Do the Messages Matter? An Investigation of Classroom Messages and College Students’ Personal Theories about Education

by

Michael M. Barger

Department of Psychology & Neuroscience
Duke University

Date:_______________________

Approved:

___________________________
Rick Hoyle, Supervisor

___________________________
Lisa Linnenbrink-Garcia

___________________________
Mark Leary

___________________________
Makeba Wilbourn

___________________________
Jeffrey Greene

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology & Neuroscience in the Graduate School of Duke University

2016
ABSTRACT

Do the Messages Matter? An Investigation of Classroom Messages and College Students’ Personal Theories about Education

by

Michael M. Barger

Department of Psychology & Neuroscience
Duke University

Date:_______________________

Approved:

___________________________
Rick Hoyle, Supervisor

___________________________
Lisa Linnenbrink-Garcia

___________________________
Mark Leary

___________________________
Makeba Wilbourn

___________________________
Jeffrey Greene

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology & Neuroscience in the Graduate School of Duke University

2016
Abstract

Students hold a number of personal theories about education that influence motivation and achievement in the classroom: theories about their own abilities, knowledge, and the learning process. Therefore, college instructors have a great interest in helping to develop adaptive personal theories in their students. The current studies investigated whether specific messages that instructors send in college classroom might serve as a mechanism of personal theory development. Across 2 studies, 17 college instructors and 401 students completed surveys assessing their personal theories about education at the beginning and end of college courses. Students and instructors reported hearing and sending many messages in the classroom, including instructor help messages, conciliatory messages, uncertainty in the field messages, differential ability messages and generalized positive and negative feedback. Between-class and within-class differences in message reports were associated with students’ personal theories at the end of their courses, controlling for initial personal theories. Students’ initial personal theories were also related to the messages students reported hearing. The findings demonstrate the utility of assessing non-content messages in college classrooms as potential mechanisms for changing students’ personal theories in college. Implications for research and practice are discussed.
Contents

Abstract ......................................................................................................................................... iv

List of Tables .............................................................................................................................. viii

List of Figures ............................................................................................................................ ix

Acknowledgements .................................................................................................................... x

1. Introduction and Literature Review ....................................................................................... 1
   1.1 Personal Theories about Education........................................................................... 2
   1.2 Development of Personal Theories .............................................................................. 6
   1.3 Messages and Beliefs ........................................................................................................ 8
      1.3.1 Existing evidence ......................................................................................................... 8
      1.3.2 Messages in college as mechanisms of personal theory change: Gaps in the literature .............................................................................................................................. 13

2. Current Studies ........................................................................................................................ 17

3. Study 1 ...................................................................................................................................... 26
   3.1 Method ............................................................................................................................. 26
      3.1.1 Participants ................................................................................................................. 26
      3.1.2 Procedure .................................................................................................................... 27
      3.1.3 Measures ..................................................................................................................... 28
         3.1.3.1 Students’ personal theories ............................................................................... 28
         3.1.3.2 Instructor personal theories ............................................................................. 29
         3.1.3.3 Messages for students and instructors ........................................................... 29
         3.1.3.4 Demographics ..................................................................................................... 31
4.2.4.3 Using post-exam message codes to predict changes in personal theories. 77

4.3 Discussion of Study 2 ........................................................................................................ 83

4.3.1 Relations between messages and beliefs .................................................................... 84

4.3.2 Differences between classes ...................................................................................... 89

4.3.3 Development of personal theories in college ............................................................. 91

5. General Discussion ......................................................................................................... 96

5.1 Comparing and Contrasting Results from Study 1 and 2 ........................................ 97

5.2 Insights into the Development of Personal Theories ................................................... 102

5.3 Limitations and Future Directions ............................................................................. 106

5.4 Implications for Practice .............................................................................................. 108

6. Conclusion ...................................................................................................................... 110

Appendix A ................................................................................................................................ 111

Appendix B ................................................................................................................................ 115

Appendix C ................................................................................................................................ 116

References ................................................................................................................................ 117

Biography ................................................................................................................................... 128
List of Tables

Table 1: Comparison of Studies ................................................................................................ 19

Table 2: Research Questions and Planned Analyses............................................................... 21

Table 3: Hypothesized Relations for Inclusion in Study 1 Multilevel Models....................... 25

Table 4: Descriptive Statistics for Messages at the Class Level in Study 1 ......................... 32

Table 5: Full Results of Multilevel Models ........................................................................... 36

Table 6: Full Results of Instructor Message Models in Study 1 .......................................... 40

Table 7: Prediction of Student-reported Messages Using Initial Personal Theories .......... 42

Table 8: Descriptive Statistics for Personal Theories and Student-Reported Messages by Class .............................................................................................................................................. 52

Table 9: Differences in Uncertainty in the Field Messages by Initial Cluster and Shift .... 69

Table 10: Differences in Instructor Help Messages by Initial Cluster and Shift................. 70

Table 11: Frequency of Student-Reported Instructor Post-exam Messages....................... 76

Table 12: Hierarchical Regressions Predicting T2 Personal Theories from Post-exam Messages....................................................................................................................................... 79

Table 13: Hierarchical Regressions Predicting T3 Personal Theories from Post-exam Messages....................................................................................................................................... 82

Table 14: Complete Summary of Between- and Within-Class Findings from Study 1 & Study 2 .......................................................................................................................................... 97
List of Figures

Figure 1: A visual depiction of how class-level average beliefs changed from the beginning to the end of the summer course in Study 1 ........................................................ 33

Figure 2: The path model tested in Study 2 ................................................................. 50

Figure 3: Supplemental model for Study 2 .................................................................. 53

Figure 4: Full model of personal theories and messages for chemistry classroom .... 59

Figure 5: Full model of personal theories and messages for neuroscience classroom .. 59

Figure 6: Supplemental model distilling messages ....................................................... 63

Figure 7: The five cluster solution. All responses were given on a five point scale with an unlabeled midpoint .......................................................... 66

Figure 8: Interaction between generalized positive feedback and exam performance on T2 incremental theories of intelligence ............................................................ 80

Figure 9: Interaction between reasoned conciliatory messages and exam performance on T3 perceived competence ................................................................. 83
Acknowledgements

The world that scientists wish to describe is complex and messy. It would be impossible to conduct science in all of this mess without the assistance of enthusiastic and supportive colleagues, family members, and friends. I have many people to thank for the completion of this research. First, my advisors, Rick Hoyle, Lisa Linnenbrink-Garcia, who have given so much time and energy to help me succeed. They have also helped me develop my own “personal theories” about the world of academia for which I am forever grateful. I then must thank my other committee members Mark Leary, Makeba Wilbourn, and Jeff Greene, who have strengthened and broadened this project with their own expertise and by getting me to dig deep into developmental theory. I would also like to thank my colleagues, Stephanie Wormington, Madeleine George, and Kristin Johnson, who provided valuable feedback and endless emotional support throughout this long process.

I thank the Duke Graduate School for a Summer Research Fellowship that allowed me the freedom to conduct the Study 1 during the summer of 2015. Without the numerous instructors and students who were willing to give up their valuable class time, this research would not have been possible. A number of research assistants made collecting and coding data possible, including Anita Restrepo, Stephanie Banks, and Carolyn Fishman.
And finally and most importantly, I thank my wife, Anna Doolittle-Barger. While she has supported this accomplishment more than any other person, earning her support remains my greatest accomplishment.
1. Introduction and Literature Review

College educators take pride in being able to change the way their students think. The purpose of higher education is not merely to provide content knowledge, but to fundamentally change students’ thinking and learning (Astin, 1993; Perry, 1970). The information that college education provides to students can be fleeting, but the beliefs that develop in college have the potential to last beyond the confines of the educational context. Beliefs about one’s own abilities, the nature of intelligence, and knowledge play a critical role in how people adapt in a complex and ever-changing world (Barger & Linnenbrink-Garcia, under review).

But how can college instructors change their students’ beliefs, specifically? Imagine sitting in a college classroom. The instructor might tell students that they should ask the instructor for help if they get confused, or that it is not very serious if someone performed poorly on one test. While on the surface, these messages seem commonplace, innocuous, and even positive, they actually tell students something about what the instructor believes about the education process. Specifically, these messages might signal that the students require the assistance of instructors in order to construct knowledge and that a bad grade is okay because students cannot do much to change their level of ability. Many of the things instructors say in the classroom have little to do with content, but everything to do with their own beliefs about the education process. If the objective of college education is to create life-long learners and critical thinkers,
researchers and educators must aim to find how certain messages instill adaptive or maladaptive systems of beliefs that last beyond a single classroom.

Students have many beliefs about education (Dweck & Leggett, 1988; Harter, 2012; Nicholls, 1992; Schommer, 1990), and there are many possible messages that instructors send in class that have the potential to change these beliefs. Given the multifaceted nature of belief development, I begin by organizing the process using a theoretical model that treats beliefs about education as personal theories and describing three general types of personal theories students form about education. I then outline the evidence supporting the hypothesis that college students use messages to inform and change their personal theories. Finally, I describe the remaining gaps in the literature and introduce the current studies.

1.1 Personal Theories about Education

In order to understand the development of college students’ beliefs about education, I utilize a psychological framework that is often applied describing the cognitive development of much younger learners: theory theory (Gopnik, 2012; Gopnik & Wellman, 1994; Wellman, 1990). According to theory theory, children form abstract cognitive representations about the physical world, social relationships, and the mental states of others by gathering evidence from their interactions with the world. While this theoretical approach has traditionally been applied to children, it is my assertion that the same cognitive process drive personal theory change throughout the lifespan. In this
view, students of all ages have abstract cognitive representations that guide their thinking and action, and while personal theories become more solidified over time, new evidence is continuously folded into their personal theories (Barger & Linnenbrink-Garcia, under review).

Personal theories of education can be divided into three general categories: theories about the self, theories about knowledge, and theories about the learning process. Theories about the self have been studied in many forms (Leary & Tangney, 2003), though perhaps the best studied within a school context is academic self-concept, or the beliefs students have about their abilities in school (Bong & Skaalvik, 2003; Harter, 1988; Marsh & Seaton, 2013). A central component of academic self-concept is perceived competence (Harter, 1982), such that students with positive academic self-concept would believe that they are generally competent in school. Students might also have more specific self-concepts about particular topic areas (Byrne & Shavelson, 1996; Marsh, Byrne, & Shavelson, 1988; Marsh, Craven, & Debus, 1998) or make specific predictions about their ability to do well on a task, often called self-efficacy (Bandura, 1977).

Students also possess theories about the structure and source of knowledge, called personal epistemology (Hofer & Pintrich, 2002) or epistemic/epistemological beliefs (Muis, 2007; Schommer, 1990, 1993). There are many different frameworks for conceptualizing the way students’ view knowledge and the field seems to be expanding to include all aspects of epistemic cognition (e.g., Chinn, Buckland, & Samarapungavan,
2011; Greene, Azevedo, & Torney-Purta, 2008). However, one framework that fits well with theory theory emanates from Schommer’s (1990) model and Hofer and Pintrich’s (1997) model, in which students’ cognitive representations of knowledge are divisible into independent beliefs (see also Greene, Torney-Purta, & Azevedo, 2010; Hofer, 2000). For example, some students believe that knowledge is composed of a collection of unchanging facts, while others might think knowledge is more complex and evolving. Other beliefs deal less with the nature of knowledge and more with how it is justified. Students might believe that knowledge is handed down by authority figures, or that knowledge is justified through personal experiences, or both. Students’ personal theories about knowledge also have domain general and domain specific components (Muis, Bendixen, & Haerle, 2006) such that students have overarching theories about knowledge that also vary for knowledge in different disciplines.

The final category of personal theories involved the learning process, the transmission of knowledge to the self. Two well-studied beliefs that fall under this category are beliefs that if learning is a quick process (referred to here as quick learning theories), and that if something is not learned quickly than it will not likely be learned at all (Schommer, 1990) and beliefs about the nature of intelligence (Dweck & Leggett, 1988). Some students endorse an entity theory of intelligence, believing that intelligence is a fixed trait that does not change very much, while other students believe that intelligence can grow through effort, an incremental theory.
Personal theories matter in the classroom because they are the blueprints students use to predict what kind of actions to take in the classroom in order to succeed, avoid needless effort, and meet any other educational or social goals. In educational contexts, students’ personal theories relate to a number of important predictors of academic success, including motivation (Bråten & Strømsø, 2004; Chen & Barger, 2016; Chen & Pajares, 2010; Dweck & Leggett, 1988; Greene, Costa, Robertson, Pan & Deekens, 2010; Paulsen & Feldman, 1999, 2005), emotion (King, McInerney, & Watkins, 2012; Muis et al., 2015), and self-regulation (Bråten & Strømsø, 2005; Burnette, O’Boyle, VanEpps, Pollack, & Finkel, 2012; Cury, Elliot, Da Fonseca, & Moller, 2006; Muis, 2007), as well as learning and achievement (Barger, Wormington, Huettel, & Linnenbrink-Garcia, under review; Blackwell, Trzesniewski, & Dweck, 2007; Marsh, 1990a; Qian & Alvermann, 1995; Schommer, 1993; Stipek & Gralinski, 1996; Trautwein & Lüdtke, 2007). Generally speaking, the beliefs that tend to be most adaptive are positive theories about the self, that knowledge is complex and changing, that knowledge is derived not only from authority figures, but from evaluating multiple knowledge claims, incremental beliefs about ability or intelligence, and the theory that learning takes time, though there are some notable exceptions to these patterns in certain contexts (e.g., Chen, Metcalf, & Tutwiler, 2014; Park & Kim, 2015). Because certain personal theories tend to be more adaptive in educational contexts, researchers have taken an interest in how personal theories develop and change.
1.2 Development of Personal Theories

Personal theories develop over time as students gain experience in the classroom. Much like a scientist gathers data to inform their own theories, children implicitly use their experiences in the classroom to form theories about themselves, about knowledge, and about the learning process. Because students’ experiences in the classroom widely differ, and students may interpret or experience this evidence differently as well, students form different theories over time.

Students’ experiences differ in part because different evidence is afforded to them by different contexts. Educational contexts vary by societal culture, the culture of a field of study, or even more specifically the climate of the individual classroom. At the broadest level of analysis, researchers can examine differences in personal theories as a product of broad, cultural forces. There are differences in people’s personal theories across different races and ethnicities (e.g., Chen, 2012; Chen & Pajares, 2010; Nichols, White, & Price, 2006) and countries of origin (e.g., Li, Fung, Bakeman, Rae, & Wei, 2014; Markus & Kitayama, 1991). In college, students are inundated in a new kind of broad context: the culture of their field of study. Beliefs about knowledge and the nature of intelligence differ across fields (Leslie, Cimpian, Meyer, & Freeman, 2015; Paulsen & Wells, 1998; Trautwein & Lüdtke, 2007) because the epistemic assumptions of different fields coalesce around a unified social epistemology (Goldman, 2011) and the abilities valued in each field vary considerably. For example, the way that truth is constructed in
science is very different from how truth is established in history. These differences in students’ personal theories in various cultural contexts supports the notion cultural contexts afford students with different evidence for the way education works.

At a smaller unit of analysis, individual classrooms also vary in the kinds of theories they support. For example, a classroom’s epistemic climate describes the way knowledge is portrayed and spoken about in the classroom and the interplay between teachers’ and students’ beliefs (Bendixen & Rule, 2004; Feucht, 2010). Different epistemic climates foster different kinds of personal theories about knowledge (Muis & Duffy, 2012). Researchers have also examined “perceptions of environmental entity theory” (Good, Rattan, & Dweck, 2012, p. 708) in classrooms, the extent that the classroom environment supports a fixed theory of intelligence, as measured by students’ perceptions of the personal theories of those around them.

While differences between cultures, fields of study, and classroom climates are important to understanding how students develop different personal theories, these findings fail to describe the mechanism through which differences in the environment cause changes in students’ beliefs over time. It is not clear what specific experiences and forms of evidence students have in these different contexts and environments that lead them to construct different personal theories. One potential form of evidence that student implicitly use to adjust their theories, and thus a potential mechanism for how contexts shape students’ beliefs, are the messages they hear. The words that parents and
teachers say to their children and students convey information about the way the world works, which students then use to update their personal theories about education.

1.3 Messages and Beliefs

There is both correlational and experimental evidence to suggest that subtle messages can change the way individuals think about education and knowledge. There are many possible messages, but in many cases, researchers focus on one type of message at a time.

1.3.1 Existing evidence

An increasing number of studies have demonstrated that manipulating the messages children hear can have broad effects on their personal theories about intelligence and the actions associated with these beliefs. Some of the earliest evidence for this came from the study of the effects of different forms of praise on fifth-grade students (Mueller & Dweck, 1998). When children received person-praise in a series of laboratory experiments, such as “you are so smart,” as opposed to process praise, such as “you worked so hard,” they endorsed entity beliefs about intelligence, that intelligence is a fixed trait, more strongly. The implication of this study is that the type of praise provides evidence to the student about whether a person’s intelligence is fixed or can change. One study of 10-year-old children extended these laboratory findings to the real world by measuring daily maternal praise and children’s beliefs about intelligence (Pomerantz & Kempner, 2013). They found that mothers’ self-reported use
of daily person-praise (sending the message that the person’s ability is more worthy of
praise than the effort) predicted their children’s entity theory six months later, even
while controlling for initial beliefs. In another longitudinal study that followed
preschool-aged children into elementary school, researchers found that parents’ process
praise with their preschool-aged children predicted children’s incremental beliefs about
intelligence at ages seven-eight (Gunderson, Gripshover, Romero, Dweck, Goldin-
Meadow, & Levine, 2013). Seemingly miniscule experimental manipulations in phrasing
can influence student motivation, such as generic language (e.g., “boys are good at this”)
versus non-generic language (e.g., “that boy is good at this;” Cimpian, 2010), and even
overinflated praise (e.g., “You made an incredibly beautiful drawing!”) versus non-
inflated praise (e.g., “You made a beautiful drawing!” Brummelman, Thomaes, de
Castro, Overbeek, & Bushman, 2014). Furthermore, many studies have found that
experimentally manipulated messages in the form of brief psychosocial interventions
can change personal theories about the changeability of traits, like intelligence and
personality in the transition to middle school and high school (Blackwell et al., 2007; Miu
& Yeager, 2014; Yeager et al., 2014). These studies demonstrate how brief messages, in a
laboratory, through an intervention, or accumulated over time in real world settings,
have the potential to change the way that children think about education in different
contexts and long after the message was actually delivered.
Unlike the preceding research on theories about intelligence, researchers have not studied specific messages about knowledge as developmental mechanisms for personal theories about knowledge. However, some studies do suggest this is the case. First, one study of four- to six-year-old children found that when parents used evaluativist talk (weighing multiple sources) with their children about science topics, the children were more likely to talk about evidence when making claims (Luce, Callanan, & Smilovic, 2013). In another study of linguistic cues (Bonawitz et al., 2011), researchers examined whether preschool-aged children use evidence present in demonstrative language (e.g., “This is what the toy does”) versus surprised language (e.g., “Oh! The toy does this.”). Intriguingly, children in the latter condition explored the toy more than children who heard demonstrative language. One possible interpretation, from an epistemological perspective, is that the surprise from the experimenter provided evidence that the adult did not know everything about the toy, meaning it would be a good use of the children’s time to learn about the toy by exploring it on their own. Therefore, it is possible that subtle differences in messages may cause differences in theories about the source of knowledge construction as well.

