor spent in each condition on each focus area. The average teaching time across all conditions was 30 minutes and 26 seconds. Professor 1 averaged 30 minutes and 35 seconds (SD = 72 seconds) while Professor 2 averaged 30 minutes and 17 seconds (SD = 35 seconds). The difference between these times is not significant, $t(8) = 0.509$, $p = .625$, indicating that neither professor had an advantage in available time to spend implementing one of the focuses.

In the control condition, both professors naturally used some of Professor Self-Disclosure and Interpersonal Connections. Professor 1 spent 15.82% of the teaching time incorporating those focuses while Professor 2 only spent 3.1% of the time doing so.

Table 2

A Comparison Between Professors of the Time Spent in Each Condition Implementing the Three Areas of Focus

	Professor 1		Professor 2		
Area of Focus by Condition	Time (seconds)	% of time	Time (seconds)	% of time	Percentage Difference
Control	1915		1775		
Professor Self-Disclosure	132	6.89	18	1.01	5.88*
Intellectual Connections	0	0.00_a	0	0.00_b	0.00
Interpersonal Connections	171	8.93	37	2.08_c	6.85*
Total focuses	303	15.82	55	3.10	12.72*
Overlap time	0	0.00	0	0.00	0.00
Experimental 1	1797		1822		
Professor Self-Disclosure	**322**	**17.92**	**1213**	**66.58**	**48.66***
Intellectual Connections	35	1.95	88	4.83	2.88*
Interpersonal Connections	331	18.42	49	2.69_c	15.73*
Total focuses	628	34.95	1332	73.11	38.16*
Overlap time	60	3.34	18	0.99	2.35
Experimental 2	1776		1801		
Professor Self-Disclosure	61	3.43	551	30.59	27.16*
Intellectual Connections	**322**	**18.13**	**1108**	**61.52**	**43.39***
Interpersonal Connections	217	12.22	141	7.83	4.39*
Total focuses	557	31.36	1426	79.18	47.82*
Overlap time	43	2.42	374	20.77	18.35*
Experimental 3	1777		1817		
Professor Self-Disclosure	281	15.81	601	33.08	17.27*
Intellectual Connections	0	0.00_a	0	0.00_b	0.00
Interpersonal Connections	**539**	**30.33**	**723**	**39.79**	**9.46***
Total focuses	798	44.91	1312	72.21	27.30*
Overlap time	22	1.24	12	0.66	0.58
Experimental 4	1912		1871		
Professor Self-Disclosure	**192**	**10.04**	**615**	**32.87**	**22.83***
Intellectual Connections	**112**	**5.86**	**651**	**34.79**	**28.93***
Interpersonal Connections	**251**	**13.13**	**496**	**26.51**	**13.38***
Total focuses	514	26.88	1630	87.12	60.24*
Overlap time	41	2.14	132	7.06	4.92*

Note. The proper focuses for each condition are in boldface for convenience. "Overlap time" indicates the time the professor spent implementing more than one focus at a time. "Total focuses" is calculated by adding up the time the professor spent on each focus and subtracting the overlap time. Percentages sharing a common subscript denote no statistical difference between the control and experimental conditions at $\alpha = .001$ according to a difference between proportions test; all other percentages are significant.
*$p < .001$.

As expected, the professors' natural, baseline ways of teaching (without preparation) carried through each experimental condition to some extent as can be seen in Professor Self-Disclosure and Interpersonal Connections having at least a small percentage of time spent on them. This is especially the case for Professor 1.

The pattern demonstrated in the table shows that the manipulation mostly worked, particularly for Professor 2. Professors spent significantly more time focusing on the proper emphases in the corresponding conditions compared to the control condition as determined by two-tailed difference in proportions tests. For example, while neither professor spent time implementing Intellectual Connections in the control condition, both spent the most time on that focus in the second—the proper—experimental condition (18.13% for Professor 1 and 61.52% for Professor 2). However, the data show that the manipulation check did not work as cleanly as expected. The professors taught using some non-focus areas in the improper experimental conditions. For example, even though Professor 1 spent significantly more time implementing focus one in the proper condition than in the control condition (17.92% versus 6.89%), he implemented Interpersonal Connections (18.42%) more than Professor Self-Disclosure.

The numbers also reveal that Professor 2 incorporated the preparation much more successfully than Professor 1: he spent more time emphasizing the focuses within and across all four experimental conditions, and he did so more cleanly. That is, he placed more emphasis on the desired focus in each condition. In the Professor Self-Disclosure, Intellectual Connections, and Interpersonal Connections conditions, Professor 2 spent 66.58%, 61.52% and 39.79% (respectively) of the time implementing the correct focus while Professor 1 only spent 17.92%, 18.13% and 30.33% (respectively) of the time implementing the correct focus. This is especially revealing given the baseline rates for each professor because it shows that Professor 2

implemented spirituality more than Professor 1. The results from the combined condition indicate that the professors successfully implemented each focus in this condition more than in the control condition. However, here again, Professor 2 spent more time implementing these focuses (87.12% of the time) than Professor 1 (26.88% of the time) and maintained a better balance of all three focus areas. Because there are significant differences between professors on time spent on the focuses, I made it a point to observe and address the significant differences in the analyses that follow by including professor as an independent variable. I later discuss the implications of these differences.

Data Screening

The examination and evaluation data were collected in different formats at different times, making it impossible to link those data together. Since the nature of the collection did not allow for the data to be screened and analyzed together, they are reported separately. All data were screened and analyzed using Stata 14.

Demographic analyses. There were no significant differences in demographic data across conditions. However, I found significant differences in the proportions of students between professors on two demographic qualities: gender and college class. Professor 1 had significantly more male student participants (57.5%) than did Professor 2 (43.2%), $\chi^2(1, N = 198) = 3.89$, $p < .05$. Notably, the ratio of male to female undergraduate students at BYU in 2014 was 55% to 45%, respectively. This matches well with Professor 1's proportion of students but not with Professor 2's proportion. Professor 1 also had significantly more juniors (33.3%) than Professor 2 (16.7%) while Professor 2 had more freshmen (43.1%) than Professor 1 (24.6%), $\chi^2(4, N = 171) = 9.75$, $p < .05$. Currently, demographic information for college class at BYU is unavailable. These results indicate that randomly assigning the order of conditions was not

sufficient to balance the differences in gender and college class across professors. As a result, Professor 1's student participants were more likely to be male and further along in their education than Professor 2's students.

Evaluation data screening. In order to prepare the evaluation data for factor analysis I screened the data for outliers, missing data, issues with nonlinearity and heteroscedasticity, and normality. The only potential issue seemed to be in normality. A Shapiro-Wilk test, $W = .983$, $p = .016$, and a multivariate test for univariate normality, joint $\chi^2 (2) = 9.35$, $p = .009$, indicated that the data were probably not normally distributed, likely due to a kurtosis issue. Upon applying a recommended square root transformation, the data became even less normally distributed ($W = .975$, $p = .001$), so I did not make any changes. The minimum amount of data for factor analysis was satisfied, with a final sample size of 203, with over 200 cases in each of 59 items.

Examination data screening. The examination data were also screened for outliers, missing data, issues with nonlinearity and heteroscedasticity, and normality. No issues were found so the data remained unchanged.

Data Reduction

I first analyzed the Teacher Rating Questionnaire using confirmatory factor analysis (CFA) in order to see if the items fit onto three factors in the same way as they did in the pilot studies. Unfortunately, this model had poor fit ($\chi^2 (854) = 2394.72$, $p < .001$; RMSEA = .094; CFI = .786; TLI = .774) even after applying suggested modification indices. Consequently, I used exploratory factor analysis (EFA) to see which items grouped together and to verify whether those factors matched with the three focus areas.

I tested the assumptions for EFA and found that Barlett's test of sphericity was significant, $\chi^2 (1711) = 10247.67$, $p < .001$, indicating that there was sufficient correlation among

the items on the Teacher Rating Questionnaire. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was marvelous, KMO = .942, above the recommended value of .6, indicating that the degree of common variance was high. In other words, any extracted factors in an EFA would account for a large amount of variance. Given these overall indicators, factor analysis was conducted with all 59 items.

Principal axis factor analysis was used because the primary purpose was to analyze the shared variance among observed variables in order to find which items grouped together and to compare whether those factors matched with the three focus areas. An initial analysis showed that the first factor explained 69% of the variance, the second factor 12% of the variance, and a third factor 8% of the variance. The fourth, fifth, and sixth factors had eigenvalues of just over one, each factor explaining 4% of the variance. In examining the factor solutions, I found that the fourth and fifth factors had no loadings above .4; the sixth factor had only one. A scree plot showed a leveling off of eigenvalues after three factors. Given this information and the theoretical support, I chose to retain three factors in the principal axis analysis. I also chose to use a promax rotation because, as explained above, the factors are considered to be correlated.

During several steps, a total of nine items (i.e., items 12, 14, 33-36, 44, 46, 56) were eliminated because they failed to meet a minimum criteria of having a primary loading factor of .4 or above or cross-loaded at .4 or above.

A principal axis factor analysis of the remaining 50 items using promax rotations was conducted, with the three factors explaining 89% (100% when three factors are forced) of the variance. All items had primary loadings over .4 and most communalities were above .5. The factor loading matrix for this final solution is presented in Table 3.