Together, these studies suggest that young children are folding the underlying content of messages into their personal theories, but do older students do this as well? Although it is likely true that early experiences are particularly foundational in forming children’s personal theories, theories have the potential to change as new evidence is
presented. This might be particularly true when students progress to a new level of education and find themselves with new frames of reference, such as the transition to middle school, high school, or college (Davis, 1966; Eccles et al., 1993; Perry, 1970). When students reach a new stage, their existing beliefs might be challenged by the new context (Brownlee, Walker, Lennox, Exley, & Pearce, 2009; Paulsen & Wells, 1998). Changes in personal theories have been observed in undergraduate students (Barger et al., under review; Flanigan, Peteranetz, Shell, & Soh, 2015; Muis & Duffy, 2013; Perry, 1970; Trautwein & Lüdtke, 2007). In fact, the “less-crystallized” nature of college students’ attitudes and identity has been used as a criticism of the widespread use of undergraduate students in psychological research (i.e., Sears, 1986). A great deal of work has determined that late adolescence and early adulthood are critical times of identity development (Arnett, 2000; Erikson, 1963; Harter, 2012), suggesting that students’ representation of themselves, including their self theories, are changing in the college setting.

This also seems to be true for beliefs about knowledge (Perry, 1970). For example, students who begin college concentrating in “hard” fields increasingly report that knowledge is simple and certain, whereas students concentrating in “soft” science courses, such as arts and social sciences increasingly reject this belief (Trautwein & Lüdtke, 2007). As another example, one quasi-experimental intervention study in two college statistics classrooms found that the instructor that consistently sent messages
about the nature of knowledge as changing, complex, and constructed led students’ to more strongly believe that knowledge is complex and changing over the course of a semester (Muis & Duffy, 2013).

In terms of beliefs about intelligence, researchers have suggested (Dweck, 2002) and found that these beliefs are relatively stable during college (across years, \(0.64 < r < 0.74\); Robins & Pals, 2002). However, one study found a substantial increase in entity beliefs in a single semester in a sample of computer science students (Flanigan et al., 2015). Furthermore, one laboratory study found that entity messages about giftedness have an impact on achievement-related behavior, specifically likelihood to self-handicap after failure, even in undergraduate students (Snyder, Malin, Dent, & Linnenbrink-Garcia, 2014), similar to the effects of praise on younger children (Mueller & Dweck, 1998). Thus, the assertion that theories about intelligence do not change beyond early education is far from proven. In summary, students’ various personal theories about education change within the college context. I propose that the nature of these changes resemble how changes in personal theories occur earlier in life, and that one form of evidence for personal theory change comes in the form of messages from instructors.

Little research has been done on the messages that college instructors send in class, though there is reason to believe that they differ from the messages that change younger children’s beliefs (after all, most professors do not spend class time praising students for how clever they are). Some studies have investigated specific messages
from the college classroom that may play a role in shaping students’ personal theories. For example, when an instructor uses conciliatory language after their students perform poorly, such as “It’s OK that you had a bad test,” students perceive this as a sign of the instructors’ entity theories of intelligence (Rattan, Good, & Dweck, 2012). Therefore, messages such as these may foster entity theories in the students that hear them as well. Researchers have also observed micro-aggressions based on intelligence in multiple college classrooms (Suárez-Orozco et al., 2015). Micro-aggressions (e.g., “You need to do it like you are in kindergarten”) might send subtle messages about what the instructor really thinks about students’ ability or the nature of intelligence. However, examining the kinds of evidence-laden messages that are sent in by college instructors deserves much closer investigation as part of the study of undergraduate students’ personal theory development.

1.3.2 Messages in college as mechanisms of personal theory change: Gaps in the literature

While a great deal of evidence shows that messages change students beliefs about education, both in immediate ways and over the course of time, several questions remain when considering the role that teachers may play in changing college students’ personal theories. In line with the theoretical assumptions of theory theory that have been investigated in earlier developmental stages (Gopnik, 2012; Gopnik & Wellman, 1994), I assert that college students’ personal theories change through the same basic mechanism: by collecting evidence from their social environments and adjusting their
theories accordingly. The objective of the current studies is to demonstrate that students’ various beliefs about education function as personal theories; that students use evidence from the messages their teachers and adjust their personal theories accordingly. Thus, the more messages an instructor sends with implicit or explicit information about students’ abilities, the nature of knowledge, and the nature of learning, the greater the changes in students’ personal theories about education.

There are four large gaps in the extant literature that need to be addressed to support and test the college-student theory theory: (1) what kinds of messages are sent in college classrooms, (2) the complexity of messages in the classroom and the various kinds of evidence they provide in a real-world context, (3) a need to distinguish between classroom-level effects and individual effects of messages, and (4) the role individuals’ initial personal theories play in the process. First, there is a clear distinction in the literature between experimental work in the laboratory and field work in the classroom, particularly in college classrooms. Messages in the lab change beliefs, and students’ beliefs seem to change over the course of schooling, but it is unclear what kinds of messages college instructors naturally send their students. There is virtually no evidence of types of epistemic messages college instructors send to their students. Do instructors send messages with information about the nature of knowledge, intelligence, and the self, and do students’ receive these messages?
Second, researchers have not looked at a broad range of messages as a mechanism of change in undergraduates students’ different types of personal theories, particularly those about knowledge. A single message may contain information that adjusts a wide range of personal theories, but many of the existing studies link a single message to a single personal theory. The classroom is a very rich environment, and studies should consider this. Which instructor messages play a role in changing which of students’ personal theories?

Third, much of the existing research assumes a model in which a teacher sends a message and students in the class hear that message and its underlying content identically, which is subsequently assimilated into each student’s personal theory about education. However, a more appropriate metaphor for classroom messages may be a message in a bottle thrown out to sea (Perrenoud, 1998); there is no guarantee that any one message will reach any particular student. Therefore, instructor messages might change personal theories because instructors send different messages (between-class effects), because students in the class hear different messages (between-class effect), or because individual students within a class hear or remember messages differently from their classmates (within-class effect). Through which of these processes do messages change personal theories?

The fourth gap in the literature relates to the understudied possibility of within-class differences. If students within the same class differ on how likely they are to
remember certain messages, researchers need to understand why. One possibility is that students with particular personal theories at the beginning of class are more likely to remember messages with content that aligns with that theory. This would align with psychological research that demonstrates people are more likely to attend to and remember information that corresponds to their existing beliefs, sometimes referred to as confirmation bias (Nickerson, 1998). As an example, a student who believes that knowledge changes over time might be more likely to report hearing an instructor say that issues in the field are unsettled. Because college students already have over a decade of educational experiences that they have used to form their personal theories, their theories likely tint the evidence available to students, complicating the development of personal theories. Do students with different personal theories remember classroom messages differently? The studies proposed here will begin to fill these gaps necessary to support the college student theory theory and stimulate new research on how college teachers can foster different beliefs in their students by sending specific messages.
2. Current Studies

In order to address these gaps in the literature, I conducted two complementary studies. Each study involves measuring undergraduate students’ beliefs at the beginning and end of a college course at Duke University as well as the messages that instructors might send throughout the course through student self-report. Study 1 investigated instructors and students of science, technology, engineering, and mathematics (STEM) courses taught during Duke’s summer sessions. Study 2 examined students in two large lecture classes in chemistry and neuroscience.

Each of the studies offer unique advantages in terms of both study design and theory (see Table 1 for a comparison of the studies). Study 1 was uniquely situated to examine both between-class and within-class differences in student-reported messages, instructor-reported messages, and students’ changes in beliefs. Summer courses also have the benefit of being short, intensive courses in which most students at Duke are only taking one class at a time\(^1\), so that changes in beliefs might be more logically tied to the events in the course. However, the numerous domains of the courses are an unavoidable confound in this study. Study 2 compliments the findings from Study 1 as a way of best addressing the four gaps in the literature. Study 2 was conducted in two larger classes with a single core domain each to better explore the interplay of within-

\(^1\) There was a small sample of students enrolled in more than one class. Students who were enrolled in multiple classes that were a part of Study 1 were asked not to participate a second time.
class differences in perceived messages and students’ initial and final personal theories. The larger number of students across the two classes ($N = 239$) allowed for the use of multiple-group path analysis. Both classrooms were introductory courses, such that for the majority of the students, this class will be their first college experience with chemistry and neuroscience. One classroom (chemistry) also contains almost exclusively first-year students, which possesses the added benefit of having a homogenous group of students at a critical time of transition into college. The structure of these courses also allowed for interesting supplemental analyses as well.
Table 1: Comparison of Studies

<table>
<thead>
<tr>
<th></th>
<th>Study 1: Chemistry Class</th>
<th>Study 2: Neuroscience Class</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>Instructors ($N = 15$) and students ($N = 162$) in 15 summer courses</td>
<td>Students ($N = 112$) in large class</td>
</tr>
<tr>
<td><strong>Participant grade-level</strong></td>
<td>10% first-year, 40% second-year, 31% third-year, 19% fourth-year/+</td>
<td>Primarily first-year students</td>
</tr>
<tr>
<td><strong>Course structure</strong></td>
<td>6-week summer courses</td>
<td>Project-based learning with lecture component</td>
</tr>
<tr>
<td><strong>Domain(s)</strong></td>
<td>Various STEM courses: biology, chemistry, math, psychology, statistics</td>
<td>Chemistry (Introductory)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>Randomly selected audio recordings in 9 classes</td>
<td>Targeted audio recordings</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td>Measure instructor’s reported messages</td>
<td>Supplemental post-exam survey</td>
</tr>
</tbody>
</table>

Through these complimentary studies, I sought to answer three principle questions: (1) what are the messages college instructors send in class?, (2) do messages in class predict changes in students’ personal theories about education?, and (3) do students’ initial beliefs predict whether they report class messages?
These questions were answered using multiple analytical techniques (see Table 2 for a summary of the analyses and their corresponding research questions). Research question 1 involved factor analyzing self-report items from a pilot of college instructors and students and Study 1 participants to find reliable message factors. To answer the second and third research question, a series of multilevel models was used in Study 1 and structural equation models were used in Study 2.
Table 2: Research Questions and Planned Analyses

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Hypotheses</th>
<th>Data Source(s)</th>
<th>Analytic Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RQ1 What messages are send in class?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(RQ1a) What messages do instructors report sending in class?</td>
<td>Up to 10 hypothesized factors</td>
<td>Pilot data</td>
<td>Exploratory factor analysis</td>
</tr>
<tr>
<td>(RQ1b) What messages do students report hearing in class?</td>
<td>Up to 10 hypothesized factors</td>
<td>Pilot data, Study 1, Study 2</td>
<td>Confirmatory factor analysis</td>
</tr>
<tr>
<td>(RQ1c) Do instructor-reported, student-reported, and observed messages correspond to one another?</td>
<td>Yes, though degree of correspondence may vary</td>
<td>Study 1</td>
<td>Correlation</td>
</tr>
<tr>
<td>(RQ1d) What messages about exam performance due students remember immediately after receiving feedback?</td>
<td>Conciliatory messages, generalized feedback</td>
<td>Study 2</td>
<td>Qualitative coding</td>
</tr>
<tr>
<td>Research Questions and Planned Analyses (cont.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RQ2 Do messages predict belief changes?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(RQ2a) Do class-average student-reported messages predict changes in students' personal theories?</strong></td>
<td>Yes, see Table 3 Study 1 Multilevel modeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(RQ2b) Do within-class differences in student-reported messages predict changes in students' personal theories?</strong></td>
<td>Yes, see Table 3 Study 1 Multilevel modeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(RQ2c) Do instructor-reported messages predict changes in students' personal theories?</strong></td>
<td>Yes, see Table 3 Study 2 Structural-equation modeling, ANOVA with ISOA shifts, Regression analyses</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(RQ2d) Are the effects of instructor support and spurring messages distinguishable from instructor help and conciliatory messages?</strong></td>
<td>Yes Study 2 Structural-equation modeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RQ3 Do initial beliefs predict the messages students recall hearing?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(RQ3a) Do students’ initial beliefs predict what messages they report hearing in class?</strong></td>
<td>Yes Study 1 Multilevel modeling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
I predicted that college students would report hearing many different types of messages, that instructors would report saying multiple types of messages, and that qualitative observational data will corroborate the self-reports. The messages investigated in these studies were chosen based on theoretical assumptions about different types of personal theories, discussion with undergraduates, and pilot testing with undergraduate students and instructors. As very few studies exist about the kinds of messages college instructors send, I began by pilot testing a broad set of messages that could theoretically send information that a student could use to adjust their personal theories about education. Specifically, the messages of interest fall into six categories: instructor help messages (e.g., “Ask me for help if you don’t understand.”), differential ability messages (e.g., “This is more difficult for some than others.”), conciliatory messages (e.g., “It’s OK that you got it wrong.”), uncertainty in the field messages (e.g., “We don’t really have an answer to that yet.”), generalized negative feedback (e.g., “This class is not performing up to my standards.”), and generalized positive feedback (e.g., “Everyone did a good job on the test.”).

I hypothesize that the frequency of these messages in class will predict changes in personal theories in systematic ways based on the evidence each message might provide a student (see Table 3 for a summary of which messages are expected to relate to changes in specific messages and will be included in initial models). Some of these
hypotheses were highly speculative, particularly in predicting changes in personal theories about knowledge, given the paucity of research on this topic. Instructor help messages, which students may interpret as evidence that they will need help because they are not smart enough, or because learning takes time, are expected to decrease incremental theories of intelligence, the belief in quick learning, and perceived competence. Constantly telling students to ask the instructor for help may also imply that the instructors are a more valid source of knowledge than solving problems on one’s own or with classmates, providing reason for students to increase their beliefs in justification by authority and decrease their beliefs in the personal justification of knowledge. Differential ability messages could provide students with evidence that intelligence is a fixed trait and that if learning does not happen quickly then it will not happen for some, decreasing incremental theories of intelligence and perceived competence while increasing quick learning beliefs, but also might provide evidence to students that only certain people (e.g., experts) should be trusted to justify knowledge, increasing justification by authority but decreasing personal justification beliefs. Conciliatory messages are expected to correspond with reductions in incremental theories of intelligence (Rattan et al., 2012), but also might communicate to students that they were not expected to be capable of high achievement, predicting decreases in beliefs in personal justification and perceived competence. Messages that there is uncertainty in the field implies a constructivist view of the world (Taylor, Fraser, &
Fisher, 1997), in which knowledge is constructed by people, as opposed to facts objective reality, thus these messages are expected to predict decreases in beliefs about simple/certain knowledge and justification by authority but increase personal justification beliefs. Generalized negative and positive feedback is expected to predict decreases and increases in students’ perceived competence, respectively, in line with research on feedback (e.g., Craven, Marsh, & Debus, 1991; Pintrich & Blumenfeld, 1985). Generalized feedback might also present evidence that learning happens quickly, such that negative feedback predicts decreases the belief in quick learning and positive feedback increases this belief.

Table 3: Hypothesized Relations for Inclusion in Study 1 Multilevel Models

<table>
<thead>
<tr>
<th>Concept</th>
<th>HELP</th>
<th>DIFF</th>
<th>CONC</th>
<th>UNCE</th>
<th>NEGF</th>
<th>POSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental Theories</td>
<td></td>
<td></td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td></td>
</tr>
<tr>
<td>Quick Learning</td>
<td>⬤</td>
<td>⬤</td>
<td></td>
<td>⬤</td>
<td>⬤ ⬤</td>
<td></td>
</tr>
<tr>
<td>Simple/Certain Knowledge</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Justification by Authority</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Personal Justification</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
<tr>
<td>Perceived Competence</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
<td>⬤</td>
</tr>
</tbody>
</table>

*Note.* Red boxes indicate a hypothesized negative relation and green boxes indicate a hypothesized positive relation. HELP = Instructor help. DIFF = Differential ability. CONC = Conciliatory messages. UNCE = Uncertainty in the field. NEGF = Generalized negative feedback. POSF = Generalized positive feedback.
3. Study 1

3.1 Method

3.1.1 Participants

Duke summer term instructors ($N = 18$) in STEM fields (biology, chemistry, math, psychology, statistics) and their students were recruited to participate. Instructors of STEM courses with at least 8 students enrolled in their course over the summer received emails inviting them to participate in the project (response rate of 42%). Three classes were excluded because of the high enrollment of high school students and a shorter course length, giving a final sample of 15 classes. Instructors included faculty members ($n = 4$), post-doctoral instructors ($n = 4$), and graduate students ($n = 7$).

Student participants were undergraduates at Duke enrolled in the summer course who were at least 18 years old. Of the students enrolled in the 15 summer classes, 162 agreed to participate and completed surveys at the beginning and end of the class (81% retention rate). Classes included between 3 and 38 students who completed both surveys, though when excluding the minimum and maximum, all other classes contained 5 to 14 students. These students may not have been a perfect representation of the university as a whole; even so, students have a wide range of reasons to take summer courses. For example, students might be more likely to be double-majors trying to get ahead of their requirements or be more likely to participate in varsity sports, making an effort to lighten their course-load during their competitive season.
3.1.2 Procedure

Instructors were notified via email several weeks before the beginning of the summer session. Instructors were given opportunities to ask questions about the study and worked with the researcher to set aside class time for the study. During the first week of the course, a researcher was introduced by the instructor, described the project to students, and provided them a link to participate in the first survey. This survey measured students’ personal theories and demographic information. Students were informed that their instructor had agreed to participate, but that their own participation was optional and would not affect their grades.

Instructors received an email in the fourth week of the six-week summer course asking them to complete an instructor survey. The survey assessed instructor beliefs about education and the messages they remember saying in the class they were teaching. All instructors completed the survey within one week of receiving the email.

A subset of instructors (n = 9) additionally consented to have three of their classes audio recorded, accounting for approximately 10% of the course’s lectures. Classes were randomly selected and the instructor was notified of the selection three days in advance. In the event that the instructor was not lecturing on the randomly selected day (e.g., exam), a new day was selected. To record the class, the researcher placed an audio recorder near the instructor, left the classroom to reduce reactivity, and returned at the end of class to retrieve the recorder.
In the final weeks of the semester, a researcher returned to the classroom to provide students a link to the final student survey. The final survey assessed students’ beliefs about education as well as the messages they recalled their instructors’ saying during the term.