Table 3

Factor Loadings and Communalities Based on a Principal Axis Analysis With Promax Rotation for 50 Items from the Teacher Rating Questionnaire (N = 203)

Item	Factor 1	Factor 2	Factor 3	Communality
Item1: This teacher seemed to be helpful.	.57			.53
Item2: This teacher presented the material in an unclear way (reverse scored).	.50			.36
Item3: It would be difficult to succeed in a class if it were taught by this teacher (reverse scored).	.59			.51
Item4: Overall, this is a quality teacher.	.67			.64
Item5: This lesson was difficult to understand (reverse scored).	.57			.37
Item6: This lesson was interesting.	.56			.46
Item7: The lesson materials and activities were effective.	.83			.60
Item8: This teacher was competent.		.43		.45
Item9: This teacher was intelligent.		.56		.41
Item10: I did not like the way this teacher taught (reverse scored).	.68			.54
Item11: I would recommend this teacher to a friend.	.80			.72
Item13: This teacher was enthusiastic.	.67			.51
Item15: This teacher did a good job stimulating students' interest.	.87			.72
Item16: If this teacher were to teach a class, I would sign up to take a class from them next semester.	.74			.64
Item17: This teacher was spiritually inspiring.			.78	.76
Item18: This teacher has an ongoing personal relationship with God.			.80	.68
Item19: This teacher integrated the gospel into the subject.			.88	.71
Item20: This teacher developed a good working relationship with the students.	.70			.57
Item21: If I had this teacher for a class, I could develop good rapport with him/her.	.50			.56
Item22: I would feel uncomfortable approaching this teacher for help (reverse scored).		.46		.25
Item23: This teacher was open to questions.		.52		.38

(continued)

Item	Factor 1	Factor 2	Factor 3	Communality
Item24: This teacher allowed us to connect with each other.	.73			.53
Item25: If I had this teacher for a class, I could get closer to the other students.	.74			.54
Item26: This teacher has interest in student learning.	.48			.58
Item27: This teacher gave opportunities for us to get help.	.43			.41
Item28: This teacher responded to students respectfully.		.77		.58
Item29: This teacher explained concepts effectively.	.57			.53
Item30: This teacher promoted active student involvement.	.72			.51
Item31: My testimony was strengthened because of this teacher.			.72	.67
Item32: My intellectual skills were developed because of this teacher.	.71			.58
Item37: Creative and Interesting	.71			.62
Item38: Effective Communicator	.51			.47
Item39: Enthusiastic About Teaching and About Topic	.67			.61
Item40: Flexible/Open-Minded		.51		.62
Item41: Good Listener		.66		.62
Item43: Humble		.57		.47
Item45: Prepared		.41		.49
Item47: Promotes Class Discussion	.72			.48
Item48: Promotes Critical Thinking/ Intellectually Stimulating	.82			.67
Item49: Punctuality/Manages Class Time	.44			.40
Item50: Rapport	.83			.66
Item51: Respectful		.94		.73
Item52: Sensitive and Persistent	.46			.47
Item53: Understanding		.67		.52
Item54: Spiritual			.89	.80
Item55: Spiritually intellectual			.90	.84
Item57: Genuinely Cares for Learning	.42			.59
Item58: Overall lesson rating	.85			.74
Item59: Overall teacher rating	.71			.75

Note. Factor loadings < .40 are omitted. The table is divided into three section based on the corresponding sections on the Teacher Rating Questionnaire. For more item details see Appendix A. Items that did not load above .40 on any factor are not displayed.

The factor labels did not follow as closely with the three areas of focus as desired. Because the first factor consisted of so many items (33) covering a wide range of teaching qualities it was labeled as "General Teaching Skills". The second factor, consisting of 11 items, was labeled as "Openness and Respect". The third factor, consisting of six items, was the most clear. It was labeled as "Spirituality" due to the nature of the items that loaded on to the factor. Internal consistency for each of the factors was examined using Cronbach's alpha. The alphas were high: .97 for General Teaching Skills, .89 for Openness and Respect, and .94 for Spirituality.

Composite scores were created for each of the three factors based on the mean of the items in each factor. Higher scores indicated higher ratings on each factor. The average rating for General Teaching Skills was 5.80 (SD = 1.51), for Openness and Respect 7.29 (SD = 1.16), and for Spirituality 5.91 (SD = 2.07). Once these factors were determined and composite scores were created, each factor was ready to be screened in preparation for further analyses.

Data screening by factor. Once the factors were determined, I screened the data in each factor to ensure that the data were ready for further analysis. There was one outlier for General Teaching Skills in the combined condition for Professor 2 that was an extreme score on the low end, thus altering the distribution. To fix this, I fenced the outlier using the criterion of the median plus or minus two interquartile ranges. This brought the low score (2.68) up to a more operational score of 2.969. Taking care of this outlier helped the normality, even though the distribution still had an issue with kurtosis. The recommended transformation did not make the distribution more normal, so no further changes were made to the data in the General Teaching Skills factor.

A similar issue arose with the Openness and Respect factor. There were two outliers found in this factor. One was in the Professor Self-Disclosure condition for Professor 2. This was an extreme score (4.23) on the low end and was fenced to 4.454 using the criterion of the median plus or minus two interquartile ranges. The other outlier was a low score (3.91) in the Intellectual Connections condition for Professor 2. It was fenced to 3.955 using the same criterion. These changes made the data more normally distributed, yet there still was a negative skew. The recommended transformation was an identity transformation, so no transformations were used.

Lastly, there were a few outliers in Spirituality. There were two outliers in the control condition for Professor 1. These high scores (8.83 and 5.67) were fenced to 5.334. There were also three low scores (4.0, 4.0, and 4.17) in the Professor Self-Disclosure condition for Professor 2 which were fenced to 4.25. These changes made the data more normally distributed, yet there were still issues with normality. The recommended transformation was an identity transformation, so no transformations were used. Once these data were screened, all the data were ready for analysis.

Significance Tests

Teaching quality. Basic descriptive data for the evaluation data for each group are listed in Table 4. This table gives a snapshot of the ratings across the three extracted factors for each professor and condition. Table 4 also displays the rating scores for the experimental conditions as they differ from the control condition within each professor.

Table 4

Basic Descriptive Statistics and Difference Scores for Evaluation Data Across the Three Extracted Factors

	Professor 1				Professor 2			
Condition	*n*	GTS	OR	S	*n*	GTS	OR	S
Control	19	5.86[a]	6.99	3.12	18	4.38	7.09	2.75
		(1.44)[b]	(1.46)	(1.18)		(1.21)	(1.13)	(1.30)
Professor Self-Disclosure	12	7.49	8.09	6.56	28	4.71	6.96	7.04
		(0.87)	(0.95)	(1.09)		(1.26)	(1.08)	(1.36)
		1.63[c]*	**1.10***	**3.44**		**0.33**	**-0.13**	**4.29**
Intellectual Connections	29	6.40	7.40	6.20	27	5.18	7.20	6.93
		(1.42)	(1.17)	(1.63)		(1.22)	(1.19)	(1.28)
		0.54	**0.41**	**3.08**		**0.80**	**0.11**	**4.18***
Interpersonal Connections	8	6.67	7.35	5.15	17	6.81	7.91	5.86
		(1.30)	(0.86)	(1.91)		(1.10)	(0.95)	(1.63)
		0.81	**0.36**	**2.03**		**2.43***	**0.82**	**3.11**
Combined	15	6.30	7.35	5.57	30	5.87	7.17	7.38
		(1.25)	(0.96)	(1.69)		(1.22)	(1.15)	(1.23)
		0.44	**0.36**	**2.45**		**1.49***	**0.08**	**4.63***

Note. Difference scores are in boldface. The asterisks (*) indicate significant differences in difference scores between professors as determined by Tukey's HSD tests. GTS = General Teaching Skills. OR = Openness and Respect. S = Spirituality.
[a]Mean. [b]SD. [c]Mean difference from the control condition
*$p < .05$

In order to compare the extent to which incorporating the areas of focus improved upon student participants' perceptions of teaching quality I ran a 5 (condition) X 2 (professor) multivariate analysis of variance (MANOVA) with the three factors from the factor analysis as dependent variables to compare evaluations of professor quality both between conditions and between professors. The main effect for condition was significant, $F(4, 193) = 18.11$, $p < .0001$, partial $\eta^2 = .270$. The main effect for professor was also significant, $F(1, 193) = 33.07$, $p < .0001$, partial $\eta^2 = .342$. There was also a significant interaction between the two independent variables, $F(4, 193) = 4.12$, $p < .0001$, partial $\eta^2 = .079$, indicating that the effect of condition on perceptions of teaching quality depends on the professor.

Two-way analysis of variance (ANOVA) follow-ups for each factor (dependent variable) revealed significant differences between both condition and professor for General Teaching Skills and Spirituality but not for Openness and Respect.

General Teaching Skills. On the General Teaching Skills variable the main effect for condition was significant, $F(4, 193) = 6.51$, $p < .001$, partial $\eta^2 = .119$, indicating a medium to large effect. The main effect for professor was also significant, $F(1, 193) = 36.30$, $p < .0001$, partial $\eta^2 = .158$, indicating a large effect. In this case, Professor 1 had significantly higher ratings on General Teaching Skills than Professor 2. There was a significant interaction between the two independent variables, $F(4, 193) = 5.91$, $p < .001$, partial $\eta^2 = .109$, indicating a medium to large effect.

Figure 1 graphically displays the evaluation ratings across professors and conditions on the General Teaching Skills factor. Post hoc tests—namely, Tukey's honest significant difference (HSD) pairwise comparisons—indicated that there were a few significant differences for this variable. Average professor evaluations for the control condition ($M = 5.14$) were

significantly lower than the Interpersonal Connections (*M* = 6.76) and combined (*M* = 6.02) conditions, lending partial support for Hypothesis 1 which says that there will be higher ratings on teaching quality in the experimental conditions than the control condition. The average evaluations for the Interpersonal Connections condition were also significantly higher than the Professor Self-Disclosure (*M* = 5.54) and Intellectual Connections (*M* = 5.82) conditions, lending some support for Hypothesis 4 which says that there will be higher ratings on teaching quality in the Interpersonal Connections condition compared to the Professor Self-Disclosure and Intellectual Connections conditions. No support for Hypothesis 3 was found on the General Teaching Skills quality because there were not higher ratings on teaching quality in the combined condition compared to the other experimental conditions.