3.1.3 Measures

A complete list of measures can be found in Appendix A for Study 1 and Study 2, along with the reliabilities of the measures at the various time points in Study 1.

3.1.3.1 Students’ personal theories

Students completed six scales at the beginning and end of the course assessing their beliefs about education. Incremental theories of intelligence measured the extent to which the student believed that intelligence can increase through effort instead of just a fixed trait (8 items, e.g., “You can always substantially change how intelligent you are,” Dweck, 1999). Quick learning beliefs assessed the extent to which the student believed that learning happens quickly or not at all (6 items; e.g., “If something can be learned, it will be learned immediately,” DeBacker, Crowson, Beesley, Thoma, & Hestevold, 2008; Wood & Kardash, 2002). Domain-specific perceived competence measured the extent to which an individual believed they were able to succeed in the course’s subject area (5 items, e.g., “I can do even the hardest work in [subject] if I try,” Midgley et al., 2000).

Beliefs about knowledge were assessed by three scales adapted from two preexisting scales (Greene, Azevedo, Torney-Purta, 2008; Hofer, 2000) that have
demonstrated reasonable reliability in previous studies with similar populations (Barger et al., under review; Barger, Perez, Canelas, & Linnenbrink-Garcia, 2015). Simple/certain knowledge measures the extent to which a student believes knowledge is a collection of unchanging facts (8 items, e.g., “In [subject], what is a fact today will be a fact tomorrow.”). Justification by authority measures students’ beliefs that knowledge is handed down by authority figures (6 items, e.g., “Things written in [subject] textbooks are true.”). Personal justification measures the extent to which the student believes that what is right means different things to different people (3 items, “In [subject], the truth means different things to different people.”).

3.1.3.2 Instructor personal theories

Instructors responded to the same items as their students (with the exception of perceived competence) assessing theories of intelligence, quick learning, simple/certain knowledge, justification by authority, and personal justification.

3.1.3.3 Messages for students and instructors

Students (at the beginning of the semester) and instructors (in the middle of the semester) responded to a list of messages that instructors might say in class. Students were asked to report how often they recalled hearing the message or a message like it from their current instructor on a scale of 1 (never heard it) to 5 (hear it often). Instructors were asked to report how often they recalled saying the message or a message like it in class on a scale of 1 (never say it) to 5 (say it often). These scales were
created for the purpose of the study based on a combination of prior research findings and assumptions of what type of information might theoretically produce changes in students’ beliefs. An early version of the messages survey was pilot tested with 204 students and 81 college instructors. A factor analysis of the messages survey indicated that there were approximately 10 factors that might be reliably measured, including instructor help (e.g., “You should ask me questions if you get confused.”), lots of material (e.g., “We have a lot of material to get through today.”), more knowledge beyond course (e.g., “That’s too complicated for this class.”), independent learners (e.g., “It’s your job to figure that out.”), differential ability (e.g., “This is more difficult for some than others.”), conciliatory messages (e.g., “It’s OK that you got it wrong.”), single truth (e.g., “We know that this is true.”), uncertainty in the field (e.g., “We don’t really have an answer to that yet.”), generalized negative feedback (e.g., “This class is not performing up to my standards.”), and generalized positive feedback (e.g., “Everyone did a good job on the test.”). After the pilot study, items were added with the express intent of augmenting the measurement of these ten factors so that each subscale had at least three items.

In the present study, two subscales were not reliable for instructors or students (αs < 0.6: independent learners, single truth) and two were not related to any of the belief measures (lots of material, more knowledge beyond course), leaving the six message factors described in the hypotheses sections for subsequent analyses. A confirmatory factor analysis with the best items of the six factors (See appendix A) fit the
data well ($\chi^2(215) = 289.37, p < .001$, CFI = 0.96, RMSEA 90% CI [.031, .058], SRMR = .06), suggesting these messages were distinguishable and reliably measured.

### 3.1.3.4 Demographics

Both students and instructors reported their gender and race/ethnicity. Students were also asked about their year in school and instructors were asked to report their number of years of teaching experience.

### 3.2 Results

Instructors and students reported varying levels of messages in class (for average reports of classes and instructors for the six message factors, see Table 4). Correspondence between teacher reports of messages and their class average student-reports of the same messages varied ($N = 15$; instructor help, $r = .38$; differential ability, $r = .57^*$; conciliatory messages, $r = .47^!$; uncertainty in the field, $r = .76^*$; generalized negative feedback, $r = .13$; generalized positive feedback, $r = .52^*$; $^* p < .05$, $^! p < .10$).
Table 4: Descriptive Statistics for Messages at the Class Level in Study 1

<table>
<thead>
<tr>
<th></th>
<th>Class Averages of Student-Reports</th>
<th>Instructor-Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Instructor help</td>
<td>4.53</td>
<td>0.28</td>
</tr>
<tr>
<td>Differential ability</td>
<td>2.40</td>
<td>0.58</td>
</tr>
<tr>
<td>Conciliatory messages</td>
<td>3.08</td>
<td>0.65</td>
</tr>
<tr>
<td>Uncertainty in the field</td>
<td>2.96</td>
<td>0.79</td>
</tr>
<tr>
<td>Generalized negative feedback</td>
<td>1.49</td>
<td>0.31</td>
</tr>
<tr>
<td>Generalized positive feedback</td>
<td>3.24</td>
<td>0.79</td>
</tr>
</tbody>
</table>

*Note. The range of all scales is from 1 to 5.*

The average of students’ personal theories in each class also seemed to change to differing degrees in the different classes (see Figure 1), supporting the claim that personal theories change during college based on varied classroom contexts. A repeated-measures MANOVA using time of measurement (beginning vs. end) and class revealed a significant interaction between time and class at the multivariate level ($F(13, 84) = 1.299, p = .044, \eta^2 = .12$) despite the small average number of subjects per class, suggesting significant heterogeneity in personal theory changes by class.
Multilevel modeling with disaggregated between- and within-class effects (Curran & Bauer, 2011) was utilized to examine whether messages related to changes in students’ personal theories during the summer term. Due to the relatively small number of level 2 groups and the small conditional ICCs (0-0.11) for students’ final personal theories (controlling for initial personal theories), a procedure with repeated-measures mixed models using compound symmetry assumptions was followed (Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). Three different metrics of messages were used as predictors: between-class differences in student-reported messages (using the class average of each
student-reported message), within-class differences in student-reported messages (using the difference between an individual student’s report and the class average), and instructor-reported messages, which were examined in separate models due to the strong relation between and student and instructor reports and the small number of level 2 groups. These three variables helps to distinguish whether the relation between messages are attributable to differences in messages from the instructors’ perspectives, differences between overall student perceptions in each class, or individual differences in message perceptions within each class. All predictor variables were mean-centered. A different set of models was conducted for each of the six student beliefs that were being measured. An example of this model can be seen below for predicting changes in incremental theories of intelligence:

\[
\text{Level 1: } \text{INCR}_{T2,ij} = \beta_{0j} + \beta_{1j}\text{HELP}_{ij} + \beta_{2j}CÑC_{ij} + \beta_{3j}\text{INCR}_{T1,ij} + r_i
\]

\[
\text{Level 2: } \begin{align*}
\beta_{0j} &= \gamma_{00} + \gamma_{01}\text{HELP}_j + \gamma_{02}\text{CONC}_j + u_{0j} \\
\beta_{1j} &= \gamma_{10} \\
\beta_{2j} &= \gamma_{20} \\
\beta_{3j} &= \gamma_{30}
\end{align*}
\]

In this model, HELP\(_{ij}\) represents group-mean centered student reports of instructor help, or the reports of instructor help for student \(i\) compared to the average class level of class \(j\). The class mean for student-reported instructor help in class \(j\) is represented by HELP\(_j\). Controlling for initial beliefs is accounted for by the term that includes INCR\(_{T1,ij}\).
Table 3 shows the messages that were hypothesized to relate to changes in particular personal theories. Messages that were not related to personal theories were removed from the final models (see Table 5 for results from the final models and Table 7 for a summary of the results). Marginal results are also reported due to the study’s exploratory nature.

For every personal theory, initial personal theories were strong and significant predictors of final personal theories (see Table 5). However, even when accounting for initial personal theories, student-reported messages explained significant variance in final personal theories.
<table>
<thead>
<tr>
<th>T2PT</th>
<th>Incremental Theories of Intelligence</th>
<th>Quick Learning</th>
<th>Simple/ Certain Knowledge</th>
<th>Justification by Authority</th>
<th>Personal Justification</th>
<th>Perceived Competence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\Gamma$</td>
<td>SE</td>
<td>$\gamma$</td>
<td>SE</td>
<td>$\gamma$</td>
<td>SE</td>
</tr>
<tr>
<td>T1PT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.84* 0.03</td>
<td></td>
<td>0.83* 0.04</td>
<td></td>
<td>0.88* 0.03</td>
<td></td>
</tr>
<tr>
<td>HELP</td>
<td>B -0.29* 0.12</td>
<td>-0.01 0.09</td>
<td>X</td>
<td>-0.03 0.11</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W 0.05 0.05</td>
<td>-0.08* 0.03</td>
<td>X</td>
<td>-0.09* 0.04</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>DIFF</td>
<td>B X X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W X X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONC</td>
<td>B -0.13* 0.06</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W 0.01 0.03</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNCE</td>
<td>B -0.08* 0.03</td>
<td>X</td>
<td>0.10* 0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W -0.02 0.03</td>
<td>X</td>
<td>0.06* 0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEGF</td>
<td>B 0.03 0.07</td>
<td></td>
<td></td>
<td>0.12 0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W 0.05* 0.03</td>
<td></td>
<td></td>
<td>-0.07* 0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSF</td>
<td>B X</td>
<td></td>
<td></td>
<td>0.10* 0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W X</td>
<td></td>
<td></td>
<td>0.05 0.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Note. Each column represents a separate multilevel model, with the final personal theory (T2PT) listed at the top as the dependent variable. T1PT represents the corresponding beginning of the semester personal theory for each model. Blank cells were not included in the models and cells marked with an “X” were originally included in the model but were removed because neither the between- nor within-class results were marginally significant. HELP = Instructor help. DIFF = Differential ability. CONC = Conciliatory messages. UNCE = Uncertainty in the field. NEGF = Generalized negative feedback. POSF = Generalized positive feedback.

* $p < .05$, † $p < .10$
Differences in student-reported frequency of instructor help messages and conciliatory messages between classes both negatively predicted incremental theories at the end of the course, as hypothesized. Differences in help messages within classes negatively predicted quick learning theories, as hypothesized, while within-class differences in negative feedback marginally positively predicted quick learning, contrary to hypotheses. Differences in uncertainty in the field messages between classes negatively predicted simple/certain knowledge beliefs, as hypothesized. None of the messages included in the model predicted justification by authority. Differences in instructor help within classes negatively predicted personal justification, while differences in uncertainty in the field between classes, and to some extent within classes, positively predicted personal justification, as hypothesized. Generalized feedback predicted perceived competence in the hypothesized directions, though the effect was from within class differences for negative feedback and between-class differences for positive feedback. Because both between-class and within-class differences predicted final personal theories while controlling for initial personal theories, this suggests that some messages have the same effect on all students’ personal theories, but that the students varying remembered experiences might produce changes in beliefs for students in the same class as well. In other words, the messages that individual students perceive (regardless of what the class perceives as a whole) act as subjective evidence for
adjusting students’ personal theories. This further supports the assertion that not all messages an instructor says reach all students in the same way.

Some similar and some unique results were found when using instructor-reported messages as level 2 variables in six separate models (Table 6). In line with the student-report results and the hypotheses, instructor help messages negatively predicted students’ incremental theories of intelligence at the end of the semester, uncertainty in the field messages negatively predicted students’ belief in simple/certain knowledge, and instructor help negatively predicted and uncertainty in the field marginally positively predicted personal justification. In addition to the findings that were in line with the student-report results, instructor-reported differential ability messages negatively predicted justification by authority, counter to the hypotheses, and conciliatory messages marginally positively predicted personal justification, as hypothesized.
Table 6: Full Results of Instructor Message Models in Study 1

<table>
<thead>
<tr>
<th>T2PT: Incremental Theories of Intelligence</th>
<th>Quick Learning</th>
<th>Simple/ Certain Knowledge</th>
<th>Justification by Authority</th>
<th>Personal Justification</th>
<th>Perceived Competence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma$</td>
<td>SE</td>
<td>$\gamma$</td>
<td>SE</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>T1PT</td>
<td>0.84*</td>
<td>0.03</td>
<td>0.85*</td>
<td>0.04</td>
<td>0.89*</td>
</tr>
<tr>
<td>HELP</td>
<td>-0.22*</td>
<td>0.09</td>
<td>-0.05</td>
<td>0.12</td>
<td>0.00</td>
</tr>
<tr>
<td>DIFF</td>
<td>-0.03</td>
<td>0.03</td>
<td>-0.02</td>
<td>0.03</td>
<td>-0.04*</td>
</tr>
<tr>
<td>CONC</td>
<td>-0.01</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.07*</td>
</tr>
<tr>
<td>UNCE</td>
<td></td>
<td></td>
<td>-0.06*</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>NEGF</td>
<td>-0.05</td>
<td>0.11</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.05†</td>
</tr>
<tr>
<td>POSF</td>
<td>0.06</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Each column represents a separate multilevel model, with the final personal theory (T2PT) listed at the top as the dependent variable. T1PT represents the corresponding beginning of the semester personal theory for each model. HELP = Instructor help. DIFF = Differential ability. CONC = Conciliatory messages. UNCE = Uncertainty in the field. NEGF = Generalized negative feedback. POSF = Generalized positive feedback. 

* $p < .05$, † $p < .10$
To further investigate within-class differences in students’ reported messages, I conducted additional multi-level models using messages as the dependent variables and students’ initial beliefs (focusing on within-class differences in initial beliefs) as predictors. In multiple cases, students’ initial beliefs predicted how often students’ remembered hearing instructor messages (see Table 7 for the complete model findings). Students were less likely to report hearing instructor help messages if they had initial beliefs in quick learning and but marginally more likely to report these messages if they had high levels in justification by authority. Students were more likely to report hearing conciliatory messages if they began the semester with strong beliefs in personal justification of knowledge. Students were less likely to report generalized negative feedback if they had high beliefs in justification by authority. Students were also marginally more likely to report generalized positive feedback if they began the semester with high personal justification beliefs. Therefore, some of the variability in the messages students reported hearing in class (which one might assume to be homogenous within the same classroom) could be predicted by their initial personal theories.
Table 7: Prediction of Student-reported Messages Using Initial Personal Theories

<table>
<thead>
<tr>
<th></th>
<th>HELP</th>
<th>DIFF</th>
<th>CONC</th>
<th>UNCE</th>
<th>NEGF</th>
<th>POSF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\gamma)</td>
<td>SE</td>
<td>(\gamma)</td>
<td>SE</td>
<td>(\gamma)</td>
<td>SE</td>
</tr>
<tr>
<td>Incremental Theories</td>
<td>-0.03</td>
<td>0.05</td>
<td>0.00</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.07</td>
</tr>
<tr>
<td>Quick Learning</td>
<td>-0.34*</td>
<td>0.10</td>
<td>-0.22</td>
<td>0.15</td>
<td>-0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Simple/Certain Knowledge(^i)</td>
<td>-0.17</td>
<td>0.12</td>
<td>0.15</td>
<td>0.18</td>
<td>0.28</td>
<td>0.17</td>
</tr>
<tr>
<td>Justification by Authority(^i)</td>
<td>0.18\†</td>
<td>0.10</td>
<td>-0.05</td>
<td>0.15</td>
<td>-0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>Personal Justification(^i)</td>
<td>0.01</td>
<td>0.08</td>
<td>0.14</td>
<td>0.12</td>
<td>0.33*</td>
<td>0.11</td>
</tr>
<tr>
<td>Perceived Competence(^i)</td>
<td>-0.04</td>
<td>0.08</td>
<td>-0.03</td>
<td>0.11</td>
<td>-0.08</td>
<td>0.11</td>
</tr>
</tbody>
</table>

*Note. All personal theories were mean-centered, with the domain-specific personal theories (\(^i\)) measured as within-class variance in beliefs to remove variability in messages that might vary by field of study.*

\* \(p < .05\), \† \(p < .10\)
3.3 Discussion of Study 1

Study 1 provides evidence that instructors’ send educational messages in class and these messages are related to changes in students’ personal theories. This supports the proposed college-student theory theory (Barger & Linnenbrink-Garcia, under review) in which college students, like younger students, continue to use messages in class as evidence to adjust their personal theories about education, even over the course of a single, six-week class. Different messages contain different kinds of information for students. Instructor help messages provided evidence that intelligence is fixed, that learning takes time, that knowledge is not personally justified, and that students are capable of succeeding. Differential ability messages provided evidence that authorities are not the only source of knowledge. Conciliatory messages provided evidence that intelligence is fixed, but also that knowledge can be personally justified. Messages of uncertainty in the field provided evidence that knowledge is not a collection of unchanging facts. However, the evidence in some of these messages does not seem to reach students the same way, as within-class differences in remembered messages also related to final personal theories and initial personal theories predict what messages students are likely to hear.

Despite the promising results from Study 1, several questions remain. Specifically, would the relations between messages and personal theories changes, particularly the domain-specific theories about knowledge, still be observed if the
classes had all been in the same domain? It remains unclear whether changes in domain-specific beliefs can be directly attributed to the messages or confounding factors, such as field of study, of which there were several in Study 1. While controlling for initial beliefs and analyzing within-class differences in messages alleviates some of this concern, it is still possible that differences in messages between classes are actually a marker for the way different fields’ portray knowledge, instead of a mechanism for belief change. Therefore, Study 2 will need to test and compare models within single-fields of study, which was not possible due to the small sample sizes within any one domain in Study 1. There was also a great deal of within-class variability in students’ remembered messages (ICCs ranging from .04-.48). This variability was at least partially explained by students’ initial beliefs, meaning that the process of personal theory development is not as simple as evidence being provided to and processed by all students the same way; initial beliefs may serve as a lens through which students experience and remember the messages from the instructor. This complex process might be better studied using structural equation modeling, which was not possible in Study 1 due to the small number of students in each class. Study 2 will be able to do this in two larger, more homogenous samples and examine all of these relations simultaneously.

Study 1 also raises additional questions that would be better investigated in Study 2. First, messages such as instructor help and conciliatory messages may appear positive at first glance, but based on the data, might have more pernicious effects. It
might be possible to send messages with similar intentions (i.e., letting students know that they have the instructors' support and that there will be future opportunities for achievement in class) without providing underlying evidence to students that their intelligence is fixed or that they cannot play any role in the formation of knowledge. Second, many of the messages that seem to matter are involved with achievement (conciliatory messages, generalized positive and negative feedback). Students' individual achievement might then play a moderating role in the messages students remember and subsequently their changes in beliefs. For example, students who perform well on the exam probably interpret conciliatory messages and generalized feedback about the whole class differently from students who perform poorly.

In order to address these issues and questions, I conducted Study 2 in two larger lecture classes. This allowed a richer exploration of the interplay between students' initial personal theories, the within-class differences in perceived messages, and students' personal theories at the end of the class using structural equation modeling.
4. Study 2

4.1 Method

4.1.1 Power Analysis

Because structural equation modeling was used in Study 2, a minimum of 200 total participants was preferred to assure accuracy of estimators (Hoyle & Gottfredson, 2015). Using estimates of variances and path coefficients based on the data from Study 1, a Monte Carlo simulation study was conducted for a sample of 220 anticipated participants to determine the power of testing the path model proposed for Study 2. Results suggested that the average fit of the proposed path model would be excellent (RMSEA = .015, SD = .017; SRMR = .024, SD = .003) and there would be sufficient power (1 - β > .80) to find multiple significant path coefficients that were in line with the within-class results from Study 1. It is likely that this simulation gave fairly conservative estimates, given the additional error in Study 1 from the numerous courses and domains, which was substantially reduced by the more homogeneous design of Study 2.

4.1.2 Participants

Participants were undergraduate students in two introductory-level classes, honors introductory chemistry (n = 112) and introductory neuroscience (n = 127). The chemistry class was exclusively first-year students, many of whom typically have taken AP chemistry in high school. The neuroscience class enrolls older students (3.3% first-year students, 72.7% second-year, 15.7% third-year, 8.3% fourth-year) and is required for
the major. The sample was mostly female (53.3% female, 38.2% male, 8.5% did not report) and demographically similar to the university as a whole in race/ethnicity (39.4% white (non-Hispanic), 32.1% Asian, 9.3% black, 6.1% Hispanic, 3.6% multiracial/other, 9.3% did not report). Among the students who reported additional demographics at the final assessment, 66% had attended a public high school and 30.4% attended a private high school. Only 8.9% of these students considered themselves to be a first generation college student and the median responses on parental education was between a bachelors and a master’s degree.

The retention rate for the final survey was comparable to Study 1, with 76.1% of students who completed the first survey also completing the second survey. The retention rate for the post-exam survey was somewhat lower (60.3%), likely due to the fact that students had to complete the survey outside of class time.

4.1.3 Procedure

The procedure was structurally identical to Study 1 stretched from six-weeks to a full fifteen-week semester. During the first two weeks of the course, the instructor of the class introduced the researcher, who then directed students who wanted to participate in the online survey to complete it during class. The first survey (T1) assessed students’ initial personal theories and demographic information.

Through coordination with the instructor, observational data were collected on the day students received their first exam grades and instructors discussed the exams
with students. The chemistry class was audio recorded while the neuroscience class was video recorded by Tableau. These events were within a week of one another in the two classes. On this day, students were emailed a brief online survey (T2) to be completed outside of class; a majority of students (n = 144) completed the survey.

During the final two weeks of the semester, researchers returned to the classroom to direct students to a third online survey to be completed during class time (T3). This survey assessed students’ final beliefs, the messages they remember hearing from the principle instructor, what other classes they took during the semester, and an estimate of their attendance in class. This survey asked students an open-ended question about what messages they remembered the instructor saying in class about the exam, what grade they received, and their incremental theories of intelligence, perceived competence, and effort beliefs.

4.1.4 Measures

The measures for Study 2 were identical to the measures for Study 1, with the additions described below. For a complete list of items, see Appendix A.

4.1.4.1 Personal theories

General personal theories (theories of intelligence, speed of learning) and domain-specific beliefs (perceived competence, simple/certain knowledge, justification by authority, personal justification) in chemistry and neuroscience were measured at the beginning and end of the semester.
In addition to these beliefs about education, beliefs about the importance of effort in achievement (8 items, e.g., “The harder you work at something, the better you will be at it,” Blackwell, 2002) were also included.

4.1.4.2 Messages

Six message subscales from Study 1 were administered at the end of the semester (instructor help, differential ability, generalized negative feedback, generalized positive feedback, conciliatory messages, and uncertainty in the field, see Appendix A). In addition to the six message factors, Study 2 attempted to determine if there are ways to send instructor help and conciliatory messages without a negative impact on students’ beliefs, as was found in the first study. Two additional subscales were included in the final surveys that delineate instructor help (e.g., “Ask me for help if you don’t understand.”) from instructor support (e.g., “My door is always open.”) and conciliatory messages (e.g., “It’s OK if you had a bad test.”) from spurring messages (e.g., “Struggling on the test just means it’s not clicking yet.”), or giving reasoned conciliatory messages, a distinction that was suggested by exploratory factor analyses below.

4.1.4.3 Demographics

Students were asked to report their year in school, gender, and race/ethnicity on the first survey. Students were asked to report additional demographics on the final survey, including educational history, class enrollment, and parental education.
4.1.5 Statistical analysis plans

The relations between personal theories and messages in class were investigated using structural equation modeling. The multi-group path model depicted in Figure 2 was tested for the chemistry and neuroscience classes. Given the complexity of the model and the difficulties in measuring certain personal theories (DeBacker et al., 2008), a path model was determined to be more appropriate than a model with latent variables. A measurement model was tested separately to examine measurement equivalence across the two classes and subjects.

![Figure 2: The path model tested in Study 2.](image)

*Note.* As shown here, the model will include autoregressive paths from initial beliefs to final beliefs, paths from messages to final beliefs, and paths from initial beliefs to messages.
Because the two classes and subjects were substantially different, as were the demographics of students in the classes, it was prudent to begin the model selection process by assuming that the variables in the two classes would have different mean structures (as demonstrated by the descriptive statistics shown in Table 4) and path coefficients. Starting with the assumption that all parameters were free to vary across the two groups in the model (Model A), I then systematically tested whether imposing equality constraints on components of the model would substantially degrade the model fit. In lieu of testing the 78 path coefficients individually, which would lead to a virtually infinite number of combinations of fixed and free parameters across the groups, I followed a stepwise approach constraining groups of related paths. I first tested imposing equality constraints on the autoregressive paths (Model B), then constrained groups of paths related to each T3 personal theory (Models C1-C6), and then constrained groups of paths related to each of the T1 personal theories (Models E1-E6). When constraining a set of paths to equality across classes did not degrade the model fit significantly using a chi-square difference test, this set of constraints was retained to produce the final model (Model F). CFI, RMSEA, and SRMR were used to evaluate the overall fit of the final model, following recommendations from multiple researchers (Bentler, 1990; West, Taylor, & Wu, 2013) that a good-fitting model should have a CFI greater than .95, an RMSEA below .08 and ideally below .06, and SRMR less than .08.
Table 8: Descriptive Statistics for Personal Theories and Student-Reported Messages by Class

<table>
<thead>
<tr>
<th></th>
<th>Chemistry</th>
<th></th>
<th></th>
<th>Neuroscience</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>α</td>
<td>M</td>
<td>SD</td>
<td>α</td>
</tr>
<tr>
<td>T1 Incremental Intelligence</td>
<td>3.91</td>
<td>1.03</td>
<td>.93</td>
<td>3.64</td>
<td>0.93</td>
<td>.92</td>
</tr>
<tr>
<td>T1 Quick Learning</td>
<td>1.77</td>
<td>0.45</td>
<td>.62</td>
<td>1.87</td>
<td>0.50</td>
<td>.68</td>
</tr>
<tr>
<td>T1 Simple/Certain Knowledge</td>
<td>2.54</td>
<td>0.63</td>
<td>.83</td>
<td>1.85</td>
<td>0.50</td>
<td>.78</td>
</tr>
<tr>
<td>T1 Justification by Authority</td>
<td>3.43</td>
<td>0.68</td>
<td>.87</td>
<td>3.17</td>
<td>0.69</td>
<td>.83</td>
</tr>
<tr>
<td>T1 Personal Justification</td>
<td>2.52</td>
<td>0.76</td>
<td>.73</td>
<td>3.29</td>
<td>0.64</td>
<td>.44</td>
</tr>
<tr>
<td>T1 Perceived Competence</td>
<td>4.13</td>
<td>0.72</td>
<td>.91</td>
<td>4.02</td>
<td>0.71</td>
<td>.87</td>
</tr>
<tr>
<td>T1 Effort Beliefs</td>
<td>4.80</td>
<td>0.58</td>
<td>.68</td>
<td>4.58</td>
<td>0.51</td>
<td>.53</td>
</tr>
<tr>
<td>Instructor Help</td>
<td>3.95</td>
<td>0.77</td>
<td>.81</td>
<td>4.27</td>
<td>0.71</td>
<td>.74</td>
</tr>
<tr>
<td>Conciliatory Messages</td>
<td>3.99</td>
<td>0.67</td>
<td>.72</td>
<td>3.14</td>
<td>0.98</td>
<td>.83</td>
</tr>
<tr>
<td>Differential Ability</td>
<td>3.27</td>
<td>0.80</td>
<td>.69</td>
<td>3.08</td>
<td>0.89</td>
<td>.80</td>
</tr>
<tr>
<td>Uncertainty in the Field</td>
<td>2.16</td>
<td>0.88</td>
<td>.88</td>
<td>4.25</td>
<td>0.58</td>
<td>.72</td>
</tr>
<tr>
<td>Generalized Negative Feedback</td>
<td>1.65</td>
<td>0.67</td>
<td>.85</td>
<td>1.46</td>
<td>0.65</td>
<td>.78</td>
</tr>
<tr>
<td>Generalized Positive Feedback</td>
<td>2.61</td>
<td>0.81</td>
<td>.66</td>
<td>3.77</td>
<td>0.92</td>
<td>.82</td>
</tr>
<tr>
<td>Instructor Support</td>
<td>3.40</td>
<td>0.82</td>
<td>.78</td>
<td>3.47</td>
<td>0.91</td>
<td>.83</td>
</tr>
<tr>
<td>Reasoned Conciliatory Message</td>
<td>2.80</td>
<td>0.88</td>
<td>.64</td>
<td>2.65</td>
<td>0.85</td>
<td>.66</td>
</tr>
<tr>
<td>T3 Incremental Intelligence</td>
<td>3.99</td>
<td>1.01</td>
<td>.93</td>
<td>3.84</td>
<td>1.00</td>
<td>.92</td>
</tr>
<tr>
<td>T3 Quick Learning</td>
<td>1.97</td>
<td>0.62</td>
<td>.83</td>
<td>1.97</td>
<td>0.54</td>
<td>.69</td>
</tr>
<tr>
<td>T3 Simple/Certain Knowledge</td>
<td>2.52</td>
<td>0.66</td>
<td>.86</td>
<td>1.83</td>
<td>0.53</td>
<td>.77</td>
</tr>
<tr>
<td>T3 Justification by Authority</td>
<td>3.29</td>
<td>0.65</td>
<td>.83</td>
<td>3.05</td>
<td>0.68</td>
<td>.83</td>
</tr>
<tr>
<td>T3 Personal Justification</td>
<td>2.68</td>
<td>0.67</td>
<td>.56</td>
<td>3.12</td>
<td>0.71</td>
<td>.60</td>
</tr>
<tr>
<td>T3 Perceived Competence</td>
<td>3.85</td>
<td>0.87</td>
<td>.92</td>
<td>3.85</td>
<td>0.80</td>
<td>.92</td>
</tr>
<tr>
<td>T3 Effort Beliefs</td>
<td>4.54</td>
<td>0.64</td>
<td>.72</td>
<td>4.64</td>
<td>0.59</td>
<td>.67</td>
</tr>
</tbody>
</table>

A supplemental model was also tested to distinguish between the effects of instructor help/support and conciliatory/spurring messages on a subset of personal theories (See Figure 3). Exploratory factor analyses were first conducted to verify that
the new scales (instructor support and spurring messages) were empirically distinguishable from the previously tested scales. The same model building procedure was followed, starting with the assumption that all paths were free to vary between classes and then systematically testing whether constraining path coefficients to be equal across classes degraded the model fit.

Figure 3: Supplemental model for Study 2

Given the burgeoning research on epistemic beliefs as a system of cognitions as opposed to orthogonal beliefs (Barger et al., under review; Chen & Barger, 2016; Chinn et al., 2011; Schommer-Aikins, 2004), I also conducted person-centered analyses (Magnusson, 1998) to explore whether the messages that students-reported hearing in class were associated with shifts in students’ holistic systems of personal theories about knowledge. Two person-centered analyses were considered for investigating how
students shift in their beliefs over time: latent transition analysis (LTA; Lanza, Flaherty, & Collins, 2003) and i-states as objects cluster analysis (ISOA; Bergman & El-Khoury, 1999). Latent transition analysis is a mixed modeling procedure that identifies latent profiles across multiple time points that underlie variatiability in the observed variables (in this case, personal theories about knowledge). It then simultaneously models the likelihood of shifts between those groups at a latent level, and covariates can be included in the same model to predict these shifts. One clear benefit is the ability to answer the current questions within a single statistical model. Unfortunately, given the sample size, it was unlikely to find a stable LTA solution with more than two classes (early analyses supported this). Previous research suggests that far more than two groups are likely necessary to characterize students’ systems of personal theories about knowledge (Barger et al., 2014; Buehl & Alexander, 2005; Chen, 2012; Greene et al., 2008, 2010). Therefore, LTA was not used for the current study, but might prove useful in future research with larger samples.

ISOA, on the other hand, groups individuals according to the similarity in their observed variables, categorizing individuals not on a latent level but based on their observed values. Unlike in structural equation modeling, where fit statistics can be used to identify the best fitting number of clusters, a number of criteria are compared to triangulate the best cluster solution, including the agglomeration schedule and dendrogram, cluster size, the proportion of variance in clustering variables explained by
the cluster solution, and theoretical considerations. Previous research using latent profile analysis and ISOA suggest a number of theoretically likely clusters that develop in a predictable pattern. This includes realists (high endorsement of all epistemic beliefs), who then progress to dogmatism (lower endorsement of simple/certain knowledge, high justification by authority, low personal justification) or skepticism (low endorsement of simple/certain knowledge, low justification by authority, high personal justification), and finally rationalists (low simple/certain knowledge, balanced justification by authority and personal justification beliefs). One study also found a large cluster of students who were near the average on all scales, which was termed uncommitted to represent the idea that these students did not have a particularly strong stance towards knowledge (Barger et al., under review).

Because ISOA only identifies clusters and cluster membership, separate analyses were necessary to investigate whether reported messages correspond with changes in students’ systems of personal theories about knowledge. Individuals would be separated first by their initial membership. Then, within each initial cluster, ANOVAs were conducted comparing students who ended up in different clusters on uncertainty in the field and instructor help messages, those most closely aligned with personal theories about knowledge. This would help point to whether individuals who begin with similar systems of beliefs but end up with different systems differ in the kinds of messages they report hearing. This would support the idea that messages are associated
with changing personal theories. Given the predictable progression of development and the stable tendency of epistemic beliefs over short periods of time (Chandler, Hallett, & Sokol, 2002), it was expected that there would be very few students in some of the cells. To deal with the small cell sizes statistically, cells with fewer than five cells were excluded from these exploratory analyses.

For post-exam data, students reported what they could recall their instructor saying about their exam. Two independent coders examined the students’ open-ended responses to what the instructor said in class about the exam for four categories that correspond with the Likert-scale message items: generalized positive feedback, generalized negative feedback, conciliatory messages, and spurring messages. Students who reported a particular message were scored as 1 and students who did not report the message were coded as 0. A series of hierarchical regression analyses were utilized to examine whether student-reported messages assessed more proximally to a relevant in-class event (receiving their first exam scores) were predictive of changes in personal theories. Interaction terms between students’ self-reported exam grades and the binary message codes were also included to determine whether the relation between messages and personal theories varied by exam performance. Grades were standardized in each classroom and centered in order to calculate four separate interaction terms. Separate regressions were conducted for the following criterion variables: T2 and T3 incremental theories of intelligence, T2 and T3 perceived competence, T2 and T3 effort beliefs, T3
quick learning, and T3 personal justification beliefs. The first step of the regressions included the initial belief and class membership (chemistry = 0, neuroscience = 1). The second step included the standardized exam score. The third step added the four binary message codes. The fourth and final step added the four interaction terms between message and mean-centered exam performance.

4.2 Results

The primary objective of Study 2 was to examine the interplay between students’ beliefs and the within-class variability in student-reported messages through structural equation modeling. The results below begin with the procedure that was used to reach the final model. I then describe the supplemental model that distinguished between instructor help/support and conciliatory/spurring messages. Following the primary analyses, ancillary analyses focused on the person-centered analyses of students’ personal theories about knowledge, the coding of responses on the post-exam survey, and the analyses using the coding of the open-ended responses.

4.2.1 Final model results

A multi-group path analysis was conducted with students in both classes (see Figure 3). It was critical to test differences between the two classrooms. I began with the assumption that all parameters were free to vary across the two groups in the model (Model A) and then systematically tested whether imposing equality constraints on components of the model would substantially degrade the model fit using change in chi-
square tests. Constraining autoregressive paths (Model B) did not significantly degrade the model fit. Therefore all autoregressive paths were fixed in the subsequent model. While testing constraining groups of paths related to each T3 personal theory (Models C1-C6), the paths associated with T3 personal justification and T3 justification by authority were the only ones that degraded model fit when fixed across the two classes. Therefore, other paths from messages to T3 personal theories were constrained across the two classes in subsequent models (Model D). The process of testing constraining groups of paths related to each of the T1 personal theories (Models E1-E6) revealed that only paths radiating from T1 personal justification differed between classes, so that all other paths were constrained across the classes in the final model (Model F). For the model fit statistics for each step of the model-building process, see Appendix B.

The final model constrained most paths to be equal across groups (neuroscience, chemistry), with the exception of the groups of paths that connected T1 personal justification to reported messages, reported messages to T3 personal justification, and reported messages to T3 justification by authority. The final model fit the data well ($\chi^2(120) = 154.981$, CFI = .959, RMSEA = .057 (95% CI: [.025, 0.081]), SRMR = .069) and is represented with all significant pathways displayed in Figure 4 for chemistry and Figure 5 for neuroscience. Non-significant pathways and correlated errors between scales measured at the same time are not shown for clarity.
There was significant stability in students’ personal theories, as the autoregressive paths ranged from .34 to .71. Incremental theories proved to be the most

Figure 4: Full model of personal theories and messages for chemistry classroom.