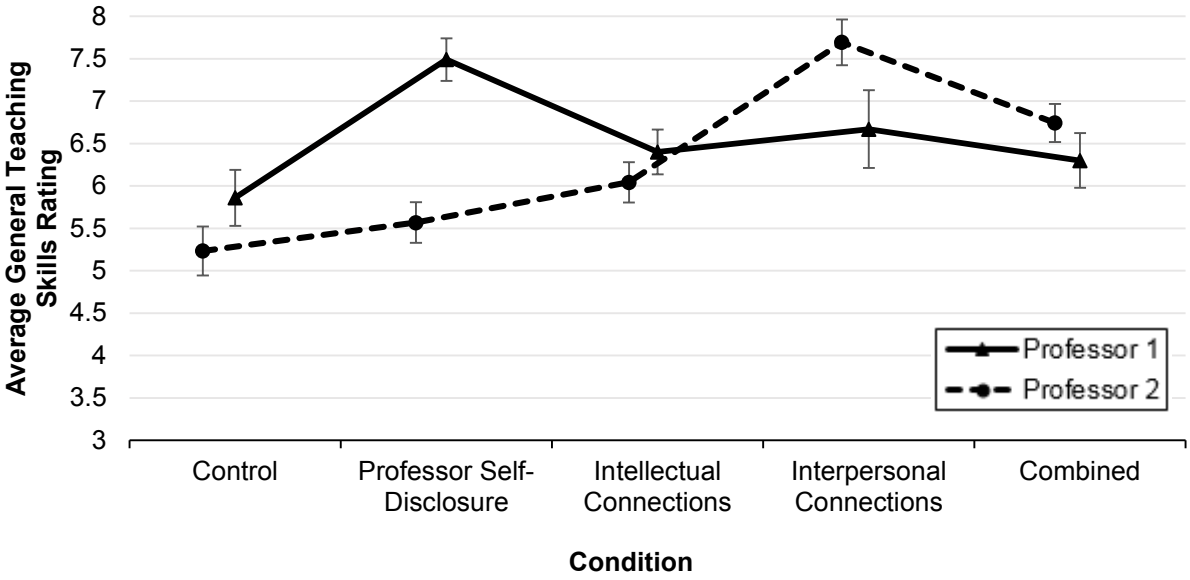

Figure 1. Average student evaluation ratings on the General Teaching Skills variable across professors and conditions. Error bars represent standard errors. Points are offset horizontally so that error bars are visible. Recall that the order of the first three experimental conditions were randomized; therefore, this is does not reflect chronological order.

Post hoc tests were also done to compare the means for professor in each condition on the General Teaching Skills variable. Tukey's HSD tests indicated that the mean scores for evaluations of Professor 1 in the control, Professor Self-Disclosure, and Intellectual Connections conditions were significantly higher than the mean scores for Professor 2 in those conditions.

Since Professor 2 differed from Professor 1 in the control condition, this suggests he is not as good of a teacher—at least in regards to perceived General Teaching Skills—as Professor 1 was initially. To take this into account, I did follow-up analyses comparing the relative change in teaching quality ratings from the control to experimental conditions within each professor. Table 4 displays these differences. Post hoc analyses showed that Professor 1 had significantly more positive change in the Professor Self-Disclosure condition, but significantly less change in the Interpersonal Connections and combined conditions compared to Professor 2.

Spirituality. The main effect for condition on the Spirituality factor was significant, $F(4, 193) = 48.58$, $p < .0001$, partial $\eta^2 = .502$, indicating a very large effect. The main effect for professor was also significant, $F(1, 193) = 9.79$, $p = .002$, partial $\eta^2 = .048$, indicating a small to medium effect. In this case, Professor 2 had significantly higher ratings on Spirituality than Professor 1. There was a significant interaction between the two independent variables, $F(4, 193) = 2.94$, $p = .022$, partial $\eta^2 = .057$, indicating a small to medium effect.

Figure 2 graphically displays the evaluation ratings across professors and conditions on the Spirituality factor. For this variable, post hoc comparisons using Tukey's HSD tests indicated that the mean score for evaluations for the control condition ($M = 2.94$) was significantly lower than all of the experimental conditions ($M = 6.89, 6.55, 5.63, 6.78$, respectively), offering strong support for Hypothesis 1 which says that there will be higher ratings on teaching quality in the experimental conditions than the control condition. In addition,

the Interpersonal Connections condition (*M* = 5.63) significantly differed from the other three experimental conditions, having the opposite effect than what was predicted in Hypothesis 4. At least across conditions, I found no support for Hypothesis 3 with the Spirituality quality because there were not higher ratings on teaching quality in the combined condition compared to the other experimental conditions.

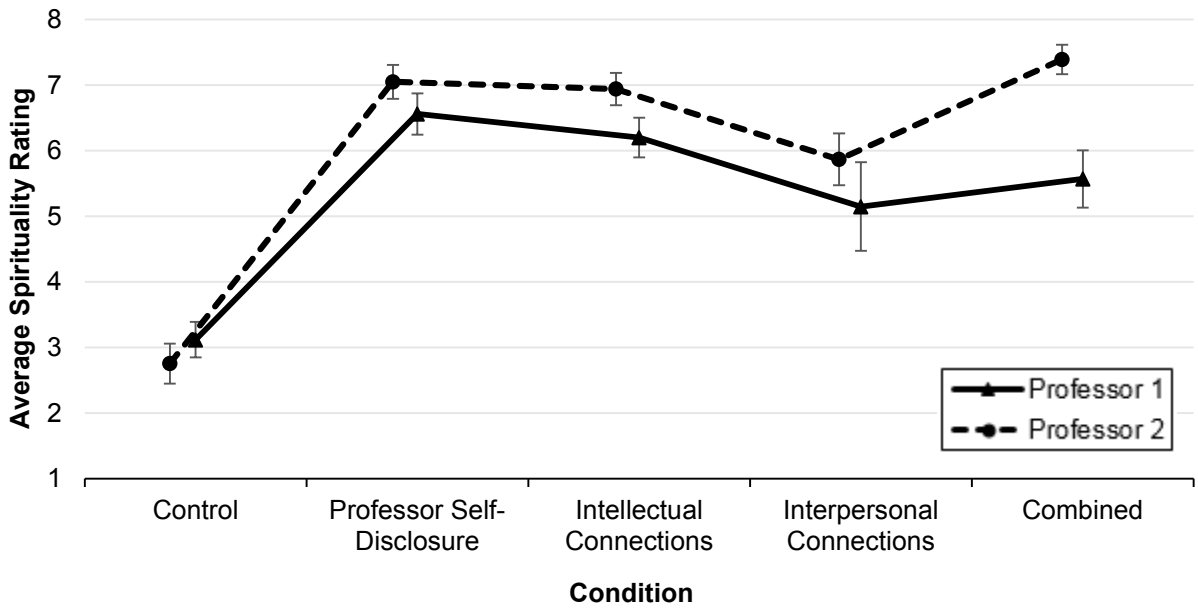

Figure 2. Average student evaluation ratings on the Spirituality variable across professors and conditions. Error bars represent standard errors. Points are offset horizontally so that error bars are visible. Recall that the order of the first three experimental conditions were randomized; therefore, this is does not reflect chronological order.

Post hoc tests were also performed to compare the means for professor in each condition on the Spirituality variable. Tukey's HSD tests indicated that the mean score for evaluations for Professor 2 in the combined condition (*M* = 7.38) was significantly higher than the mean score for Professor 1 (*M* = 5.57).

Post hoc analyses were also performed to compare the relative change in Spirituality ratings from the control to experimental conditions within each professor. The results showed

that Professor 2 had significantly more positive change in the Intellectual Connections and combined conditions compared to Professor 1.

Openness and Respect. Figure 3 graphically displays the evaluation ratings across professors and conditions on the Openness and Respect factor. As mentioned above, there were no main effects nor an interaction effect for this factor. There was no difference between conditions averaging across condition, $F(4, 193) = 1.28$, $p = .280$, partial $\eta^2 = .026$, and there was no difference between professor averaging across professor, $F(1, 193) = 0.99$, $p = .320$, partial $\eta^2 = .005$. There was also no significant interaction between the two independent variables, $F(4, 193) = 2.20$, $p = .071$, partial $\eta^2 = .044$. None of the hypotheses were supported for the Openness and Respect teaching quality.

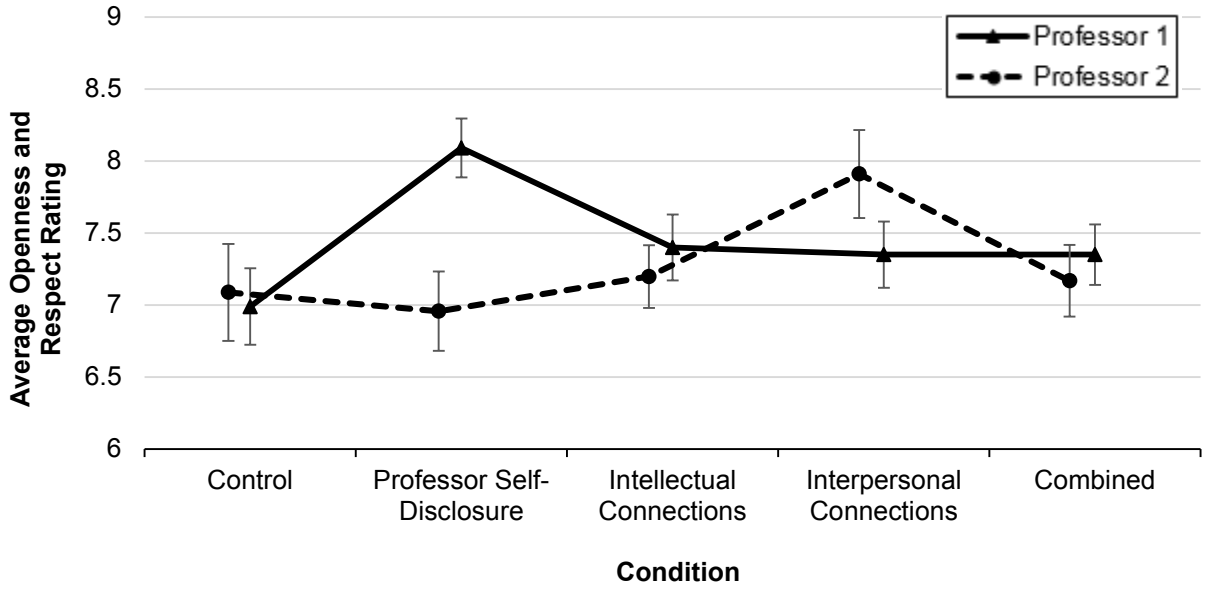

Figure 3. Average student evaluation ratings on the Openness and Respect variable across professors and conditions. Error bars represent standard errors. Points are offset horizontally so that error bars are visible. Recall that the order of the first three experimental conditions were randomized; therefore, this is does not reflect chronological order.

To determine whether professors differed in the relative change in Openness and Respect that was perceived by students from the control to experimental conditions, post hoc analyses were also performed. The results showed that Professor 1 had significantly more positive change in the Professor Self-Disclosure condition compared to Professor 2.

Retention. Basic descriptive data for the examination data for each group are listed in Table 5 and graphically displayed in Figure 4.

In order to compare the extent to which incorporating the areas of focus improved retention scores across conditions I conducted a factorial ANOVA that examined the main effects of condition and professor as well as the interaction effect between condition and professor on student retention scores. To control for demographic differences between professors, gender and college class were entered as fixed factors. The main effect for condition was not significant, $F(4, 156) = 1.05$, $p = .383$, partial $\eta^2 = .026$, nor was the main effect for professor, $F(1, 156) = 1.25$, $p = .265$, partial $\eta^2 = .008$. There was no significant interaction between the two independent variables, $F(4, 156) = 2.23$, $p = .068$, partial $\eta^2 = .054$. Gender and college class also failed to produce significant test statistics for the examination scores. These results indicate that students in the experimental conditions did not have higher nor lower retention scores than students in the control condition, offering no support for Hypotheses 2, and subsequently Hypotheses 3-4 which state that retention scores will be highest in the combined condition followed by the Interpersonal Connections condition.

Table 5

Basic Descriptive Statistics for Examination Data

	Professor 1		Professor 2		Total	
Condition	*n*	*M (SD)*	*n*	*M (SD)*	*n*	*M (SD)*
Control	17	7.41 (2.17)	18	8.33 (2.11)	35	7.89 (2.13)
Professor Self-Disclosure	12	8.42 (1.46)	27	7.39 (2.23)	39	7.71 (2.09)
Intellectual Connections	28	7.70 (2.25)	27	6.89 (2.35)	55	7.30 (2.31)
Interpersonal Connections	8	6.44 (2.37)	17	7.79 (1.84)	25	7.36 (2.07)
Combined	15	8.37 (2.31)	30	7.67 (1.99)	45	7.90 (2.10)
Average	80	7.74 (2.17)	119	7.55 (2.15)	199	7.63 (2.15)

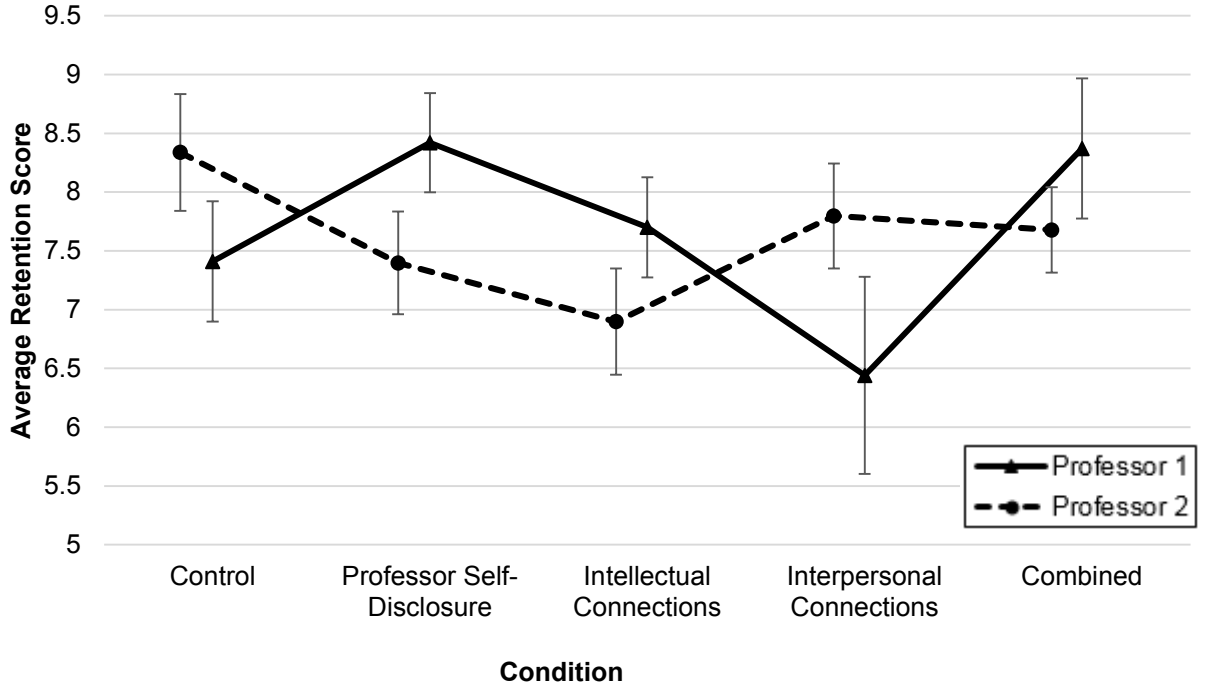

Figure 4. Average student retention scores across professors and conditions. Error bars represent standard errors. Points are offset horizontally so that error bars are visible. Recall that the order of the first three experimental conditions were randomized; therefore, this does not reflect chronological order.

Professor Self-Evaluations

Because there were only two faculty participants, broad conclusions were not made from the analysis of the professor self-evaluations. Instead, I reviewed each professor's response to each question and made simple observations from content that seemed salient and meaningful using data from Table 2 at times to expound on their comments.

There were a few responses on the Evaluation for Professors (see Appendix D) that seemed salient and meaningful in understanding the professors' experiences in participating in this study. In response to the first question, which asked, "To what extent did you feel the Spirit as you taught?," Professor 1 said that he felt the Spirit "extensively" as he taught. He expanded on this saying, "I felt that, at certain times, I was prompted to say certain things. Similar to my experience in most class settings." This indicates that his spiritual experience was not too different than when he teaches his normal classes. The percentage of time spent on the emphases bears this out; Professor 1 did not change his teaching very much. In fact, Professor 1 spent an average of 34.5% of teaching time on the focuses in total in the experimental conditions compared to 15.8% of the time in the control condition. That is a difference of only 18.7%, meaning independent raters observed only a very slight change in professor 1's teaching in the experimental conditions. Professor 2, on the other hand, spent an average of 77.9% of teaching time on the focuses in total in the experimental conditions compared to 3.1% of the time in the control group. This means raters observed him spending 74.8% more time dedicated to spirituality in the experimental conditions than he did in the control condition. This is a big difference from Professor 1. Therefore, it is expected that the professors would have differing experiences.

Professor 2's response to the first question was more powerful and specific. He indicated that he had stronger spiritual experiences when he focused on Professor Self-Disclosure: "Moderate but powerful at certain points of sharing personal beliefs and experiences." Professor 2, as a matter of fact, repeatedly mentioned the impact of the Professor Self-Disclosure focus. In response to question two about the extent to which he thought the students felt the Spirit, he said, "Probably moderate in most sessions but more strongly in…[the] self-disclosure [condition]." To question three about what he will remember most from the teaching experience he replied, "I learned to be more comfortable sharing personal and gospel insights and testimony—also that students applauded after…[the] self-disclosure [condition]." Again, it is clear here that Professor 2 reported making genuine changes to his teaching and noted the impact of the difference because there was one.

The greatest impact for Professor 1 seemed to be the Intellectual Connections focus. He responded to question three saying, "The value of connecting spiritual concepts to secular concepts" was the thing he will remember most from the teaching experience. In conjunction with this response he also mentioned in question eight that "The analogy between average speed and progression throughout life I tried to merit in the '[Intellectual] connection' sessions went over much better than I expected." Interestingly, both professors mentioned specific emphases that they intend to focus on in their actual teaching after this study. Professor 2 responded to question eight, "Yes, I will be more open to the Spirit and sharing relevant personal and gospel insights. Actually, I have already started doing it this semester and got positive student feedback."

Discussion

Several important findings emerged from the results. First, independent ratings and the professors' own reports indicate that the professors incorporated the three areas of focus into their teaching, with even minimal incorporation paying notable dividends. Second, the professors had differing perceptual experiences when implementing the focus areas. Additionally, incorporating the three areas of focus led to a positive increase in students' perception of the teaching quality. Finally, including the focuses did not result in an improvement in retention of the material; however, it may not have hurt retention either.

Effective Preparation

One of the most practical findings is that the professors' preparation was successfully incorporated into their teaching. The findings displayed in Table 2 are particularly relevant on this matter and warrant careful attention. For both professors in this study, there was a significant increase in time spent emphasizing the areas of focus from the control condition to the experimental conditions.

There was, however, a difference between the professors' emphasis on the focuses. In this study, Professor 2 fared much better than Professor 1 at spending time incorporating the proper focus in each condition. One reason for this could be that the professors may have had different comfort levels with the focus areas. Nonetheless, the findings show that even with only 19% more attention dedicated to the focuses (on average) in the experimental conditions, Professor 1's perception of being spiritual significantly increased. This may imply that professors who might be uncomfortable with knowing how to incorporate spirituality into their classes can make only minor tweaks in one or more of these focus areas and still reap great rewards (e.g., on teaching quality). As an example, in the Professor Self-Disclosure condition,

Professor 1 only increased his focus on Professor Self-Disclosure by about 11 percentage points from the control condition (and Interpersonal Connections by about 9.5 percentage points as well). His ratings on Spirituality then increased markedly, as well as his ratings on General Teaching Skills and Openness and Respect. Even the retention scores of his students increased in the Professor Self-Disclosure condition, though not to a significant degree.

This finding may help to bridge the gap between student and faculty expectations around faith and learning. It demonstrates that even a slight increase in time spent on the spiritual emphases—for example, a slight increase in disclosures about their personal relationship with God—may: 1) help ease the mind of professors who have expressed concerns and discomforts around integrating spirituality into their teaching, and 2) help students who would like to see more spirituality in their classes feel like their expectations are being met.

Difference in Professors' Perceived Experience

The findings from the Evaluation for Professors indicate that professors have differing perceptual experiences in implementing the three areas of focus. Professor 1 mentioned how his experience in this study (especially in relation to Professor Self-Disclosure) was at times akin to his experience in most regular class settings at BYU. This is not a surprise as professors, in general, at BYU are expected to work on being spiritually strengthening and intellectually enlarging to their students (see Aims of a BYU Education). Additionally, even though Professor 1 claimed to self-disclose most of the time while he is in front of a classroom, the results from Table 2 indicate that that may not be the case, at least for this study. This suggests that though professors may see themselves as already implementing these focuses into their normal teaching, they may not be doing so as much as they think they are.