Figure 5: Full model of personal theories and messages for neuroscience classroom.
stable, followed by the justification domains, perceived competence, and then quick learning and simple/certain knowledge. The strong correspondence between initial and final theories suggests that the incremental theories of intelligence might be operating more as traits than developing theories.

Even when controlling for initial personal theories, messages were significant predictors of final personal theories. This is displayed in Figures 4 and 5 as the arrows drawn from the message scales in the middle of the figure to the T3 personal theories. Specifically, conciliatory messages positively predicted incremental theories, negative feedback positively predicted quick learning and simple/certain knowledge, and uncertainty in the field messages negatively predicted simple/certain knowledge. In the chemistry class, justification by authority was predicted by positive feedback while personal justification was predicted by uncertainty in the field messages. In the neuroscience class, justification by authority was predicted positively by instructor help and negative feedback and negatively by differential ability.

Furthermore, students’ initial personal theories predicted variability in students’ reported messages. This can be seen in Figures 4 and 5 with the arrows from the T1 personal theories to the message scales in the middle of the figure. Specifically, students who initially believed that learning must occur quickly reported more frequent generalized negative feedback. Students with higher perceived competence reported more frequent conciliatory, instructor help, and uncertainty in the field messages.
Finally, personal justification predicted more frequent uncertainty in the field messages in the neuroscience class.

There were no significant indirect effects of initial personal theories on final personal theories through any of the messages.

**4.2.2 Model separating unique messages from instructor help and conciliatory messages**

Given the results from Study 1 that suggested potentially harmful effects of the well-intentioned instructor help and conciliatory messages, a second objective of Study 2 was to try to distinguish the underlying positive components of these messages from the implicit negative message. The positive intention of the instructor help message is that the instructor is available for support. A positive intention behind the conciliatory message might be to remind students that they can do better on future opportunities, thus spurring the student on for future coursework.

First, exploratory factor analyses were conducted to determine whether such messages were distinguishable from one another in the self-report measures. An exploratory factor analysis of instructor help and instructor support message items revealed that there were generally two factors that separated new support items from the original instructor help items. Two new support message items that cross-loaded on both factors in initial EFAs were removed from further analyses. The biggest difference between the two factors was that the instructor help items all had the implication that students would struggle in some way and then need to seek help from the instructor.
(e.g., “You should ask me questions if you get confused”), as opposed to the support items that more ambiguously informed the students that support is available in the course (e.g., “My door is always open.”). The new instructor support subscale of five items demonstrated reasonable reliability ($\alpha = .79$). The two scales were moderately correlated ($r = .56, p < .001$).

The conciliatory and spurring messages proved more difficult to disentangle empirically. An initial exploratory factor analysis did not find a clear distinction between the original conciliatory message items and the new spurring message items. Instead, two factors seemed to form around items that mentioned “exams” and “answers” or “assignments.” Nevertheless, by removing one of the original conciliatory message items and keeping three new items that focused on specific reasons for why a student might have performed poorly, two clear factors emerged. Therefore, a new scale for conciliatory exam messages ($\alpha = .87$) and reasoned conciliatory messages ($\alpha = .66; r = .41, p < .001$) were computed for use in the subsequent path analyses. The term “reasoned conciliatory messages” replaces “spurring messages” for the remainder of these analyses.

The next step was to test whether the positive effects of instructor support and reasoned conciliatory messages can be distinguished from the sometimes detrimental effects of instructor help and conciliatory messages found in Study 1. The model depicted in Figure 3 was used to examine this question. In testing this model, I used the
same procedure as the full model (see previous section). As before, equality constraints
on the paths involving personal justification were determined to significantly degrade
the model fit (see Appendix C), as were the paths involving the quick learning belief,
and so these paths were allowed to vary across classes. Marginally significant
degradations in model fit were not originally excluded from subsequent models (Model
F'), but including these changes together did significantly improve the model fit, and so
the final model removed these marginally degrading equality constraints as well (Model
F). The final model fit the data well ($\chi^2(44) = 67.559$, CFI = .962, RMSEA = .077 (95% CI:
[.036, 0.112]), SRMR = .067). The significant standardized paths for both classes can be
seen in Figure 6. Again, pathways that did not significantly differ from zero and
correlated errors between scales in each column are not shown for clarity.

\[ \]

Figure 6: Supplemental model distilling messages.

Note. The first path coefficient on each line is for the chemistry classroom and the
second is for the neuroscience classroom. Significant paths are in bold while
marginally significant effects marked with †.
From this model, it became clear that even though the two sets of similar messages are related, their relation to student’s initial and final beliefs were distinguishable. In both classes, instructor support message predicted higher perceived competence, but instructor help did not. In the neuroscience class, instructor help had the expected negative relation to personal justification, while instructor support related positively to personal justification. In the chemistry class, conciliatory exam messages negatively related to quick learning beliefs while instructor help messages positively related to quick learning beliefs, leaving the new message categories unrelated. Initial beliefs also predicted different messages in different patterns. Specifically, incremental theories of intelligence predicted more frequent reports of reasoned conciliatory messages and higher perceived competence was associated with more frequent reports of instructor help messages in both classes. Personal justification beliefs predicted less frequent instructor help in the chemistry class but more frequent instructor support in the neuroscience class. Quick learning beliefs predicted less frequent instructor support messages in the chemistry class and more frequent reasoned conciliatory messages in the neuroscience class. Therefore, these similar sets of messages seem to provide different information about the underlying nature of education in these classes.
4.2.3 Person-centered analyses of epistemic beliefs

All survey responses at T1 and T3 with valid simple/certain knowledge, justification by authority, and personal justification score were included in the ISOA \((n = 415)\). No outliers were found on any variable. I examined cluster solutions with two to six clusters. The agglomeration schedule suggested that four or five clusters were the most appropriate cluster solutions. The five cluster solution was the smallest number of clusters that accounted for at least 50% of the variance in each of the clustering variables. The five-cluster solution divided a cluster from the four-cluster solution that did not have a clear, theoretical position into hypothesized rationalist and skeptic clusters. A six-cluster solution divided dogmatic students from the five-cluster solution into more extreme and moderate dogmatic individuals, resulted in a small cluster with 16 i-states (approximately 8 individuals across two time points). This was not considered to be theoretically important enough to justify such a small cluster. Using all of these methods, the five-cluster solution was ultimately selected.

4.2.3.1 Description of clusters

There were five clusters that roughly corresponded with either theoretical epistemic positions (Chandler et al., 2002; Greene et al., 2008) or prior empirical work on profiles of epistemic beliefs (Barger et al., under review; Greene et al., 2010). The profiles can be seen in Figure 7.
The sample size described here for each cluster represents the number of i-states, such that 2 i-states represent roughly one individual in the same profile at the beginning and end of the semester. The first cluster represented rationalists \((n = 70)\), with very low agreement with simple/certain knowledge, and high-moderate, but balanced, patterns of justification by authority and personal justification. The second cluster was labeled skeptics \((n = 51)\) due to their equally low agreement with simple/certain knowledge, rejection of justification by authority, and strong agreement with personal justification. The third cluster \((n = 141)\) was near the average on all scales, mirroring a large cluster labeled *uncommitted* that has also been found in previous ISOA analyses using the same clustering variables (Barger et al., under review). The fourth cluster represented
dogmatists \((n = 67)\), with moderate simple/certain knowledge (similar to the dogmatist profile in the Barger et al. study, though slightly different from the hypothesized pattern for dogmatists in Greene et al., 2008), high justification by authority, and low personal justification beliefs. The final cluster \((n = 86)\) represented a mix between the hypothesized profiles for realism (high-all beliefs) and dogmatism, and is therefore labeled \textit{realist/dogmatist}. Students in the realist/dogmatist cluster were similar to students in the dogmatist cluster except in having moderate agreement with personal justification as well.

A series of chi-square tests of independence investigated whether certain clusters were more likely at the beginning or end of the semester and whether certain clusters were more likely in the chemistry and neuroscience domain. Profile membership did not appear to depend on time point as there were not significantly more students in any cluster at T1 or T3 \((\chi^2(4) = 4.64, p = .32)\).

At T1, chemistry students were more likely to be in the dogmatist profile \((\text{ASR} = 5.1)\) and less likely to be in the rationalist \((\text{ASR} = -3.5)\) and skeptic profiles \((\text{ASR} = -2.6)\) than neuroscience students \((\chi^2(4) = 38.69, p < .001)\). The pattern was similar at T3, with chemistry students more likely to be dogmatists \((\text{ASR} = 3.8)\) and realist/dogmatists \((\text{ASR} = 2.7)\) and less likely to be rationalists \((\text{ASR} = -3.1)\) and skeptics \((\text{ASR} = -2.7)\) than neuroscience students \((\chi^2(4) = 32.74, p < .001)\). This corresponds to theory, in that the
more advanced students who are more aligned with the epistemology of the social sciences would have more developmentally advanced positions.

4.2.3.2 Examining shifts in cluster membership

Over the course of the semester, the slight majority of students with data at both time points remained in the same cluster (50.6%). However, a number of students shifted to a different way of thinking about knowledge. The following analyses sought to differentiate between those students who shifted clusters based on the messages they reported hearing in class.

At T1, there was significant variability in students’ personal theories about knowledge. This was also reflected in the number of students in each cluster at T1 (36 rationalists, 21 skeptics, 78 uncommitted, 40 dogmatists, 48 realist/dogmatists). Different messages might have stronger effects on different students, so in order to investigate whether hearing certain messages predict changes in cluster membership, I focused on differences in reported messages between students in different clusters at T3 within students who started the semester in the same cluster (e.g., comparing students who started the semester as dogmatists and stayed in the dogmatist cluster vs. shifted to the skeptic cluster). To compare cells, ANOVAs with Tukey’s post hoc tests were conducted. Cells with fewer than five individuals were excluded from analyses. Full results of these analyses can be seen in Tables 9 and 10.
Table 9: Differences in Uncertainty in the Field Messages by Initial Cluster and Shift

<table>
<thead>
<tr>
<th>DV: Uncertainty in the Field</th>
<th>T2 Cluster</th>
<th>T1 Cluster</th>
<th>Rationalist</th>
<th>Skeptic</th>
<th>Uncommit.</th>
<th>Dogmatist</th>
<th>Realist/Dogmatist</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationalist</td>
<td>4.22</td>
<td>4.31</td>
<td>3.32</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4.05^b (8)</td>
<td>.09</td>
</tr>
<tr>
<td>(16)</td>
<td>(9)</td>
<td>(8)</td>
<td>(0)</td>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skeptic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(2)</td>
<td>(12)</td>
<td>(3)</td>
<td>(0)</td>
<td>(0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncommitted</td>
<td>3.33</td>
<td>3.52</td>
<td>3.50</td>
<td>2.53</td>
<td>3.30</td>
<td>3.30</td>
<td></td>
<td>.58</td>
</tr>
<tr>
<td>(6)</td>
<td>(5)</td>
<td>(32)</td>
<td>(6)</td>
<td>(10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dogmatist</td>
<td>-</td>
<td>-</td>
<td>3.20^b</td>
<td>2.01^a</td>
<td>2.49^ab</td>
<td>2.49^ab</td>
<td></td>
<td>.04</td>
</tr>
<tr>
<td>(0)</td>
<td>(1)</td>
<td>(6)</td>
<td>(13)</td>
<td>(7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Realist/Dogmatist</td>
<td>4.05^b</td>
<td>-</td>
<td>2.56^ab</td>
<td>2.08^a</td>
<td>3.12^ab</td>
<td>3.12^ab</td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>(8)</td>
<td>(3)</td>
<td>(6)</td>
<td>(5)</td>
<td>(17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Groups that do not significantly differ based on Tukey’s post-hoc tests are noted with matching superscripts.

Among students that began the semester as rationalists, at least five individuals were rationalists, skeptics, or uncommitted at T3. There was a marginally significant difference between these three groups in reported uncertainty in the field messages \(F(2, 30) = 2.586, p = .092, \eta^2 = .15\) and a significant difference in instructor help messages \(F(2, 30) = 5.016, p = .013, \eta^2 = .25\). With instructor help messages, students who remained in the rationalist cluster reported higher help messages than students who shifted to
uncommitted, which goes contrary to expectations that help messages make the instructor seem like the only tenable source of knowledge.

Table 10: Differences in Instructor Help Messages by Initial Cluster and Shift

<table>
<thead>
<tr>
<th>T1 Cluster</th>
<th>DV: Instructor Help</th>
<th>T2 Cluster</th>
<th>Realist/ Dogmatist</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationalist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationalist</td>
<td>4.60\textsuperscript{b}</td>
<td>4.30\textsuperscript{ab}</td>
<td>3.75\textsuperscript{a}</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(16)</td>
<td>(9)</td>
<td>(8)</td>
<td>(0)</td>
</tr>
<tr>
<td>Skeptic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skeptic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>(12)</td>
<td>(3)</td>
<td>(0)</td>
</tr>
<tr>
<td>Uncommitted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncommitted</td>
<td>4.06\textsuperscript{ab}</td>
<td>3.33\textsuperscript{a}</td>
<td>4.18\textsuperscript{ab}</td>
<td>4.44\textsuperscript{b}</td>
</tr>
<tr>
<td></td>
<td>(6)</td>
<td>(5)</td>
<td>(32)</td>
<td>(6)</td>
</tr>
<tr>
<td>Dogmatist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dogmatist</td>
<td>-</td>
<td>-</td>
<td>4.50</td>
<td>4.15</td>
</tr>
<tr>
<td></td>
<td>(0)</td>
<td>(1)</td>
<td>(6)</td>
<td>(13)</td>
</tr>
<tr>
<td>Realist/Dogmatist</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Realist/Dogmatist</td>
<td>4.46</td>
<td>-</td>
<td>3.87</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>(8)</td>
<td>(3)</td>
<td>(6)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

Note. Groups that do not significantly differ based on Tukey’s post-hoc tests are noted with matching superscripts.

Among students who were skeptics at T1, there was only one cell (remain skeptic) with more than five students. Therefore, no analyses were conducted on the initially skeptic students.

Among students that began the semester as uncommitted, there was no statistically significant differences by final cluster membership in reported uncertainty in the field messages ($F(4, 54) = 0.720, p = .582, \eta^2 = .05$) or instructor help messages ($F(4, 53)$)
However, it is worth noting that the direction of the latter conform with theory, in that Tukey’s tests found that the uncommitted students that became skeptics trended lower in instructor help messages \( (M = 3.33) \) while those that became dogmatists trended higher in instructor help messages \( (M = 4.44) \).

Among students in the dogmatists cluster at T1, there were significant differences between final cluster membership and uncertainty in the field messages \( (F(2, 23) = 3.739, p = .039, \eta^2 = .25) \), but not instructor help messages \( (F(2, 23) = 1.425, p = .261, \eta^2 = .11) \). Specifically, dogmatists who remained in the dogmatist cluster reported fewer uncertainty messages \( (M = 2.01) \) than students who shifted to the uncommitted cluster \( (M = 3.20) \). This aligns with theory, because the constructivist uncertainty message is associated with shifting to and uncommitted cluster that is more constructivist on each dimensions than the dogmatist cluster.

Finally, among students who began the semester in the realist/dogmatist cluster, students who ended the semester in different clusters differed in their reported uncertainty in the field messages \( (F(3, 31) = 3.769, p = .020, \eta^2 = .27) \), but not instructor help messages \( (F(3, 31) = 1.408, p = .259, \eta^2 = .12) \). Students who shifted from being realist/dogmatist to the more clearly dogmatist cluster reported fewer uncertainty messages \( (M = 2.08) \) than those that shifted to the more sophisticated rationalist cluster \( (M = 4.05) \). This again supports a theory in which uncertainty in the field messages provide evidence of a constructivist theory of knowledge.
4.2.4 Post-exam messages

Another objective of Study 2 was to approach a more sensitive measurement of students’ perceptions of messages in class after exams. In the week immediately following receiving feedback on their first exam, students were invited to participate in an online survey that asked them about what they remember hearing from their instructors about the exam.

4.2.4.1 Recorded instructor messages after exam

In order to contextualize students’ experiences after receiving their first exam, full class periods were recorded when students received exam feedback from their instructors. In the chemistry class, the primary instructor said the following about the exam (emphasis added):

Let me talk about the exam for just a couple of minutes, and then we’ll move on to the next unit. So the statistics you can see are on the whiteboard here. Uh the highest score was 88, which was almost a perfect score. There’s always somebody who does really well, so if you did really well, congratulations. Um, the good news is that you know that 9% of your score under your belt that you did well... You see that the class mean was 72, the median was 74 and the standard deviation was 11. So, some of you did not do as well as you would like, and so what I’ll say for you is that there’s good news and there’s bad news: the good news is that you have a couple of options if you didn’t do as well as you like. One is that your lowest test score is substituted by the final exam average, right? So, even if you weren’t able to get as well prepared as you would like, or maybe it was your first college test and you just weren’t sure what was gonna be there, um, it’s a certain kind of amnesty with that final exam replacement option. Um, the bad news is that usually the first exam is the highest average, um so that doesn’t mean it will be your highest average; there are people every semester that the first exam is their lowest, ok, but those people are not the majority of people in the class. Ok? So the most
important thing to remember, if you didn’t do as well as you would like on this test, is that it’s not the score that you got that’s the most important thing, ok. If your score is not as high as you would like, it’s how you react to it, right? What do you do? So you need to, hopefully, you’ve claimed your test, you’ve looked at the key, you’ve figured out what you missed – by the way, if you think your exam was graded improperly, um I have some regrade forms up here. OK? And I can look at it again, if you think it was graded improperly. Um, the other thing is that you’re thinking about what I missed, you didn’t get as many points as you like, you also have to think about what am I doing to prepare for class each day, and what can I change so that on the next test I get more points, and I’m more in line with how I would like to do, right? So it’s really under your control. Um, and how you react to the score is much more important than what the score is. OK? If your score is MUCH lower than the mean, let’s say two standard deviations below the mean or more, which would be 50 or less, on this test. Then you might want to think about another option, which is the department has an option that you can petition Dr. [sic], to move down to Chem 101. They’re doing the topics in the same order, so you haven’t missed any topics. They’ve done thermodynamics and atomic structures, the same things we’ve done, ok? So if you did very poorly, not just around the average – ’cause I know for a lot of you, you’ve never had an average score in your life and you’re freaking out because you got a bad grade, OK. The-this petition to switch classes is not for you, OK. The petition to switch classes is if you really did a lot lower than the mean, like around the low 50s or below. Is that clear to everybody? That’s an option that only the chemistry department has, other departments don’t have it. It’s only for people that it’s their first semester taking chemistry, which is why you all do. OK, so I want to make sure everybody knows about that. And if you’re going to do that, you have to do it within a week of when the test was handed back in class, which means you need to make that decision by Friday. Which means you need to have talked to Professor [sic], by Friday. Everybody got that? It doesn’t apply to most of you, but some of you I wanna make sure you know about it.