Professors also are impacted by some factors more than others. In this study, Professor 1 seemed to be more impacted by the Intellectual Connections focus whereas Professor 2 was more impacted by the Professor Self-Disclosure focus. In fact, the implementation of Professor Self-Disclosure was so successful for Professor 2 that his students applauded at the end of the teaching session. This factor had such an influence on him that he started incorporating it in his regular classes and had already received positive feedback. Professor 2 seems to have made genuine changes to his teaching (he implemented Professor Self-Disclosure into the corresponding experimental condition 66.58% of the time) and noticed the impact of the difference—compared to Professor 1 who raters observed as making only small changes. This finding suggests that after effective preparation some professors may be able to not only improve on, but excel in, at least one area of focus. That aspect can positively change the way they teach.

Teaching Quality Improvement

In the present study, the professors' implementation of the preparation improved student participants' perceptions of their teaching quality and their spirituality, when compared to the control condition. In particular, for General Teaching Skills, professors, on average, received significantly higher ratings in the Interpersonal Connections condition and the combined condition, with Interpersonal Connections having the largest effect, particularly for Professor 2. For Spirituality, professors received significantly higher ratings in all the experimental conditions, with Interpersonal Connections having the smallest effect. This is interesting because I predicted that Interpersonal Connections would have the greatest individual influence on teaching quality, but it turns out that even though it had the greatest effect on General Teaching Skills, it had the least effect on Spirituality. One possible explanation for this finding is that students are less likely to see the teachers' instruction as spiritual when it focuses on the

relationship because spirituality is not explicitly discussed as much within this focus area. This is especially likely at a religious institution like BYU where there may be a tendency to think of spirituality in religious terms and content of a class session.

Despite improvements to the scales on the Teacher Rating Questionnaire to help reduce ceiling effects found on the pilot studies, there was still no significant effect on the Openness and Respect variable. The reason why the preparation had no effect was probably because the professors were naturally adept at demonstrating those qualities that made up that factor. The baseline ratings for Openness and Respect were still very high (indicating a potential ceiling effect), so there was little room for significant improvement. Of note, however, is that the means (displayed in Table 4) show that both professors improved from the control condition to the experimental conditions on this factor (with the exception of Professor 2 in the Professor Self-Disclosure condition), even though the differences are not statistically significant. These findings offer a complement to the instructors involved in the study. Future studies may find that professors who have low ratings on the Openness and Respect factor could benefit from preparation on the three areas of focus.

The results also indicate that professors differed significantly from one another on the General Teaching Skills and Spirituality qualities. Professor 1 had higher absolute ratings on General Teaching Skills than Professor 2 in the Professor Self-Disclosure and Intellectual Connections conditions. This is not all that surprising given that Professor 1 already had higher ratings on this variable in the control condition—an indication that he may be more naturally adept at those skills than Professor 2. One reason why Professor 1's ratings on General Teaching Skills in the Interpersonal and combined conditions did not significantly differ from the control condition could be that according to the independent raters, Professor 1 made only minor

changes in implementing spirituality, unlike Professor 2 who was observed to make major changes (see Table 4). In other words, relative to their respective control conditions, Professor 2's ratings on General Teaching Skills in the Interpersonal Connections and combined conditions increase more significantly than Professor 1's. Professor 2 also had higher ratings on Spirituality than Professor 1. This, too, is not surprising given how much Professor 2 integrated the focuses into the experimental conditions and how slight Professor 1's changes were. From these findings it is important to note that even when Professor 1 implemented IFL minimally, there still is an effect on teaching quality.

While there were significant differences in these teaching qualities across both professors and conditions, significant interactions indicate that student ratings of their professors depend on the condition in which they are teaching. This again highlights how professors may excel in one focus area and struggle in others. One specific example of where this played out in this study can be found when we examine support for Hypothesis 3 (there will be higher ratings on teaching quality in the combined condition compared to the other experimental conditions) more closely. Initially, it seems that Hypothesis 3 was not supported for the General Teaching Skills nor the Spirituality variables because there were no significant differences between the combined condition and the other experimental conditions. However, in light of five facts, there may be partial support for this hypothesis when looking at these variables:

1) Professor 2 had significantly higher student ratings than Professor 1 on Spirituality in the combined experimental condition.

2) The combined condition is where the professors differed the most on implementation of the focus areas (i.e., 27% for Professor 1 and 87% for Professor 2; a difference of 60%).

3) If Professor 2 spent 87% of the time on the focus areas, he was likely teaching very differently than normal.

4) Relative to his own control condition, Professor 2's ratings on both variables are significantly higher.

5) Compared to Professor 2, Professor 1 spent less additional time on the focuses in the experimental conditions relative to the control condition.

These facts highlight a possible strength for Professor 2 and a weakness for Professor 1. The facts could also be interpreted as indicating the challenges of incorporating spirituality into classes with different disciplinary material. But, despite very different disciplines, both professors were able to increase the amount of class time spent on the focuses in such a way that increased student perceptions of general teaching skills and spirituality. These findings speak to the concern faculty have about spirituality and the fit with their discipline. Future research could look into which other disciplines produce similar results when implementing spirituality.

Thus far, the results have shown that in regards to time spent on the focuses Professor 2 made a pretty radical departure from his regular way of teaching in the experimental conditions. The students saw his general teaching skills as being of lower quality than Professor 1 at times, but they saw him as being more spiritual and equally open and respectful in the experimental conditions. The following section highlights another important finding of this study.

No Apparent Gains or Losses in Student Retention

Analyses showed that although implementing the areas of focus did not improve student retention after one week, it may not have hurt retention either. This finding does not support the survey research by Astin et al. (2011) which found that spiritual emphasis enhances academic performance. But other research has observed that "There is little evidence to suggest that this

interest in…faith in any way undermines the educational purposes of higher education; faith and learning can go hand in hand" (Braskamp, 2008, p. 117).

Notably, there are limitations in drawing conclusions from results that do not allow for the rejection of the null hypothesis. However, the descriptive statistics are supportive in their own right as a preliminary indication (needing further testing, of course) that rigor and performance may not suffer from the inclusion of spirituality, while perceptions of teaching are strengthened. These findings may be open to this interpretation, but given the effect sizes found, the retention test may have lacked sufficient power to even detect a difference (gain or loss) in retention. Post hoc power analyses indicated that for the sample sizes, the power for detecting the observed effect sizes of .026 (condition effect), .008 (professor effect), and .054 (interaction effect) is only 5%, 43%, and 67%, respectively. In order for a condition effect of the observed size to be detected with 80% power at the .05 level, a sample of approximately 10,770 participants would be required. Thus, the results of the retention test are inconclusive and future research will need to develop a test of retention with adequate power to truly measure whether Braskamp's observation actually holds: whether retention is aided or hurt by the implementation of the focuses.

If there was sufficient power, the research on teacher immediacy suggests that relational teaching and self-disclosure would lead to improved cognitive learning (e.g., Mullin, 2014; Witt & Wheeless, 2001; Witt et al., 2004; Witt et al., 2006). Even though the findings of this study do not necessarily undermine the foundational nature of relationality in spirituality, the findings cannot support the notion that the relational aspect of spirituality is the most impactful factor on student retention scores. Power may have been weak due to the test used, but also it may have been weak because the impact of the relational focus may require time. Since this study was

only a one-time, 30 minute teaching scenario, conclusions about the effect of relational teaching are also initial and inconclusive.

Limitations and Suggestions for Future Research

The present study was the first quasi-experiment to empirically compare different ways of spiritual teaching and as such, it has several limitations.

Unequal groups. First and foremost, the method of students self-selecting study times meant that the study conditions were not truly randomly assigned and also resulted in unequal group sizes. As a result, it is possible that one group of students in a particular condition may have been systematically different than another group of students and thus responded differently to the Teacher Rating Questionnaire and/or to the retention examination. The results bear this out. As previously mentioned, Professor 1 had more males and more juniors than Professor 2 while Professor 2 had more freshmen. Having significantly more students representing one group (e.g., freshmen) than students representing another (e.g., juniors) may skew the average examination and/or evaluation scores for professors toward the prevailing abilities of students in the group that is more represented (e.g., freshmen). While I was unable to test for this possibility using the perception data, I found no significant demographics differences on examination scores between conditions nor between professors. However, group differences on other factors such as motivation, attention, and intelligence could have impacted the results of this study. Future research should randomly assign participants to conditions to avoid this issue.

Lack of statistical power. Secondly, this study had limited statistical power to detect differences between conditions in retention scores. Although larger sample sizes would have helped, it is unlikely that there would have been significant differences. It could be the case that spirituality and relationality have no real effect on student retention. It could also be that I did

not adequately test retention. This point might be explained by two subsequent limitations: lack of student motivation and potential issues with the exams.

Lack of student motivation. Most students recruited for this study participated to get research credit as part of the requirements or incentives for one or more psychology courses. Even though students read the material beforehand and were taught for 30 minutes on that material, retention scores were very low, averaging 7.63 out of 13 (58.7%). This brings into question how motivated the students were to pay attention to the content of the reading and lecture material. Had they expected to be tested on the material, their forgetting may have been reduced (Oeberst & Lindner, 2015). Yet, even then, to receive credit for participation students were not required to score well on the examination—they just had to take it. Although this issue may not be completely avoidable, future research should take student motivation into consideration as a potential way to increase power.

Potential exam issues. Another possibility for the lack of statistical power could be due to issues directly related to the examinations themselves. Maybe they had too few questions. Could a test with only 10 questions provide enough statistical power to find an effect? Additionally, the exam questions were taken from the reading and it could be that the lecture did not add anything new. If at least some of the exam questions were based on what was said in the lecture that was not included in the reading, the power of the test might have been increased.

Relationship between topics and focuses. Another limitation of this study concerns the teaching content. Comparable to how Reber et al. (in press) have discussed in a study with similar methods, the professors in this study chose the reading and topics for their lectures based on their expertise so that they could focus on implementing the focus areas without being distracted by having to learn new material. Even though the pilot data suggest using new

material can still result in the same kinds of outcomes, the topics of Linear Kinematics and the First Amendment may have favored some focus areas over others. Had the topics been different—say, biology or engineering—Intellectual Connections, for example, may have been more difficult to apply. One of the goals of this study was to use faculty from different disciplines in order to counter the argument that spiritual teaching does not fit with some disciplines. Although still quite deficient, this study is one of the first to compare spiritual and relational teaching using content and topics from more than one discipline. Future research should continue this process by investigating how spiritual and relational teaching affect students in a variety of disciplines and topics.

Preparation and treatment fidelity. There were challenges to the preparation of teachers and the treatment fidelity. Although the professors made significant changes from the control condition, the professors in this study are not completely comparable because one (Professor 2) implemented the preparation a lot more effectively than the other (Professor 1). There could be several internal or external reasons for this. The difference could be due to professor motivation, ability, comfort level, or personality traits. It could also be due to external factors such as lack of time to sufficiently prepare for the conditions. Another possibility is that the nature of the discipline or topic made it more or less difficult to implement the preparation.

This difficulty in taking up the preparation also showed through in regards to treatment fidelity. Primarily, the results displayed in Table 2 suggest that professors differ on the natural emphasis they place on each of the focuses. In the control condition of this study, where professors have not received any explicit preparation on these focuses but are just asked to teach as if they were at a state university, they implemented both Professor Self-Disclosure and Interpersonal Connections to varying degrees. Professor 1 has small, yet larger, percentages on

these focuses (7% and 9%, respectively), than Professor 2, who had very small, almost non-existent, percentages on these focuses (1% and 2%, respectively). This difference between their control conditions is interesting, particularly in relation to their different disciplines. One explanation could be that Professor 2 naturally teaches with no emphasis on spirituality, and that it is uncommon for his discipline to encourage or emphasize spirituality in teaching. The reverse may also be the case for Professor 1: he naturally teaches with some emphasis on spirituality, and it is more common for his discipline (compared to Professor 2's discipline) to encourage or emphasize spirituality in teaching.

As expected, the professors' natural, baseline of time spent on these focuses (without preparation) seemed to have carried through each condition to some extent. Since a baseline control condition reveals natural teaching tendencies for these focuses, it helps explain how the focuses may have "leaked" into the experimental conditions. However, it does make treatment fidelity a concern as it makes it difficult to know how much of the change from the control to the experimental conditions is due to natural tendencies. Notably, though, this specific finding is mainly applicable to Professor 1 because Professor 2's control condition percentages are almost non-existent.

The results indicated that it was difficult for the professors to stay true to the focus in the intended condition. This finding lends support for the interconnectedness of the areas of focus mentioned previously in the literature review. In particular, this may highlight the similarity between the Professor Self-Disclosure and Interpersonal Connections focuses. Even though Interpersonal Connections emphasizes the student-professor relationship directly, Professor Self-Disclosure may also have a relational component. It is possible that in order to build a relationship with the students, the professors had to share some of themselves. This might help

explain why professors implemented some of both focuses into their teaching in each of these two experimental conditions (see Table 2). Particularly, this could explain why Professor 2 emphasized Professor Self-Disclosure 33% of the time when he was supposed to emphasize Interpersonal Connections (which he did 39% of the time) and why Professor 1 emphasized Interpersonal Connections (30% of the time) more than Professor Self-Disclosure (16% of the time) in the same condition.

This challenge of fidelity clouds the ability to be fully confident that a particular area of focus had a specific effect, for example, on teaching quality or retention scores. By improving the treatment fidelity there may be increased power to detect effects. Understandably, treatment fidelity is difficult when dealing with a live teaching session that involves participation from the students (Reber et al., in press). It is also difficult to do when one is dealing with aspects of spirituality that are interrelated, as evidenced by the literature. Although the professors were prepared to teach in a manner that fit the focus of each condition, some variation across conditions was unavoidable. Thus, like Reber et al. (in press) note in a similar study, this was a study of teaching emphasis, not teaching purity. To address some of these challenges with the preparation and treatment fidelity, future research could consider having a preparation that does not just emphasize how to implement the focuses but how to exclude other focuses as well.

Other limitations. There were a few other methodological issues that future research could look into improving. Firstly, the data from the student evaluations and the data from the student examinations and demographics could not be matched. A better system of data collection would have made it possible to pair these measures together, allowing an omnibus test to be conducted on all dependent variables at the same time. Moreover, this study lacked a proper parallel forms reliability test of the two exams. This would have allowed for further

confirmation of exam comparability. Additionally, even though the order of the single-focus teaching conditions were counterbalanced, the possibility of practice effects are not overcome with only two professors. For example, it is possible that practice effects led to an improvement in the perception of professors' General Teaching Skills from one experimental condition to another. Where feasible, future research should involve enough faculty participants so that the threat of practice effects could be sufficiently addressed by counterbalancing.

Additional directions for future research. This study is similar to many of the cognitive learning studies that Witt et al. (2004; 2006) analyzed: experimental designs and "time-bound snapshots" (p. 200; p. 160) that limited the ability to look at cognitive learning over time. Allen et al. (2006) state that the "tendency of educational research…is to focus on a point in time for the instructor and the course" (p. 29) and then advocate that this focus should be broadened. Seeing that in the current study an hour long preparation elicited improvement on the General Teaching Skill and Spirituality variables in just a 30 minute lecture to undergraduates who had no previous knowledge of the professors, it would be interesting to test what effect a similar preparation has in an actual classroom setting over the span of a whole semester where students would be more motivated to perform well on the examinations. Considering that the goal is not simply academic improvement in a single course nor is it successful completion of a program of study, but may be about becoming more spiritually strengthened and intellectually enlarged, instructor implementation of spirituality and relationality in a course and across an entire curriculum may have positive and meaningful long-term consequences. In a longitudinal study, what kind of outcomes would we see? Would the preparation still be effective? Would professors be able to effectively incorporate the areas of focus throughout the semester?

If such a classroom intervention yields fruitful outcomes, it would lead me to wonder if a large-scale preparation would be worth pursuing. A department-wide pedagogical preparation, for example, would be a good start to assess whether spirituality and relationality improve outcomes across faculty and within a discipline. This idea is supported by other researchers who have analyzed several similar studies: "The call should be to expand the vision beyond the single student and teacher in one course to a more programmatic view of a degree program across 4 years of education in college" (Witt et al., 2006, p. 163). These same researchers again mention that "such an assessment requires an entire body of longitudinal research focusing more on systemic interests about educational outcomes" (Allen et al., 2006, p. 29).

Keeping in mind that BYU is a religious (LDS) institution following the paradigm of theistic spirituality, the generalizability of this study must be carefully considered. While the preparation was well understood, supported, and implemented by the faculty from BYU recruited for this study, I wonder how well other faculty at BYU and at other universities will take the preparation and how well their students will receive the teaching. Certainly, spirituality can be implemented at any university, but the type of spirituality may be different. I wonder, how can a preparation that emphasizes spirituality be adapted in other universities? Spirituality at BYU likely will not—and maybe even cannot—look the same as spirituality at a state university, but it may be quite comparable at another religiously affiliated university. In planning a preparation that involves incorporating spirituality, researchers must first consider the mission, aims, and values of the institution. They must also consider the personal values of the individual faculty and students. But more importantly, as Astin et al. (2011) advocate,

> *how* students define their spirituality or *what* particular meaning they make of their lives is not at issue. Rather, [the] concern is that the relative amount of attention that colleges

> and universities devote to the "inner" and "outer" aspects of our students' lives has gotten way out of balance. (p. 2)

A major part of the "inner" aspect of students' lives is their spirituality.

Conclusion

Research done by HERI and by BYU demonstrates what Rice (2008) calls a "disparity between what students want and expect and what the American professoriate provides" (p. 101) in regards to spirituality in the classroom. Professors have major concerns about implementing spirituality in the classroom among which are beliefs that it might not fit with their discipline and fear that it might hurt student learning. This study showed that the faculty concerns about implementing spirituality into their teaching are not always warranted. As a quasi-experimental design never before done on such a topic, this study demonstrated that the divide between what students want and what professors provide in regards to spirituality in higher education can be bridged through a well-executed pedagogical preparation that educates willing professors on how to 1) focus on their journey, personal experiences, and the ways in which spirituality is working in his or her personal and professional life; 2) focus on intellectual connections between spirituality and the discipline, and illuminating the meaning that those connections have in students' lives; and 3) focus on their relationship with their students. In fact, by incorporating these focuses, even to a small degree, professors and students benefit because there is an increase in the perception of teaching quality and of the spirituality so many students desire.

References

Aims of a BYU Education. Retrieved from http://aims.byu.edu/aims

Allen, M., Witt, P. L., & Wheeless, L. R. (2006). The role of teacher immediacy as a motivational factor in student learning: Using meta-analysis to test a causal model. *Communication Education, 55*(1), 21-31. doi:10.1080/03634520500343368

Astin, A. W., Astin, H. S., & Lindholm, J. A. (2011). *Cultivating the spirit: How college can enhance students' inner lives.* San Francisco, CA: Jossey-Bass

Braskamp, L. A. (2008). The religious and spiritual journeys of college students. In D. Jacobsen, and R. H. Jacobsen (Eds.), *The American university in a postsecular age* (pp. 117-134). New York, NY: Oxford University Press.

Buber, M. (1958). *I and thou* (2nd ed.). New York, NY: Charles Scribner's Sons.

BYU Mission Statement. Retrieved from http://aims.byu.edu/mission_statement

Cooper, M. V (1999). Faculty perspectives on the integration of faith and academic discipline in Southern Baptist higher education. *Religious Education, 94*, 380-395. doi:10.1080/ 0034408990940402

Dawes, J. (2008). Do data characteristics change according to the number of points used? An experiment using 5-point, 7-point and 10-point scales. *International Journal of Market Research, 50,* 61–77.

Dunham, T. C., & Davison, M. L. (1991). Effects of scale anchors on student ratings of instructors. *Applied Measurement in Education, 4*, 23-35.

Edwards, M. U., Jr., (2008). Why faculty find it difficult to talk about religion. In D. Jacobsen, and R. H. Jacobsen (Eds.), *The American university in a postsecular age* (pp. 81-97). New York, NY: Oxford University Press.

Feldman, K. A., & Newcomb, T. M. (1969). *The impact of college on students.* San Francisco, CA: Jossey-Bass.

Feldman R., & Prohaska, T. (1979). The student as Pygmalion: Effect of student expectation on the teacher. *Journal of Educational Psychology, 71*(4), 485-493.