The bolded lines constitute what could be interpreted as conciliatory or spurring messages. The instructor is simultaneously attempting to blunt the emotional impact of a negative grade (a conciliatory message) and also providing a spurring message, that
there are numerous things students can do if they did not receive as high of a grade as they would like. The instructor provides general descriptive statistics about the test, but does not describe those results in a way that would meet the criteria of positive or negative feedback.

In the neuroscience class, the instructor said the following regarding the exam:

Um so quick, I posted a note on Piazza for those who have not already checked today, with the exam data, so I, basically what I did, when I post stuff, they’re not posted yet because I don’t want you spending all class looking at your exam stuff, um what I did was I took the score from your in-class quiz and just put in the box for question 16 when you get onto Sakai. So everything will be loaded up and that’s why I did the total combined possible data, because that’s how it’s going to appear when you log into Sakai. The total combined possible score is 50. The high score, which a handful of people got, was 47.5. The average was about 43.15-ish. And by “ish” I mean that it was 43.15. Uh The standard deviation was about 3. So if the world were to end today, there would not be a curve. Right? Hopefully the world doesn’t end today, so we don’t have to address that possibility. Also, if the world ended today, who cares what you got on the exam? Not the point though. Uh, so the total score, the average of 43. The in-class was a little bit funny, because as I was grading them and I’m looking through and some of the ones where people did really well, I’m like, “Oh, they only got 18 out of 18.5, I don’t remember them missing any other questions…” Yeah so the in-class portion was actually out of 19… Whoops! Um, so everybody got a point. Yay! And so I just added a point. And so what you get when you pick up your paper copy is the number of points out of 19. And then theoretically, if I didn’t screw up, the number that was typed into the box on Sakai is this plus one. Please check that, because it’s entirely possible that I screwed that up somewhere, OK? Um, and probably, I screwed it up to your deficit. So it’s really worth looking to make sure that the, and this will make sense when you look on Sakai, the that the number in the last question box for the in-class exam should be this number plus one. K? Um, reminder that all regrade requests must be in some sort of written format, email is fine, and that generally speaking, the success rate of the regrade request that are things beyond like, “Hey you forgot to type the
number in the box,” is fairly low, so please, please do not send me a regrade request for everything you lost points on. It will just make me sad and it will make me not like you. Valid regrade requests, totally reasonable. But please just don’t do that thing where you go through every single question, because that’s no fun for everybody... it just necessitates me writing you a very long email. Um, other notes, I know I asked you to put all of the links to the outside sources that you used in that last question box. I obviously am not going to go through and click through every link for every test as I’m grading the exam, but if there was something in one of those links that justifies an answer choice that I marked incorrectly, that would be a great time to be like “Hey, check the link,” and I’ll be like, “Hey, cool, all right, reasonable. Fixed it.” Um, what’s the other thing? Oh! The last thing: I don’t know how many of you have taken some sort of crazy type thing on Sakai that involved typing actual words into boxes, Sakai’s really stupid about being able to grade things where you actually input letters and words, so on the neuron labeling it’s going to mark a whole bunch of stuff incorrect that is actually correct. I hand-graded that, so make sure that the points look like you would expect, even though there’s little red X’s next to the grades. And that should be it. So yeah, check the points, ignore the little red X’s and green check marks, they’re largely irrelevant. Any questions on that even though it’s completely theoretical at this point.

Although the instructor provided general information about the class mean and standard deviation, there was not a statement that constituted generalized positive or negative feedback in this case. However, the message about the world ending means the test does not matter has a certain conciliatory undertone to it.

Most importantly in the current study, however, is not what was actually said in class, but what students remember hearing about the exam. Students were asked to recall everything they remember from these transcripts in the post-exam survey.
4.2.4.2 Coding of student reported post-exam messages

Students provided diverse responses of what they remembered the instructor saying in class, despite the fact that students were in the same classroom and the surveys were all completed within a week of the instructor discussing the exam. Independent coders looked for student descriptions that resembled the self-reported message items conciliatory and spurring messages and generalized positive and negative feedback. The coders reached strong agreement on each of the four categories ($\kappa = 0.74$-$0.95$). Disagreements were resolved through discussion. Frequencies of the types of messages students reported in each class are reported in Table 11. Several students provided only factual information from the instructors’ exam discussion (e.g., class average), which was not considered to be either positive or negative feedback.

Table 11: Frequency of Student-Reported Instructor Post-exam Messages

<table>
<thead>
<tr>
<th></th>
<th>Chemistry</th>
<th>Neuroscience</th>
<th>$\chi^2$(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conciliatory messages</td>
<td>9 (13.4%)</td>
<td>2 (2.6%)</td>
<td>5.962*</td>
</tr>
<tr>
<td>Spurring messages</td>
<td>12 (17.9%)</td>
<td>0 (0%)</td>
<td>15.045*</td>
</tr>
<tr>
<td>Generalized positive feedback</td>
<td>21 (31.3%)</td>
<td>23 (29.9%)</td>
<td>ns</td>
</tr>
<tr>
<td>Generalized negative feedback</td>
<td>2 (3.0%)</td>
<td>1 (1.3%)</td>
<td>ns</td>
</tr>
</tbody>
</table>
Note. Frequencies and proportions (in parentheses) of each type of message are reported by class. Results of chi-square analyses in the third column demonstrate significant differences of frequencies in reporting across the two classes.

* $p < .05$

Results suggested that many students ($n = 44, 30.6\%$) reported generalized positive feedback, despite the fact that both instructors focused on descriptive information about exam performance. Generalized negative feedback was rarely reported by students ($n = 3, 2.1\%$). Given the low frequency of these messages (combined with not seeing these messages sent by the instructors), this code was excluded from the final regression analyses below. Spurring ($n = 12, 8.3\%$) and conciliatory messages ($n = 11, 7.6\%$) were far more likely to be reported in the chemistry classroom than the neuroscience classroom, which corresponds to the frequency with which instructors actually sent these messages about the exam.

4.2.4.3 Using post-exam message codes to predict changes in personal theories

Using the remaining three codes for student-reported of messages, a series of regression analyses were used to examine whether similar patterns exist for predicting changes in personal theories as the T3 student-reported message survey items. The final results can be found in Tables 12-13.

The first series of regressions (shown in Table 12) were conducted to predict students’ T2 personal theories, immediately following the exam. For incremental theories of intelligence, only students’ initial incremental intelligence was a significant predictor in the first three steps. However, there was a significant increase in $R^2$ in the
fourth step ($\Delta R^2 = .04$, $F(3, 120) = 3.982, p = .01$) due to a significant interaction between exam grade and reporting generalized positive feedback ($\beta = .22, p = .008$). The nature of this interaction is shown in Figure 8. General positive feedback was more important for students who performed poorly on the exam. Specifically, among students who performed poorly, reporting generalized positive feedback corresponded to more entity theories of intelligence while those who did not report positive feedback had more incremental theories. There were no clear effects of post-exam messages on students’ perceived competence or effort beliefs immediately after the exam.
Table 12: Hierarchical Regressions Predicting T2 Personal Theories from Post-exam Messages

<table>
<thead>
<tr>
<th>Step</th>
<th>Criterion Variable</th>
<th>T2 Incremental Intelligence</th>
<th>T2 Effort Beliefs</th>
<th>T2 Perceived Competence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>T1 Personal Theory</td>
<td>0.72***</td>
<td>0.54***</td>
<td>0.54***</td>
</tr>
<tr>
<td></td>
<td>Class</td>
<td>0.08</td>
<td>0.21*</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.537***</td>
<td>0.311***</td>
<td>0.293***</td>
</tr>
<tr>
<td>Step 2</td>
<td>Exam Grade</td>
<td>-0.15*</td>
<td>-0.17†</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.000</td>
<td>0.006</td>
<td>0.017†</td>
</tr>
<tr>
<td>Step 3</td>
<td>Generalized Pos. Feedback</td>
<td>-0.12†</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Conciliatory Message</td>
<td>-0.06</td>
<td>-0.04</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>Spurring Message</td>
<td>0.07</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.004</td>
<td>0.013</td>
<td>0.006</td>
</tr>
<tr>
<td>Step 4</td>
<td>Pos. Feedback × Grade</td>
<td>0.22**</td>
<td>0.17</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Conciliatory × Grade</td>
<td>0.16†</td>
<td>0.05</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Spurring × Grade</td>
<td>0.07</td>
<td>-0.01</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.042*</td>
<td>0.017</td>
<td>0.010</td>
</tr>
<tr>
<td>Total R²</td>
<td></td>
<td>0.583</td>
<td>0.346</td>
<td>0.327</td>
</tr>
</tbody>
</table>

Note.

*** p < .001
** p < .01
* p < .05
† p < .10.
Figure 8: Interaction between generalized positive feedback and exam performance on T2 incremental theories of intelligence.

The second series of regressions (Table 13) predicted T3 personal theories farther removed from students’ post-exam feedback. The only discernible effect of post-exam messages was for T3 perceived competence. The fourth step contributed to a significant increase in variability explained ($\Delta R^2 = .07, F(3, 102) = 4.22, p = .007$) due to a significant interaction between exam performance and spurring messages ($\beta = .28, p = .001$). The nature of this interaction is displayed in Figure 9. Since spurring messages were only reported in the chemistry class, interpretation of this interaction should be limited to that context. Reporting spurring messages were associated with lower perceived competence for those who did poorly on the exam but higher perceived competence for those who did poorly on the exam. Notably, this pattern was not apparent in the T2.
Post-exam messages were unrelated to changes in T3 incremental theories, quick learning, personal justification, and effort beliefs.
Table 13: Hierarchical Regressions Predicting T3 Personal Theories from Post-exam Messages

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>T1 personal theory</td>
<td>0.70***</td>
<td>0.47***</td>
<td>0.60***</td>
<td>0.56***</td>
<td>0.57***</td>
</tr>
<tr>
<td></td>
<td>Class</td>
<td>0.05</td>
<td>-0.14</td>
<td>0.20*</td>
<td>0.18*</td>
<td>0.18*</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.521***</td>
<td>0.215***</td>
<td>0.361***</td>
<td>0.411***</td>
<td>0.307***</td>
</tr>
<tr>
<td>Step 2</td>
<td>Exam Grade</td>
<td>-0.06</td>
<td>0.05</td>
<td>-0.08</td>
<td>-0.07</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.000</td>
<td>0.003</td>
<td>0.003</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Step 3</td>
<td>General. Pos. Feedback</td>
<td>-0.05</td>
<td>-0.02</td>
<td>0.11</td>
<td>-0.12</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>Conciliatory Message</td>
<td>0.02</td>
<td>-0.18</td>
<td>0.06</td>
<td>-0.03</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Spurring Message</td>
<td>0.03</td>
<td>-0.02</td>
<td>-0.07</td>
<td>-0.00</td>
<td>-0.18</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.003</td>
<td>0.026</td>
<td>0.015</td>
<td>0.007</td>
<td>0.018</td>
</tr>
<tr>
<td>Step 4</td>
<td>Pos. Feedback × Grade</td>
<td>0.10</td>
<td>-0.04</td>
<td>0.01</td>
<td>0.13</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>Conciliatory × Grade</td>
<td>0.05</td>
<td>0.04</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Spurring × Grade</td>
<td>0.02</td>
<td>0.13</td>
<td>0.06</td>
<td>0.10</td>
<td>0.28***</td>
</tr>
<tr>
<td></td>
<td>ΔR²</td>
<td>0.008</td>
<td>0.018</td>
<td>0.003</td>
<td>0.022</td>
<td>0.074**</td>
</tr>
<tr>
<td>Total R²</td>
<td></td>
<td>0.532</td>
<td>0.263</td>
<td>0.382</td>
<td>0.440</td>
<td>0.400</td>
</tr>
</tbody>
</table>

Note.
*** p < .001
**  p < .01
*   p < .05
†   p < .10.
4.3 Discussion of Study 2

Study 2 expanded on the findings from Study 1 by modeling the complex relations between students’ personal theories of education and within-class differences in student-reported messages from instructors. Study 2 provided further evidence for the college theory theory, that students use non-content messages in class to adjust their personal theories about education, as well as the assertion that students’ incoming beliefs play a role in the types of messages they are more likely to notice or remember. Many of the relations were in the hypothesized direction (see Table 3), though there were also many unexpected patterns, as well as differences between the two classrooms.

Figure 9: Interaction between reasoned conciliatory messages and exam performance on T3 perceived competence.
4.3.1 Relations between messages and beliefs

I hypothesized that instructor help messages would positively predict justification by authority beliefs and negatively predict incremental theories of intelligence, quick learning, personal justification, and perceived competence. Instructor help messages were positively related to justification by authority and negatively related to personal justification, as hypothesized, in the neuroscience classroom. A perception that the instructor wanted students to come to her for help provides evidence that a teacher is necessary to construct knowledge and problems cannot be worked out on their own.

I hypothesized that conciliatory messages would positively predict personal justification and negatively predict incremental theories and perceived competence. However, in this study, students reporting more frequent conciliatory messages possessed more incremental theories of intelligence. Therefore, students who believe the instructors are trying to lessen the impact of a negative grade in these large courses were more likely to have more incremental beliefs, contrary to prior research (Rattan et al., 2012). However, this finding was unique because it investigated differences in perceptions of the message between students who share the same classroom. Disentangling between- and within-class effects is essential for this reason.

Differential ability was originally hypothesized to positively predict justification by authority, but it negatively predicted justification by authority, as in Study 1, in the
neuroscience class. Students who have a sense from instructors that some students are better than others were less likely to think that knowledge is coming from instructors. Perhaps these messages sent information that the ability of experts also differ, decreasing beliefs that knowledge is held by authority figures. More research is necessary to understand this somewhat consistent pattern of findings.

Uncertainty in the field messages were hypothesized to relate to constructivist beliefs: lower simple/certain knowledge and justification by authority and higher personal justification. In both classrooms, uncertainty messages were associated with lower simple/certain knowledge beliefs. When students feel that the instructor openly discusses uncertainty in the field, the theory that knowledge is a collection of unchanging facts must seem much less tenable. In the chemistry classroom, uncertainty messages were also associated with lower justification by authority and higher personal justification beliefs. Findings from the person-centered analyses also found uncertainty in the field messages implicated with shifts toward more constructivist epistemic positions.

Negative and positive generalized feedback were originally hypothesized to negatively and positively related, respectively, to quick learning and perceived competence. However, generalized negative feedback was associated with higher quick learning beliefs in both classes, in line with the findings from Study 1. Students generally reported low levels of generalized negative feedback from instructors.
Therefore, it may be that students who remember some degree of negative feedback take these messages as evidence that the efforts put in by the class have not produced the desired results and never will. Furthermore, those who initially believe that learning must happen quickly are more likely to recall negative feedback from the instructor. This suggests a self-reinforcing loop, in which students who believe learning happens quickly are particularly perceptive to generalized negative feedback, which provide them with information that the topic has not been learned, and therefore is not going to be learned from additional effort.

While there was no specific hypothesis for the relation between generalized positive feedback and personal theories about knowledge, positive feedback was related to justification by authority beliefs in the chemistry class. Students who remembered the instructor saying the class was doing well increased their beliefs that the experts were the source of chemistry knowledge. The reason for this is unclear, though one possible explanation is that students who are more apt to notice positive feedback also hold positive feelings towards the instructor. In both classrooms, instructors were rated as experts (agreement greater than 6 on a 7-point scale to the item, “The principle instructor had a thorough knowledge of the subject matter.”) It is possible that positive feedback triggered positive affectivity, contributing to making instructors seem like a more valid source of knowledge. This finding will have to be examined in future research.
In addition to quick learning predicting reports of negative feedback, students with higher perceived competence were more likely to report instructor help, conciliatory, and uncertainty in the field messages. The effects were somewhat small in magnitude, and the reason for these relations is not immediately clear. However, one possibility is that students who feel more strongly about their abilities are more likely to attend class (Van Blerkom, 2001). Higher attendance might then provide students with more opportunities to hear a variety of messages in the classroom.

Personal justification beliefs in neuroscience were predictive of uncertainty messages and instructor help messages. In the supplemental model, personal justification beliefs in chemistry predicted less frequent instructor help messages. The prediction of uncertainty messages fits with the notion that students’ with a more constructivist personal theory of knowledge are more likely to notice, recall, or imagine more constructivist messages in the classroom. In the supplemental model, believing chemistry knowledge was personally constructed related to students being less likely to notice, recall, or imagine that the instructor tells students that they should come to them when they do not understand the material. The opposite finding in the full model for the neuroscience class might be explained by the wording of the items. The instructor help items first state that students are working on a problem, and then that failure will require them to ask the instructor for help. The first half of the items might be what personal justification items are predicting: those that sense people can have different
understanding of knowledge might be more receptive to the fact that they must work through problems. The second half of these items, that instructors have the solution to these struggles, would then still be the component of the message that is related to lower personal justification. This idea is somewhat supported in the supplemental model, which distinguishes between messages that the instructor help messages negatively relate to personal justification while instructor support positively relate to personal justification in the neuroscience classroom.

Finally, the students with more incremental theories of knowledge were more likely to report reasoned conciliatory messages. The reasons given for failure in these items all suggest that their might be future success (e.g. “It’s OK if you did badly because you’re still new at this.”), and implicitly that one’s ability can improve. This fits with an incremental theory of intelligence and fits with evidence that students with different theories about intelligence make different attributions for failure (Dweck & Leggett, 1988; Hong, Chiu, Dweck, Lin, & Wan, 1999). In the neuroscience class, students who believed that learning happens quickly or not at all were also more likely to report reasoned conciliatory messages, though this effect was in the opposite direction in the chemistry classroom, and is therefore difficult to interpret without further study.

In Study 1, two messages that seem to have positive intention from instructors, instructor help and conciliatory messages, were associated with some maladaptive changes in students’ beliefs. Of practical importance to educators might be how to strip
away the subtext of these messages ("Come to me for help because I don't think you can do this on your own," and "It is OK to do badly because there is nothing you can do about how smart you are.") and still send the intended, positive message. Study 2 demonstrates that to some extent, this is possible when the instructor help message is framed as instructor support or when conciliatory messages are accompanied by more clear reasons that should spur students to keep trying. The effects of specific linguistic cues can be very precise, as other studies have found (Cimpian, 2010; Mueller & Dweck, 1998). In summary, the findings demonstrate the potential of studying classroom messages as a mechanism of personal theory change.