Fowler, J. W. (1981). *Stages of faith: The psychology of human development and the quest for meaning*. San Francisco, CA: Harper & Row.

Fried, J. (2001). Civility and spirituality. In V. M. Miller, and M. M. Ryan (Eds.), *Transforming campus life: Reflections on spirituality and religious pluralism* (pp. 261-280). New York, NY: Lang.

Frymier, A. B. (2013). Teacher immediacy. In J. Hattie, E. M. Anderman, J. Hattie, E. M. Anderman (Eds.), *International guide to student achievement* (pp. 425-427). New York, NY, US: Routledge/Taylor & Francis Group.

Gorham, J. (1988). The relationship between verbal teacher immediacy and student learning. *Communication Education, 37,* 40-53.

Gross, N., & Simmons, S. (2008). The religious convictions of college and university professors. In D. Jacobsen, and R. H. Jacobsen (Eds.), *The American university in a postsecular age* (pp. 19-29). New York, NY: Oxford University Press.

Hasker, W. (1992). Faith-learning integration: An overview. *Christian Scholar's Review, 21*, 234-248. Retrieved from https://www.apu.edu/live_data/files/219/faith_learning_ integration_an_overview.pdf.

Kelley, D. H., & Gorham, J. (1988). Effects of immediacy on recall of information. *Communication Education, 37*, 198-207. doi:10.1080/03634528809378719

Kimball, S. W. (1967, September). *Education for eternity*. Preschool Address to BYU Faculty and Staff. Provo, UT.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213-236.

Leahey, T.H. (1991). *A history of modern psychology*. Englewood Cliffs, NJ: Prentice Hall.

Lindholm, J. A., Astin, H. S., & Astin, A. W. (2005). *Spirituality and the professoriate: A national study of faculty beliefs, attitudes, and behaviors.* Los Angeles, CA: Higher Education Research Institute, UCLA.

Lindholm, J. A., Szelényi, K., Hurtado, S., & Korn, W. S. (2005). *The American college teacher: National norms for the 2004–2005 HERI faculty survey*. Los Angeles, CA: Higher Education Research Institute, UCLA.

McGinnis, P. M. (2004). *Biomechanics of sport and exercise* (2nd ed.). Champaign, IL: Human Kinetics.

Mehrabian, A. (1969). Some referents and measures of nonverbal behavior. *Behavioral Research Methods and Instrumentation, 1,* 213-217.

Mehrabian, A. (1971). *Silent messages.* Oxford, England: Wadsworth.

Moore, A., Masterson, J. T., Christophel, D. M., & Shea, K. A. (1996). College teacher immediacy and student ratings of instruction. *Communication Education, 45,* 29-39. doi: 10.1080/03634529609379030

Morling, B. (2014). *Research methods in psychology* (2nd ed.). New York, NY: WW Norton & Company, Inc.

Mullin, G. P. (2014). Why can't we be friends? Using relatedness to complete the link between teacher immediacy and student learning. *Dissertation Abstracts International,* 74(11-A)(E).

NSSE (2006). *Engaged learning: Fostering success for all students.* Annual Report of the National Survey of Student Engagement. Bloomington, IN: University of Indiana. Retrieved October 29, 2014, from http://nsse.iub.edu/NSSE_2006_Annual_Report/docs/NSSE_2006_Annual_Report.pdf.

Oeberst, A., & Lindner, I. (2015). Unannounced memory tests are not necessarily unexpected by participants: Test expectation and its consequences in the repeated test paradigm. *Cognitive Processing, 16*(3)*,* 269-278. doi:10.1007/s10339-015-0663-3

Reber, J. S., Downs, S. D., & Peterson-Nelson, J. A. (in press). Effects of three pedagogies on learning outcomes in a Psychology of Gender lecture: A quasi-experimental study. *Teaching of Psychology*.

Reber, J. S., Slife, B. D. (2013). Theistic psychology and the relation of worldviews: A reply to the critics. *Christian Psychology: A Transdisciplinary Journal. 7*(1)*,* 5-20.

Rice. R. E. (2008). Faculty priorities: Where does faith fit? In D. Jacobsen, and R. H. Jacobsen (Eds.), *The American university in a postsecular age* (pp. 99-114)*.* New York, NY: Oxford University Press.

Ridge, R. D., Reber, J. S., (2011, June). *Expectations of teachers resulting from public online evaluations: Expectancy confirmation.* Presented at the meeting of the Hawaii International Conference of Social Sciences, Honolulu, HI.

Ripley, J. S., Garzon F. L., Hall, M. E. L., Mangis, M. W., & Murphy, C. J. (2009). Pilgrims' progress: Faculty and university factors in graduate student integration of faith and profession. *Journal of Psychology & Theology, 37*, 5-14. Retrieved from http://digitalcommons.liberty.edu/ccfs_fac_pubs/43/

Sherr, M., Huff, G., & Curran, M. (2007). Student perceptions of salient indicators of integration of faith and learning (IFL): The Christian vocation model. *Journal of Research on Christian Education, 16*, 15-33. doi:10.1080/10656210701381080

Sibii, R. (2010). Conceptualizing teacher immediacy through the 'companion' metaphor. *Teaching In Higher Education, 15*(5), 531-542. doi:10.1080/13562517.2010.491908

Silva, K. M., Silva, F. J., Quinn, M. A., Draper, J. N., Cover, K. R., & Munoff, A. A. (2008). Rate my professor: Online evaluations of psychology instructors. *Teaching of Psychology, 35*(2), 71-80. doi:10.1080/00986280801978434

Slife, B. D., & Gantt, E. (1999). Methodological pluralism: A framework for psychotherapy research. *Journal of Clinical Psychology, 55*(12), 1453-1465.

Slife, B. D., & Reber, J. S. (2009). Is there a pervasive bias against theism in psychology? *Journal of Theoretical and Philosophical Psychology, 29*(2), 63-79. doi:10.1037/a0016985

Slife, B. D., Reber, J. S., & Lefevor, G. (2012). When God truly matters: A theistic approach to psychology. *Research in the Social Scientific Study of Religion, 23*, 213-237.

Slife, B. D., & Whoolery, M. (2006). Are psychology's main methods biased against the worldview of many religious people? *Journal of Psychology and Theology, 34*(3), 217-231.

Speck, B. W. (2005). What is spirituality? *New Directions for Teaching and Learning, 2005*(104), 3-13. doi:10.1002/tl.207

Sorenson, R. (1997). Doctoral student's integration of psychology and Christianity: Perspectives via attachment theory and multidimensional scaling. *Journal for the Scientific Study of Religion, 36*, 530-548. doi:10.2307/1387688

Wallace, M., Campbell, S., Grossman, S. C., Shea, J. M., Lange, J. W., & Quell, T. T. (2008). Integrating spirituality into undergraduate nursing curricula. *International Journal of Nursing Education Scholarship, 5*(1). Retrieved from digitalcommons.fairfield.edu/ nursing-facultypubs/8

Watson, N. J. (2006). Martin Buber's I and thou: Implications for Christian psychotherapy. *Journal of Psychology and Christianity, 25*(1), 34-43.

Wilkins, A. L., & Birch, A. J. (2011). *Spiritually strengthening and intellectually enlarging faculty: What students want.* Retrieved from http://faithandlearning.weebly.com

Wilson, R. C, Bavry, J. L., Gaff, J. G., Wood, L., & Dienst, E. R. (1975). *College professors and their impact on students*. New York, NY: Wiley.

Witt, P. L., & Wheeless, L. R. (2001). An experimental study of teachers' verbal and nonverbal immediacy and students' affective and cognitive learning. *Communication Education, 50*(4), 327-342. doi:10.1080/03634520109379259

Witt, P. L., Wheeless, L. R., & Allen, M. (2004). A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs, 71*(2), 184-207. doi:10.1080/036452042000228054

Witt, P. L., Wheeless, L. R., & Allen, M. (2006). The relationship between teacher immediacy and student learning: A meta-analysis. In B. M. Gayle, R. W. Preiss, N. Burrell, & M.

Allen (Eds.), *Classroom communication and instructional processes: Advances through meta-analysis* (pp. 149-168). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Appendix A

Teacher Rating Questionnaire

Instructions:

- **Mark all of your answers on the scantron <u>in order</u>**
- **Complete each section of the questionnaire <u>in order</u>**
- **Pay attention to the scales on which you will be rating—they change for each section**

Part A:

For the following statements about the teacher and the lesson, please indicate the degree to which you agree or disagree. We expect that there may be areas where this teacher was particularly strong as well as areas where this teacher could improve. We encourage you to use all the points on the scale as needed in order to accurately indicate your level or agreement of disagreement.

Disagree (1 to 3)			**Agree (4 to 6)**			**Strongly Agree (7 to 9)**		
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

1. This teacher seemed to be helpful.
2. This teacher presented the material in an unclear way.
3. It would be difficult to succeed in a class if it were taught by this teacher.
4. Overall, this is a quality teacher.
5. This lesson was difficult to understand.
6. This lesson was interesting.
7. The lesson materials and activities were effective.
8. This teacher was competent.
9. This teacher was intelligent.

10. I did not like the way this teacher taught.

11. I would recommend this teacher to a friend.

12. This teacher was well organized.

13. This teacher was enthusiastic.

14. This teacher had a poor knowledge of the topic.

15. This teacher did a good job stimulating students' interest.

16. If this teacher were to teach a class, I would sign up to take a class from them next semester.

17. This teacher was spiritually inspiring.

18. This teacher has an ongoing personal relationship with God.

19. This teacher integrated the gospel into the subject.

20. This teacher developed a good working relationship with the students.

21. If I had this teacher for a class, I could develop good rapport with him/her.

22. I would feel uncomfortable approaching this teacher for help.

23. This teacher was open to questions.

24. This teacher allowed us to connect with each other.

25. If I had this teacher for a class, I could get closer to the other students.

26. This teacher has interest in student learning.

27. This teacher gave opportunities for us to get help.

28. This teacher responded to students respectfully.

29. This teacher explained concepts effectively.

30. This teacher promoted active student involvement.

31. My testimony was strengthened because of this teacher.

32. My intellectual skills were developed because of this teacher.

33. This teacher contributed to the BYU aims (Spiritually Strengthening, Intellectually Enlarging, Character Building, Lifelong Learning and Service).