### 4.3.2 Differences between classes

One interesting finding from Study 2 involves the unique patterns of findings across the two classes. These differences can be summarized into three categories: mean differences between the students and classrooms, differences in the relations of justification beliefs, and different relations of quick learning beliefs.

First, there were some modest mean differences in initial personal theories, final personal theories, and student-reported messages (see Table 8). For instance, conciliatory messages were more common in the chemistry classroom, uncertainty messages and generalized positive feedback were more likely to be reported in the neuroscience class, and epistemic beliefs were more constructivist in the neuroscience. The person-centered analyses also validated the domain-specificity of epistemic beliefs,
demonstrating that a hard field like students thinking about knowledge chemistry
demonstrated less developmentally advanced beliefs (e.g., dogmatism, realism) while
those thinking about neuroscience, housed within a social science program, reported
more developmentally advanced beliefs (e.g., rationalism, skepticism).

Second, the paths from messages to students’ justification beliefs varied in the
two classes, while the paths to the domain-general beliefs (incremental intelligence,
quick learning), simple/certain knowledge, and perceived competence did not.
Specifically, justification by authority beliefs and personal justification were predicted
by uncertainty in the field in the chemistry classroom, but were predicted by help,
conciliatory, differential ability, and generalized negative feedback in the neuroscience
classroom. There are several possible explanations for this, including the domain
differences and classroom structure differences. For instance, frequency of messages and
the variability between messages differed between the two classes as did epistemic
beliefs at the beginning and end of the semester. These differences highlight the
importance of examining epistemic cognition within specific domains.

Third, students’ incoming personal justification of knowledge was more
predictive of the messages students reported hearing from the instructor in the
neuroscience class in the full model. One explanation that fits with the current theory is
that the more experienced students possessed more solidified personal theories at this
point in their educational development than the largely freshman group in the chemistry
class. Alternatively, it may be the case that personal justification is more relevant in an introductory neuroscience classroom than an introductory chemistry classroom. Students may have had more opportunities to see conflicting views reading neuroscience studies (a common practice in classes taught in psychology departments), whereas introductory chemistry classes might focus on teaching basic concepts that are essential for later chemistry courses with less opportunities to see differences of opinion.

4.3.3 Development of personal theories in college

The findings also provide a number of other insights into the development of personal theories in college students. First, the autoregressive paths, while not of primary interest in the investigation of messages, were substantially varied between the different beliefs. Specifically, incremental theories of intelligence were the most stable, followed by the knowledge justification domains, perceived competence, and subsequently the speed of learning and simple/certain knowledge which were the least stable. This indicates that beliefs like quick learning and simple/certain were much less stable, and therefore open to change during college, whereas theories of intelligence were much more resistant to change during this time period. Therefore, certain personal theories seem to develop and solidify earlier (theories of intelligence) while others continue to adjust substantially in college (quick learning, simple/certain knowledge). In considering the two least stable personal theories, these are indeed theories that might be particularly influenced by the college environment. Many students come to college
without the necessary study skills (Nist & Simpson, 2000). If college is the first time students have to study, this may produce volatility in the quick learning personal theory. The lower stability might also be explained by the apparent sensitivity to generalized negative feedback and subsequent prediction of higher quick learning beliefs. College has also long been held up by proponents of liberal education as a time to focus on understanding over valuation of facts (e.g., Astin, 1993; Goheen, 1961), which corresponds to the relative mutability of the simple/certain knowledge theories.

Therefore, quick learning and simple/certain knowledge personal theories might be more open to change in the college context.

Interestingly, perceived competence was not more stable than many other beliefs, but was not predicted by messages in the main model. This is likely due to the fact that there are many other factors that contribute to changes in perceived competence (Bandura, 1977), including mastery experiences. For example, the analyses using post-exam messages found that the effects of messages depend on other individual differences, such as exam performance. When investigating self-theories, researchers should consider how messages interact with individual factors.

Developmentally, the timing of measurement is not only a methodological dimension, but a theoretically interesting one. Post-exam messages were generally less predictive of personal theory change than the T3 reported messages. There are several possible explanations for this, some methodological and some theoretical.
Methodologically, the analyses using post-exam messages had lower power due to a smaller sample of students completing the T2 survey and the decreased variability in the independent variable (binary instead of the Likert-scales; closer to objective responses to a single recent event). The null results might also be due to the time of measurement, given the distance between the T2 and T3 responses, though this seems less likely given that the number of significant results was equivalent when the criterion variable was measured at T2 or T3. A more theoretical rationale explains that the accumulation of messages over an entire semester are more important for changes in personal theories. College students have already formed somewhat stable theories about knowledge, as evidenced by the autoregressive paths in the path models and the stability in cluster membership in the ISOA. Therefore, evidence from a single message that undercuts or contradicts an existing personal theory may not be powerful enough to change it. The self-report measure at T3 measured students’ overall sense of the frequency of a variety of related messages over an entire semester, which according to a developmental model of personal theory change is a better assessment of the compounded evidence that can change a students’ thinking about education. However, there were interactions between messages and exam performance on certain personal theories, suggesting there may be some immediate and long-lasting effects of messages for specific individuals. Interestingly, this effect was immediate for theories of intelligence (on T2 TOI), but delayed for PC (on T3 PC). There is also evidence that theories of intelligence act like
mindsets that can be primed (e.g., Burns & Isbell, 2007). It is possible that asking students to consider the exam feedback primed themselves, leading to an immediate effect on the post-exam survey that disappeared over time, given the high stability in theories of intelligence over the semester. Another interpretation would be that the most pressing evidence for informing the incremental theories was the exam feedback, but as students collected more evidence throughout the semester, this first piece of evidence became less proportionally substantial. The effect of spurring messages on perceived competence also corresponds to research of the effect overinflated praise has on young children (Brummelman et al., 2014). A positive message about the overall class’s performance likely comes across the same way to students who did poorly on the exam. The detrimental effect of spurring messages on low exam performers took time to unfold. The delayed effect might suggest that the students in the chemistry class were initially resilient. This was their first college chemistry exam, and so a spurring message might be interpreted neutrally at first, but reinterpreted as the instructor merely trying to manage their emotional response, degrading perceived competence by the end of the semester.

Finally, the person-centered analyses add to a burgeoning literature of investigating systems of epistemic beliefs (Barger et al., under review; Chen 2012; Greene et al., 2008; Schommer-Aikins, 2004). Students’ responses fell into predictable profiles that map onto developmental theory. While the sample size was relatively small
when broken down into five clusters, this study provides the first evidence that specific messages are associated with developmental shifts from one way of thinking about knowledge in the classroom to another. Furthermore, the effect sizes were substantial, suggesting a fruitful area of future research. The findings from the person-centered correspond with and strengthen the variable-centered analyses. This serves as a strong reminder about the importance of conceptualizing the nature of development.
5. General Discussion

The current studies sought to answer three questions. First, what messages do instructors send in class? These studies demonstrate that students and instructor report a number of non-content messages, including instructor help and support messages, differential ability messages, conciliatory messages, uncertainty in the field messages, generalized positive feedback, and to a much lesser extent, generalized negative feedback. Second, I asked whether messages in class predict changes in students’ personal theories about education. The answer to that basic question is unequivocally yes, though the nature of these relations is often quite complex. Third, I investigated whether students’ initial personal theories predict whether they report class messages. There is also some evidence that students’ initial personal theories predict whether or not individual students recall certain types of messages in the classroom, though again, the nature of these relations was not consistent across studies. The findings from both studies have been summarized in Table 14.
Table 14: Complete Summary of Between- and Within-Class Findings from Study 1 & Study 2

<table>
<thead>
<tr>
<th></th>
<th>HELP</th>
<th>DIFF</th>
<th>CONC</th>
<th>UNCE</th>
<th>NEGF</th>
<th>POSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental Theories of Intelligence</td>
<td>B I</td>
<td></td>
<td></td>
<td>B 2CN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quick Learning</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td>W 2CN</td>
<td></td>
</tr>
<tr>
<td>Simple/Certain Knowledge</td>
<td></td>
<td></td>
<td>B I 2CN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justification by Authority</td>
<td>2N I 2N</td>
<td></td>
<td></td>
<td></td>
<td>2N 2C</td>
<td></td>
</tr>
<tr>
<td>Personal Justification</td>
<td>W I 2N</td>
<td></td>
<td></td>
<td>B WI 2C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived Competence</td>
<td></td>
<td></td>
<td></td>
<td>W</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Significant between-class effects in Study 1 are noted with a “B” and within-class effects are noted with a “W” and the direction of the effect is noted by the color, with green indicating a positive effect and red indicating a negative effect. Instructor effects from Study 1 are noted with an “I.” Within-class effects from Study 2 are noted with a “2” with subscripts corresponding to the chemistry (C) and neuroscience (N) classrooms. Effects in bold were significant ($p < .05$) while effects in italics were marginally significant ($p < .10$). Effects in black were in the hypothesized direction while those in white were in the opposite direction to what was hypothesized. HELP = Instructor help. DIFF = Differential ability. CONC = Conciliatory messages. UNCE = Uncertainty in the field. NEGF = Generalized negative feedback. POSF = Generalized positive feedback.

5.1 Comparing and Contrasting Results from Study 1 and 2

Study 1 focused on between- and within-class effects of messages, while Study 2 focused exclusively on within-class differences in two large classrooms. Therefore, there were a number of different results across the two studies. However, a number of findings were similar across the two studies.
Perceptions of uncertainty in the field messages were associated with lower simple/certain knowledge beliefs at the end of the semester in both classes in Study 2. This corroborates the finding from Study 1, which found that classes in which these messages were more prevalent predicted lower simple/certain messages. Taken together, this supports the hypothesis that students use uncertainty in the field messages as information to adjust their beliefs about the nature of knowledge as a collection of unchanging facts downward.

Generalized negative feedback was a relatively strong predictor of higher quick learning beliefs in both classes in Study 2. This corroborates the within-class findings from Study 1. This supports the assertion that increased perceptions of the negative feedback to the class at large reinforces they idea that if learning does not happen quickly, it will not be able to occur. To some extent, this idea seems to be self-reinforced, as students who come into the classroom with quick learning beliefs are also more likely to report hearing generalized negative feedback in the classroom. If students already interpret failure to learn something immediately as a sign of failure, they may be particularly attuned to signs of failure from their instructor. Given the relatively low occurrence of such feedback in the classroom \((M < 2.0\) in each study), it is possible that these students are interpreting more innocuous messages as being negative feedback. These two findings linking negative feedback to quick learning beliefs and uncertainty
in the field messages to simple/certain knowledge beliefs (and to some extent personal justification beliefs) appear to be robust to context and situation.

Other findings from Study 1 were replicated in one of the Study 2 classrooms, but not the other. The hypothesized link between help messages and decreased personal justification beliefs were found in the neuroscience class. Additionally, uncertainty in the field messages were associated with increased personal justification beliefs in the chemistry class. Both fit with the theory that instructor help messages suggest that students should not be expected to figure things out on their own while uncertainty messages suggest that students will have to figure things out on their own. There was also evidence in the neuroscience class and in the instructor effects analyses of Study 1 that differential ability messages were associated with lower justification by authority. These findings suggest general patterns of messages being associated with changes in beliefs that generalize to some extent but are not completely robust.

One finding was particularly distinct between Study 1 & 2. Unlike the findings from Study 1, conciliatory messages were found to be associated with higher incremental theories of intelligence. This finding is surprising given findings from prior research (Rattan et al. 2012) and the between-class differences from Study 1. The major difference in Study 2 was the fact that this positive association is occurring within classrooms, where students presumably receive similar messages. Therefore, variability in student-reports of this message are differences in students’ perceptions and not
variability in the messages actually being spoken by the instructor. For instance, the
difference in messages reported in different classes might act as markers for
environmental differences in the class, such as perceived entity environment (Good et
al., 2012). Alternatively, if individual students are perceiving these types of messages,
this might be a marker of a psychological difference in the student. Based on these
results, one example could be students’ initial theories, such as personal justification
beliefs (Study 1) or perceived competence (Study 2, neuroscience).

Several within-class findings in Study 1 and Study 2 were unique to those
studies. There was no evidence in Study 2 for the links between generalized feedback
and perceived competence in Study 1 but instead between generalized feedback and
personal theories about knowledge. The decreased quick learning beliefs with instructor
help messages were also unique to Study 1. The measures were as similar as possible
across studies, but there were many differences between the studies that might help to
explain the lack of consistency (see Table 1). For example, the period of time was much
shorter in Study 1, but students were also in their classroom on a daily basis and
typically not enrolled in other classes simultaneously. This might explain why there
were more direct message to perceived competence relations in Study 1, where students
were only able to draw from one class experience for their personal theories of their own
ability, instead of potentially drawing from experiences in other related classes.
Furthermore, none of the paths from initial beliefs hypothesized based on Study 1 (see Figure 3) were supported in Study 2, and the paths that were identified as significant in Study 2 were unique findings. Again, the multiple differences between the studies and populations make it difficult to determine the reason for these differences.

One possibility is that certain beliefs are more strongly formed later in college, and therefore more likely to influence students’ perceptions or memories of the classroom. For example, personal justification was generally more predictive of student-reported messages in Study 1 and the neuroscience class, which included older students than in the chemistry class. Alternatively, the variability in student-reported messages might be better explained by other factors that were not taken into account in these studies. This is understandable, as students in the same class are presumably having objectively similar experiences. For instance, students that attend class more frequently or pay attention more closely in class might report more messages. The variability left to be explained by students’ initial personal theories would then be relatively small (Study 2: $\beta_s < .25$). This makes finding relations between incoming personal theories and student-reported messages a statistically more difficult task, and might explain some of the inconsistent null findings. Taken together, the disparate findings for the links between initial personal theories and student-reported messages across studies and classrooms suggest that the role initial beliefs play depend on the context and possibly the strength of the initial belief at different stages of development.
5.2 Insights into the Development of Personal Theories

The current studies not only provide preliminary evidence for how students use classroom messages as evidence for changing their personal theories but also provide insight into broad developmental processes at play in the college classroom and how researchers might assess them.

First and foremost, these studies serve as proof-of-concept and validation for the measurement of student-reported messages in class. Not only do students report hearing these messages, but the frequency of these messages vary by class (Study 1) and within classes. Furthermore, student-reported messages correspond to teacher-reported messages and predict students’ personal theories at the end of the semester while controlling for their initial beliefs. Asking students to recall what types of messages their instructors send by providing them with exemplary messages proved effective, even more effective than asking them to freely recall within one week what instructors said after a meaningful classroom experience, the first class exam. Theoretically, the presumption is that these reports reflect how often students hear these types of messages in the classroom. However, at this early stage, it would be more prudent to assume that student-reports of messages serve as “markers” for students’ perceptions of the classroom. They might mark differences in actual messages heard in class or serve as indicators to more ambiguous differences in classroom climate (Barger et al., 2015; Good et al., 2012; Feucht, 2010). In other words, when a student reports hearing more
uncertainty in the field messages, for instance, this might be because the instructor says these messages frequently, or because the nature of the subject, a teachers’ constructivist attitudes, or the constructivist design of the course creates a classroom climate in which such a message seems more likely to have been reported. Regardless of the many factors that likely produce variability in student-reported messages, they have proven useful in understanding personal theory development in college students.

The findings of this study also demonstrate that the message effects occur through multiple pathways at the class and individual level. This helps to fill a fundamental gap in a literature full of experimental studies. Specifically, some messages relate to beliefs at the class level (instructor help messages on incremental theories), others at the individual level (e.g., generalized negative feedback on a number of personal theories), and some a combination of both (e.g., uncertainty in the field messages on personal theories about knowledge). Conciliatory messages even related in opposing fashion to incremental theories of intelligence when examined at the class level (Study 1) and the individual level (Study 2). These differences are significant in studying how experiences shape personal theories. Take generalized negative feedback, for example. All of the relations between generalized negative feedback and final personal theories were within-class differences. Given that this message was not frequently reported, individuals that report such messages are likely interpreting more neutral messages as negative feedback, and the individual differences underlying those
different reports (e.g., initial personal theories, neuroticism, avoidance orientation). The current studies were not designed to specifically investigate the reasons for differences in between- and within-class effects. However, they do encourage researchers to consider both between-class and within-class effects of messages on personal theory change and begin to examine why certain messages function differently.

The differences across studies and classrooms also serve as an important reminder that many personal theories are contextualized within a particular field. With the exception of theories of intelligence and quick learning, personal theories were assessed at a domain specific level matching the course’s subject matter (i.e., statistics, chemistry, math, psychology, neuroscience, biology). The results also seemed to vary depending on the topic matter in Study 2, particularly for justification by authority and personal justification beliefs. This reaffirms the assertion that personal theories about knowledge (Muis et al., 2006; Chinn et al., 2011) and the self (Marsh, 1990b), should be measured at a domain-specific level. The three epistemic beliefs assessed in the current studies seem applicable to many domains, but it might be useful to consider dimensions of personal theories that are unique to a particular discipline (see Hammer & Elby, 2002), such as beliefs about the utility of the scientific method in constructing knowledge (Tsai, 1998).

The effects of messages not only depend upon the context, but the individual. Students in the same classroom report remembering different messages. Whether these
differences are due to initial personal theories, attention or memory biases, differences in message saliency, or something as mundane as class attendance, remains open to interpretation. Regardless, when an instructor sends a message, there is no guarantee that message will reach all students (Perrenoud, 1998). There is also no guarantee that students will interpret this message the same way as students’ change their personal theories. This was apparent from the interactions between message and exam performance in Study 2. A reasoned conciliatory message and generalized positive feedback were associated with different personal theory changes for students who performed worse on an exam than students who performed better. Altogether, it is worth considering not only the main effects of messages on personal theory change, but also how these effects change in different contexts and for different individuals.

Finally, while the participants in this study were all young adults, the findings shed light on the timing of development. Specifically, some personal theories proved more stable than others. Researchers have suggested that theories of intelligence begin to develop at an earlier age (Dweck, 2002; Kinlaw & Curtz-Costes, 2003). It might prove useful to target theories of intelligence at younger ages in addition to in college classrooms, since messages still predicted changes in theories of intelligence. Quick learning beliefs and simple/certain knowledge beliefs were much less stable in these samples, and therefore might be a better target for college instructors who wish to change their students’ beliefs.
5.3 Limitations and Future Directions

The objective of these studies was to identify theoretically interesting non-content messages that students hear in their actual classrooms and investigate the interplay between college students’ developing personal theories and these messages. As these studies were an attempt to broadly investigate messages and various personal theories about education, the results have been interpreted here through a more exploratory, rather than confirmatory, lens. While several hypotheses were confirmed, many hypothesized findings were not, and still other relations were found to be opposite of the hypothesized direction. Therefore, both the results and the shortcomings provide insight into new areas of research in personal theory development.

There were many differences between study participant and classrooms that made it difficult to clearly interpret conflicting findings. Future research might combine elements from each study to form a more ideal study for examining messages and personal theories. Specifically, an ideal study design would involve a large number of highly similar classrooms, preferably in the same topic, with many different instructors. For example, many universities have required courses that are taught to a large number of students but require small classroom forms of instruction. This design would help to isolate the effects of messages by reducing the number of confounds. Assessing differences between classes is particularly important if the ultimate goal of this research
is to provide advisors with advice on how to use linguistic cues to positively influence their students’ personal theories.

Given the considerable complexity of the findings of the current study, an alternative approach would be to bring these findings outside the classroom and into the laboratory. While there were undoubtedly theoretical benefits to studying these processes in vivo in the classroom, the researcher needs to cede a considerable amount of control over students’ experiences. Therefore, future research might complement the current approach by experimentally manipulating messages students hear and investigate whether this temporarily influences their reported personal theories or behavior that might reflect how they are using their personal theories (e.g., Snyder et al., 2014).

For many of the ancillary analyses with the post-exam messages and the person-centered perspective, the sample size was strained in the present study. When testing for interactions between multiple factors or splitting samples into many clusters, larger sample sizes will be necessary. The effect sizes in the present analyses (for instance, $\eta^2$ was quite large when looking at differences in reported messages by cluster shift) demonstrate the promise of more direct measurement of classroom messages and person-centered analyses in future research.
5.4 Implications for Practice

The research conducted here is still in its infancy; it may be premature to apply the findings in these studies alone to make specific and universal recommendations to instructors in the classroom. However, this study demonstrates that the college classroom has the potential to provide students with more than information and skill development. Students constantly recalibrate their views about education, whether it is in their first semester of college chemistry or a brief summer course in statistics. Common, non-content messages phrases can send signals to students as they change their personal theories and potentially have repercussions beyond a single instructor’s classroom. This research does present interesting considerations for practitioners in the classroom.

One general guiding principle from this research is that instructors should consider how well-intentioned messages in the classroom are interpreted by students. Conciliatory messages and instructor help messages seem affectively positive, but were found to be associated with less positive changes in students’ personal theories. These findings link up to burgeoning research on microaggressions in the classroom (Suárez-Orozco et al., 2015), in which statements in the classroom contain subtext which is motivationally damaging to students. Instructors should occasionally consider what their words mean from a students’ point of view, or even a specific student’s point of view.
One robust finding across studies is the link between uncertainty in the field messages in the field and constructivist beliefs about knowledge. Based on these findings, I would tentatively encourage college teachers who hope to create constructivist thinkers not only talk about uncertainty in the field, but add opportunities to discuss uncertainty into their curriculum (e.g., Mason & Scrivani, 2004; Muis & Duffy, 2013). Differences in epistemic beliefs across fields (e.g., Trautwein & Lüdtke, 2007) might be partially explained by the willingness of soft sciences to talk about unanswered questions in the field. In Study 2, that was certainly the case, with students far more likely to report uncertainty messages in the neuroscience class than the chemistry class. Teachers should not be afraid to say, “We don’t really have an answer to that yet.” Teachers instead should embrace opportunities to talk about uncertainty explicitly in their classroom. Perhaps this will encourage students to adopt the personal theories that will help them find the answer themselves in the classroom and beyond.
6. Conclusion

A college education is about more than providing students with information and training technical skills. College instructors also have the opportunity to help students see their educational world in different ways. When instructors speak to their classrooms, they provide students with the evidence that students need to change their personal theories about knowledge during college. Different students in different contexts notice and fold this evidence into their theories in different ways. The complexity of students’ personal theories and the classrooms in which they develop them will serve as a prolific area for future research.
Appendix A

Reverse scored items marked with an asterisk. All items scores on a 1 (strongly disagree) to 5 (strongly agree) scale except for theories of intelligence, which is assessed using a 1 to 6 scale. Reliabilities are shown for measures included in Study 1.

**Personal Theories Scales**

**Theories of Intelligence** (Dweck, 1999; $\alpha_{T1} = .94$, $\alpha_{T2} = .92$)
You can change even your basic intelligence level considerably.
Your intelligence is something about you that you can’t change very much.*
To be honest, you can’t really change how intelligent you are.*
You can learn new things, but you can’t really change your basic intelligence.*
You have a certain amount of intelligence, and you can’t really do much to change it.*
No matter how much intelligence you have, you can always change it quite a bit.
You can always substantially change how intelligent you are.
No matter who you are, you can significantly change your intelligence level.

**Speed of Learning** (DeBacker, Crowson, Beesley, Thoma, & Hestevold, 2008; Wood & Kardash, 2002; $\alpha_{T1} = .71$, $\alpha_{T2} = .79$)
If something can be learned, it will be learned immediately.
You will get almost all the information you can understand from a textbook during the first reading.
You will just get confused if you try to integrate new ideas in a textbook with knowledge you already have about a topic.
Working on a difficult problem for an extended period of time pays off only for smart students.
Usually, if you are going to understand something, it will make sense to you the first time.
If I cannot understand something quickly, it usually means I will never understand it.

**Simple/Certain Knowledge** (adapted from Hofer, 2000; Greene et al., 2010; $\alpha_{T1} = .81$, $\alpha_{T2} = .88$)
All experts in [subject] understand the field in the same way.
Truth is unchanging in [subject].
In [subject], most work has only one right answer.
Principles in [subject] are unchanging.
All [subject] professors would probably come up with the same answers to questions in [subject].
Most of what is true in [subject] is already known.
In [subject], what is a fact today will be a fact tomorrow.
[Person who studies subject’s] knowledge of the facts about [subject] does not change.

**Justification by Authority** (adapted from Hofer, 2000; Greene et al., 2010; $\alpha_{T1} = .86$, $\alpha_{T2} = .86$)
If you read something in a [subject] textbook, you can be sure it’s true.
If my personal experience conflicts with ideas in a [subject] textbook, the book is probably right.
If a [person who studies subject] says something is a fact, I believe it.
Things written in [subject] textbooks are true.
I believe everything I learn in [subject] classes.
If a [subject] professor says something is a fact, I believe it.

**Personal Justification of Knowledge** (adapted from Hofer, 2000; Greene et al., 2010; $\alpha_{T1} = .60$, $\alpha_{T2} = .62$)
In [subject], the truth means different things to different people.
In [subject], everyone’s knowledge can be different because there is no one absolutely right answer.
In [subject], what’s a fact depends upon a person’s point of view.

**Perceived Competence** (PALS, 2000; $\alpha_{T1} = .88$, $\alpha_{T2} = .92$)
I’m certain I can master the skills taught in [subject].
I’m certain I can figure out how to do the most difficult class work in [subject].
I can do almost all the work in [subject] if I don’t give up.
Even if the work in [subject] is hard, I can learn it.
I can do even the hardest work in [subject] if I try.

**Effort beliefs** (Blackwell, 2002)
To tell the truth, when I work hard at my schoolwork, it makes me feel like I’m not very smart.
It doesn’t matter how hard you work – if you’re not smart, you won’t do well.
If you’re not good at a subject, working hard won’t make you good at it.
When something is hard, it just makes me want to work more on it, not less.*
The harder you work at something, the better you will be at it.*
If you’re not doing well at something, it’s better to try something easier.
If an assignment is hard, it means I’ll probably learn a lot doing it.*
If a subject is hard, it means I probably won’t be able to do really well at it.

**Additional Items (T1)**
What is your year in school?
With which gender do you most identify?
What is your race/ethnicity (check all that apply)?
Please select your discussion/lab section.

**Study 2, chemistry class: Post-Exam Feedback Survey**

What do you remember your instructor saying about the results of the exam? [open-ended]
[Perceived competence beliefs, see above]
[Effort beliefs, see above]
What grade did you receive on the exam?
What factors do you think led to you receiving this score?

**Message Items**

**Instructor help** ($\alpha_{students} = .82$, $\alpha_{instructors} = .80$)
If you need help understanding, I’m happy to help.
You should ask me questions if you get confused.
Ask me for help if you don’t understand.

**Differential ability** ($\alpha_{students} = .88$, $\alpha_{instructors} = .88$)
This is more difficult for some than others.
Some of you are going to get this right away.
Not everyone is going to find this easy.
Some of you might struggle with this.

**Conciliatory messages** ($\alpha_{students} = .87$, $\alpha_{instructors} = .89$)
It’s OK if you had a bad test.
You don’t need to panic because of one exam score.
Don’t worry too much about one bad test.
Don’t worry if you got an answer wrong.

**Uncertainty in the field** ($\alpha_{students} = .85$, $\alpha_{instructors} = .93$)
We don’t really have an answer to that yet.
This issue is not yet settled in the field.
The answer is seldom that simple.
There is mixed evidence in the field.
There are many controversies in the field.

**Generalized negative feedback** ($\alpha_{students} = .81$, $\alpha_{instructors} = .89$)
This class is not performing up to my standards.
I was not satisfied with how students performed on the last assignment.
I expect you all to do better next time.
The performances on the last exam were not satisfactory.

Generalized positive feedback ($\alpha_{\text{students}} = .78$, $\alpha_{\text{instructors}} = .81$)
The class average was very high.
Everyone did a good job on the test.
Everyone is doing very well so far.

**Study 2: Additional Message Items**

**Instructor support**
I am happy to set up an appointment outside of office hours.*
My door is always open.*
Come see me during office hours.*
I want to see you succeed.*
I will do whatever it takes for you to learn the material.*
I want everyone in the class to get an A.
* Used in final scale based on EFA.

**Spurring messages/Reasoned conciliatory messages**
You can improve after a bad test.
There will be other assignments for you to improve on.
That assignment is only a small part of your grade.*
If you had a bad test, you can still do better on the next one.
It’s OK if you did badly because you’re still new at this.*
It’s OK to struggle because this topic is hard.
Struggling on the test just means it’s not clicking yet.*
* Used in final scale based on EFA.
### Appendix B

*Model Selection Procedure and Fit Statistics for Primary Model*

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Description</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>RMSEA [95% CI]</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Base model: All paths varying across groups</td>
<td>89.679</td>
<td>60</td>
<td>.965</td>
<td>.074 [0.039, 0.104]</td>
<td>.047</td>
</tr>
<tr>
<td>B</td>
<td>Model A with autoregressive paths constrained across groups</td>
<td>97.839</td>
<td>66</td>
<td>.963</td>
<td>.073 [0.039, 0.102]</td>
<td>.055</td>
</tr>
<tr>
<td>C1</td>
<td>Model B, paths involving T3 INCR constrained across groups</td>
<td>101.280</td>
<td>72</td>
<td>.966</td>
<td>.067 [0.032, 0.096]</td>
<td>.055</td>
</tr>
<tr>
<td>C2</td>
<td>Model B, paths involving T3 QUICK constrained across groups</td>
<td>108.041</td>
<td>72</td>
<td>.958</td>
<td>.074 [0.043, 0.102]</td>
<td>.058</td>
</tr>
<tr>
<td>C3</td>
<td>Model B, paths involving T3 SCK constrained across groups</td>
<td>102.788</td>
<td>72</td>
<td>.964</td>
<td>.069 [0.035, 0.097]</td>
<td>.058</td>
</tr>
<tr>
<td>C4</td>
<td>Model B, paths involving T3 JA constrained across groups</td>
<td>110.720*</td>
<td>72</td>
<td>.955</td>
<td>.077 [0.047, 0.105]</td>
<td>.057</td>
</tr>
<tr>
<td>C5</td>
<td>Model B, paths involving T3 PJ constrained across groups</td>
<td>111.865*</td>
<td>72</td>
<td>.953</td>
<td>.078 [0.048, 0.106]</td>
<td>.053</td>
</tr>
<tr>
<td>C6</td>
<td>Model B, paths involving T3 PC constrained across groups</td>
<td>106.689</td>
<td>72</td>
<td>.959</td>
<td>.073 [0.041, 0.101]</td>
<td>.056</td>
</tr>
<tr>
<td>D</td>
<td>Model B, only T3 PJ and JA paths free to vary across groups</td>
<td>126.042</td>
<td>90</td>
<td>.958</td>
<td>.067 [0.036, 0.093]</td>
<td>.061</td>
</tr>
<tr>
<td>E1</td>
<td>Model D, paths involving T1 INCR constrained across groups</td>
<td>127.847</td>
<td>96</td>
<td>.963</td>
<td>.061 [0.027, 0.087]</td>
<td>.062</td>
</tr>
<tr>
<td>E2</td>
<td>Model D, paths involving T1 QUICK constrained across groups</td>
<td>133.161</td>
<td>96</td>
<td>.956</td>
<td>.065 [0.035, 0.091]</td>
<td>.063</td>
</tr>
<tr>
<td>E3</td>
<td>Model D, paths involving T1 SCK constrained across groups</td>
<td>135.781</td>
<td>96</td>
<td>.953</td>
<td>.068 [0.039, 0.093]</td>
<td>.064</td>
</tr>
<tr>
<td>E4</td>
<td>Model D, paths involving T1 JA constrained across groups</td>
<td>137.701</td>
<td>96</td>
<td>.951</td>
<td>.069 [0.041, 0.094]</td>
<td>.064</td>
</tr>
<tr>
<td>E5</td>
<td>Model D, paths involving T1 PJ constrained across groups</td>
<td>141.845*</td>
<td>96</td>
<td>.946</td>
<td>.073 [0.045, 0.097]</td>
<td>.067</td>
</tr>
<tr>
<td>E6</td>
<td>Model D, paths involving T1 PC constrained across groups</td>
<td>126.981</td>
<td>96</td>
<td>.964</td>
<td>.060 [0.026, 0.086]</td>
<td>.062</td>
</tr>
<tr>
<td>F</td>
<td>Final model, all paths constrained across groups except those involving T1 PJ, T3 JA, and T3 PJ</td>
<td>154.981</td>
<td>120</td>
<td>.959</td>
<td>.057 [0.025, 0.081]</td>
<td>.069</td>
</tr>
</tbody>
</table>

*Note.* Models marked with * denote a significant degradation in model fit from the more freed model. These changes were therefore excluded from the subsequent model.
Appendix C

Model Selection Procedure and Fit Statistics for Supplemental Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Model Description</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>RMSEA [95% CI]</th>
<th>SRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Base model: All paths varying across groups</td>
<td>46.889</td>
<td>24</td>
<td>.963</td>
<td>.103 [.058, .146]</td>
<td>.052</td>
</tr>
<tr>
<td>B</td>
<td>Model A, autoregressive paths constrained across classes</td>
<td>54.836</td>
<td>28</td>
<td>.956</td>
<td>.103 [.062, .143]</td>
<td>.067</td>
</tr>
<tr>
<td>C1</td>
<td>Model B, paths to T3 incremental theories constrained</td>
<td>60.219</td>
<td>32</td>
<td>.954</td>
<td>.099 [.059, .137]</td>
<td>.066</td>
</tr>
<tr>
<td>C2</td>
<td>Model B, paths to T3 quick learning constrained</td>
<td>62.876†</td>
<td>32</td>
<td>.950</td>
<td>.103 [.065, .141]</td>
<td>.070</td>
</tr>
<tr>
<td>C3</td>
<td>Model B, paths to T3 personal justification constrained</td>
<td>66.520*</td>
<td>32</td>
<td>.944</td>
<td>.109 [.072, .146]</td>
<td>.076</td>
</tr>
<tr>
<td>C4</td>
<td>Model B, paths to T3 perceived competence constrained</td>
<td>58.631</td>
<td>32</td>
<td>.957</td>
<td>.096 [.056, .134]</td>
<td>.066</td>
</tr>
<tr>
<td>D</td>
<td>Model B, only T3 PJ paths free to vary across groups</td>
<td>73.435</td>
<td>40</td>
<td>.945</td>
<td>.096 [.060, .130]</td>
<td>.069</td>
</tr>
<tr>
<td>E1</td>
<td>Model D, paths involving T1 INCR constrained across groups</td>
<td>74.515</td>
<td>44</td>
<td>.950</td>
<td>.088 [.051, .121]</td>
<td>.070</td>
</tr>
<tr>
<td>E2</td>
<td>Model D, paths involving T1 QUICK constrained across groups</td>
<td>87.446*</td>
<td>44</td>
<td>.929</td>
<td>.104 [.072, .136]</td>
<td>.076</td>
</tr>
<tr>
<td>E3</td>
<td>Model D, paths involving T1 PJ constrained across groups</td>
<td>81.295†</td>
<td>44</td>
<td>.939</td>
<td>.097 [.063, .129]</td>
<td>.074</td>
</tr>
<tr>
<td>E4</td>
<td>Model D, paths involving T1 PC constrained across groups</td>
<td>73.983†</td>
<td>44</td>
<td>.951</td>
<td>.087 [.050, .120]</td>
<td>.069</td>
</tr>
<tr>
<td>F'</td>
<td>All paths constrained except those involving T1 QUICK and T3 PJ</td>
<td>83.420</td>
<td>52</td>
<td>.949</td>
<td>.082 [.047, .113]</td>
<td>.075</td>
</tr>
<tr>
<td>F</td>
<td>Final model: all paths constrained except those involving T1 and T3 quick learning and PJ</td>
<td>67.559</td>
<td>44</td>
<td>.962</td>
<td>.077 [.036, .112]</td>
<td>.067</td>
</tr>
</tbody>
</table>

Note. Models marked with * denote a significant degradation in model fit from the more freed model. These changes were therefore excluded from the subsequent model. Models marked with † only marginally degraded fit when equality constraints were conducted separately. However, together the marginally poor fitting equality constraints did substantially degrade Model F’ compared to the final Model F.
References


121


Biography

Michael M. Barger was born on July 10, 1989 in Cincinnati, OH. He graduated summa cum laude with honors with a Bachelors of Science in psychology from the Ohio State University in June, 2011. He obtained his Masters of the Arts in psychology from Duke University in May, 2014. He has published book chapters in the *International Handbook of Emotions in Education* and the *Handbook of Epistemic Cognition*. He has also co-authored articles titled “Identification as gifted and implicit beliefs about intelligence: An examination of potential moderators” and “A pharmacology-based enrichment program for undergraduates promotes interest in science,” and has presented his research at conferences across the Unites States. He received a summer fellowship and multiple travel scholarships from the Duke Graduate School and won the Division C Poster Award at the American Educational Research Association (AERA) conference in 2014. He is a member of the Society for Researchers on Adolescence and AERA Division C, Learning and Instruction and served as the AERA Motivation in Learning Special Interest Group’s Graduate Student Co-Chair from 2013-2015.