Part B:

Below are a number of teacher qualities and the behaviors that define them. For each statement please indicate the degree to which you agree or disagree about whether the teacher possesses these qualities and exhibits the corresponding behaviors. We expect that there may be areas where this teacher was particularly strong as well as areas where this teacher could improve. We encourage you to use all the points on the scale as needed in order to accurately indicate your level of agreement or disagreement.

Disagree (1 to 3)			Agree (4 to 6)			Strongly Agree (7 to 9)		
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

34. Approachable/Personable (Smiles, greets students, invites questions, responds respectfully to student comments)

35. Authoritative (Maintains classroom order)

36. Confident (Speaks clearly, makes eye contact, and answers questions correctly)

37. Creative and Interesting (Uses interesting, relevant, and personal examples; not monotone)

38. Effective Communicator (Speaks clearly/loudly; uses precise English; gives clear, compelling examples)

39. Enthusiastic About Teaching and About Topic (Smiles during class, prepares interesting class activities, uses gestures and expressions of emotion to emphasize important points)

40. Flexible/Open-Minded (Pays attention to students when they state their opinions, accepts criticism from others)

41. Good Listener (Doesn’t interrupt students while they are talking, maintains eye contact, and

asks questions about points that students are making)

42. Happy/Positive Attitude/Humorous (Tells jokes and funny stories, laughs with students)

43. Humble (Admits mistakes, never brags, and doesn't take credit for others' successes)

44. Knowledgeable About Subject Matter (Easily answers students' questions, does not read straight from the book or notes, and uses clear and understandable examples)

45. Prepared

46. Presents Current Information (Relates topic to current, real-life situations; talks about current topics)

47. Promotes Class Discussion (Asks controversial or challenging questions during class, and involves students in group activities during class)

48. Promotes Critical Thinking/Intellectually Stimulating (Asks thoughtful questions during class)

49. Punctuality/Manages Class Time (Presents relevant materials in class, leaves time for questions)

50. Rapport (Makes class laugh through jokes and funny stories, initiates and maintains class discussions)

51. Respectful (Does not humiliate or embarrass students in class, is polite to students [says thank you and please, etc.], does not interrupt students while they are talking, and does not talk down to students)

52. Sensitive and Persistent (Makes sure students understand material before moving to new material, repeats information when necessary, and asks questions to check student understanding)

53. Understanding (Doesn't lose temper at students, takes extra time to discuss difficult

concepts)

54. Spiritual (It is clear he/she has a testimony, demonstrates an ongoing relationship with God)

55. Spiritually intellectual (Cares about connections between the gospel and the material, shows the relevance of such connections)

56. Interpersonal (Demonstrates concern for students, develops a working relationship with students, helps students relate to each other)

57. Genuinely Cares for Learning (Interested in students' learning, willing to help students work though difficult problems related to the subject matter even outside of class)

Part C:

For the following questions, please use the following scale to rate the quality of the lesson and the teacher. We encourage you to use all the points on the scale as needed in order to accurately indicate your perception of the quality of the lesson and teacher.

Poor (1 to 3)			**Good (4 to 6)**			**Excellent (7 to 9)**		
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

58. How would you rate the overall lesson?

59. How would you rate the teacher overall?

Appendix B

Linear Kinematics Retention Examination

1. If you run 10 km in 26 min 22 s, what was your average speed?
    a. 0.63 m/s
    b. 6.3 m/s
    c. 381 m/s
    d. none of the above
2. What term refers to the description of motion, without regard to the forces causing the motion?
    a. kinetics
    b. biomechanics
    c. kinematics
    d. inverse dynamics
3. A runner that is initially moving at a velocity of -6 m/s changes her velocity in a 1-s time interval to a velocity of -10 m/s. What was her average acceleration?
    a. -16 m/s^2
    b. -4 m/s^2
    c. -2 m/s^2
    d. none of the above
4. If a ball's acceleration is -0.5 m/s^2, and it takes 7 seconds to stop, what was its initial velocity?
    a. 3.5 m/s
    b. 7 m/s

c. 14 m/s

d. none of the above

5. What is your average speed if you run a kilometer at 3 m/s and then walk a kilometer at 1 m/s?

   a. 1.50 m/s

   b. 1.67 m/s

   c. 2.00 m/s

   d. none of the above

6. What does acceleration describe?

   a. the rate of change for position

   b. the rate of change for displacement

   c. the rate of change for distance

   d. none of the above

7. If a 400-meter sprinter runs one complete lap around a 400-meter track and ends in the same place he began, what will his final distance be?

   a. 400 meters

   b. $2 \times \Pi \times$ meters

   c. 0 meters

   d. there is not enough information to answer this question accurately

8. What does velocity describe?

   a. the rate of change for position

   b. the rate of change for displacement

   c. the rate of change for distance

d. none of the above

9. Which of the following is the best way to describe position?

   a. the magnitude and sign (direction), relative to a fixed position

   b. only the magnitude, relative to a fixed position

   c. only the sign (direction), relative to a fixed position

   d. none of the above

10. Relating to improving his/her training, explain why a world class sprinter might want to consider instantaneous speeds, every 10 meters throughout a 100-meter sprint, rather than only average speed during the entire sprint.

Appendix C

First Amendment Retention Examination

1. Which of the following best describes the "check on government" rationale of the First Amendment's Speech Clause?
    a. Freedom of speech promises transparency in government so that abuses will be uncovered
    b. Freedom of speech helps individuals feel fulfilled
    c. Freedom of speech helps us find the truth
    d. Freedom of speech gives us information we need to make decisions
2. Which of the following best describes the "facilitating democracy" rationale of the First Amendment's Speech Clause?
    a. Freedom of speech allows us to let off steam so we don't resort to violence
    b. Freedom of speech gives us information we need to make decisions and govern ourselves
    c. Freedom of speech helps individuals feel fulfilled
    d. Freedom of speech helps us find the truth
3. Which of the following best describes the "search for truth" rationale of the First Amendment's Speech Clause?
    a. Freedom of speech allows us to let off steam so we don't resort to violence
    b. Freedom of speech helps individuals feel fulfilled
    c. Freedom of speech gives us information we need to make decisions
    d. Freedom of speech allows ideas to compete with one another and let people decide what is important

4. The First Amendment Speech Clause accomplishes the following:
    a. Prohibits the government from passing any law that impacts communication.
    b. Prohibits government actors from abridging speech without justification.
    c. Preserves the right to free exercise of religion.
    d. Allows the press to enjoy rights that no one else gets.
5. The word "Congress" in the First Amendment has been interpreted by the Supreme Court to mean:
    a. Federal legislators
    b. State police
    c. Local government officials
    d. All of the above
6. The First Amendment Press Clause accomplishes the following:
    a. Essentially nothing, according to the U.S. Supreme Court
    b. Allows journalists immunity from defamation lawsuits
    c. Helps reporters enjoy rights the rest of us don't have
    d. None of the above
7. The work of the federal legislative branch is found in which of the following?
    a. Judicial opinions like U.S. Reports
    b. Administrative agency materials like the U.S. Code of Federal Regulations
    c. Executive branch documents like executive orders
    d. Statutes such as the U.S. Code
8. Which of the following is least likely to be protected as freedom of speech?
    a. A woman burns a U.S. flag on courthouse steps to protest the war.

b. A person wearing a giant pizza-slice costume holds a sign advertising a pizza company and stands on a downtown sidewalk.

c. A man who is upset about the Legislature's speed-limit rule drives 20 miles per hour over the limit in order to make his point.

d. A group of people who are upset at the school board hold a vigil on a public sidewalk and don't say anything but just hold stacks of banned books.

9. Which of the following best describes "precedent" in the U.S. legal system?

   a. The law must evolve with society but should do so slowly in order to preserve predictability.

   b. The burden of proof in a criminal case is evidence beyond a reasonable doubt.

   c. Administrative agencies make law (legislative function), enforce law (executive function) and decide disputes (judicial function).

   d. When there is a law you disagree with, you have a right not to obey it.

10. A federal prisoner is told by prison staff that he can no longer write letters from prison and submit them to a newspaper for publication on the outside. His previous letters criticized prison disciplinary procedures, complained about bad prison food and contended that prisoners were mistreated. He decides to challenge the restriction as a violation of his free-speech rights. Write a paragraph in favor of his right to publish, using the "check on government" rationale.

Appendix D

Professor Self-Evaluation

Please answer the following questions as best as you can in the space provided. As you answer, feel free to address your experience with each experimental condition separately.

1. To what extent did you feel the Spirit as you taught?
2. To what extent do you think the students felt the Spirit as you taught?
3. What will you remember most from this teaching experience?
4. To what extent did you share your own personal experiences as a way of making concepts clearer, more applicable, and/or more memorable?
5. To what extent do you feel you developed a caring relationship with the students?
6. To what extent did you try to see and communicate connections between the gospel and your subject matter?
7. How well do you think the students would do when they take a quiz on the material?
8. Did the training influence how you will teach from now on? Please explain.

Appendix E

Study Description for Recruitment

There are multiple locations and times this study is being held, so check both the location and time that you sign up for. Please arrive promptly. You will arrive and first read over and sign the consent forms. You will then read over the material that will be discussed in the class session. After 20 minutes, a teacher will arrive and lead a class session on the material you just read. The teacher will teach for 30 minutes and then you will fill out a questionnaire about your experience. Upon completion of the questionnaire, you may leave. Exactly one week after the class session you will be contacted via email and directed to a website containing the second part of the study which will ask further questions about the class session. You must both attend the class session and complete the online portion one week later in order to receive any credit for your participation.

Appendix F

Preparation Handout

1) Creating a community of learning

What can you do to show the students that you care about them and about learning this material?

2) Topic 1

What gospel connections would help students to see the importance of this concept?

Think of a <u>personal</u> example or story that would show how you have experienced this concept and that would reveal something important about your values, your weaknesses, and/or your gospel commitments.

3) Topic 2

What gospel connections would help students to see the importance of this concept?

Think of a personal example or story that would show how you have experienced this concept and that would reveal something important about your values, your weaknesses, and/or your gospel commitments.

4) Topic 3

What gospel connections would help students to see the importance of this concept?

Think of a personal example or story that would show how you have experienced this concept and that would reveal something important about your values, your weaknesses, and/or your gospel commitments.