STUDENT-GENERATED QUESTIONS DURING CHEMISTRY LECTURES: PATTERNS, SELF-APPRAISALS, AND RELATIONS WITH MOTIVATIONAL BELIEFS AND ACHIEVEMENT

A Dissertation
Submitted to the
Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by
Bradley W. Bergey
January 2014

Examining Committee Members:
Jennifer G. Cromley, Advisory Chair, Educational Psychology
Avi Kaplan, Educational Psychology
James P. Byrnes, Educational Psychology
Julie L. Booth, Educational Psychology
David Shernoff, Educational Psychology

ABSTRACT

Self-generated questions are a central mechanism for learning, yet students' questions are often infrequent during classroom instruction. As a result, little is known about the nature of student questioning during typical instructional contexts such as listening to a lecture, including the extent and nature of student-generated questions, how students evaluate their questions, and the relations among questions, motivations, and achievement. This study examined the questions undergraduate students (N = 103) generated during 8 lectures in an introductory chemistry course. Students recorded and appraised their question in daily question logs and reported lecture-specific self-efficacy beliefs. Self-efficacy, personal interest, goal orientations, and other motivational self-beliefs were measured before and after the unit. Primary analyses included testing path models, multiple regressions, and latent class analyses. Overall, results indicated that several characteristics of student questioning during lectures were significantly related to various motivations and achievement. Higher end-of-class self-efficacy was associated with fewer procedural questions and more questions that reflected smaller knowledge deficits. Lower exam scores were associated with questions reflecting broader knowledge deficits and students' appraisals that their questions had less value for others than for themselves. Individual goal orientations collectively and positively predicted question appraisals. The questions students generated and their relations with motivational variables and achievement are discussed in light of the learning task and academic context.

ACKNOWLEDGEMENTS AND DEDICATION

I am grateful to the many teachers, students, family members, and friends whose support and hard work made the completion of my dissertation possible: my advisor, Jennifer Cromley, for the many, many opportunities you have provided and for being so generous with your time, knowledge, and resources; Avi Kaplan, for sharing your contagious passion and for challenging and mentoring me in important ways. Jim Byrnes, Julie Booth, and Dave Shernoff, for modeling excellent scholarship and for sharing your time and energy as I develop my own; Erin Horvat, for your encouragement, grounded perspective, and invitations to work as colleagues; Diane Ketelhut and Catherine Shifter, for helping me further my analytic skills; Jim Bloxton, for your openness and support for this study; Briana Chang, for your patience and perseverance during the tedious coding process; Tony Perez, for being an ideal academic older brother; my colleagues and friends—Ting Dai, Melissa Karakus, Nora Pillard Reynolds, Brad Litchfield, Mark Snyder, Uma Natarajan, Ted Wills, Brian Miller, Dar Seth and Nhi Tran and many others—your help and support have been an important part of my education; and my parents, Will and Glenda Bergey, for providing critical support throughout my formal education, including this last stretch.

I dedicate my dissertation to my wife and daughter. To Tonya, your patience, understanding, encouragement, and sacrifice during the past four and a half years are a testament to our enduring partnership. And to Sadelle, I cannot wait until you start asking questions and I hope you never stop.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

An important mechanism of learning is the act of identifying what is not yet understood. When a learner generates an authentic, information-seeking question, she engages in a process that is at the heart of self-regulated and self-directed learning. Much of the extant literature on student-generated questions has followed one of two patterns: a *cognitive* focus on the *generation* of questions or a *motivational* focus on the *decision to ask* questions. These literatures have identified student-generated questions as an important aspect of adaptive cognitive and social processes, including the active construction of knowledge (Steinberg, 1987), self-regulation (Zimmerman, 2000), and strategic help seeking (Karabenick, 1998). Yet because students often withhold their questions from classroom discourse (Dillon, 1988; Ropp, 2001; West & Person, 1993), much remains unknown about this important aspect of student cognition in classroom environments. The current study examined (1) the questions students generated and the self-appraisals of questions as recorded in daily question logs in the authentic learning environment of an undergraduate chemistry lecture course; (2) the relations between question characteristics and motivational antecedents; and (3) the relations among students' questions and selected outcomes, including exam performance, perceptions during lectures, and asking behaviors in class.

In the remainder of this chapter, I describe the links between student-generated questions and adaptive motivational, cognitive, metacognitive, and social processes. I highlight areas of the literature on student-generated questions that have been under-theorized and under-researched. I provide a brief justification for the instructional context of the study before defining the phenomena and terms. I conclude with the research questions for the current study.

The Importance of Student-Generated Questions for Learning

Student-generated questions have been identified as an important aspect of motivational, cognitive, metacognitive, and social processes that support learning. From a motivation perspective, student-generated questions are important because they can energize and direct information-seeking behavior, cognitive processes, and social interactions that support learning. Questions are widely theorized to result from cognitive disequilibrium, such as when a learner experiences perplexity when becoming aware of a gap or clash in understanding (Dillon, 1998; Graesser & McMahen, 1993; Otero & Graesser, 2001; Piaget, 1952; van der Meij, 1994). The desire to restore equilibrium can motivate exploratory or information-seeking behavior (Berlyne, 1960; Chinn & Brewer, 1993; Festinger, 1962; Flammer, 1981; Loewenstein, 1994). Self-generated questions can provide an opportunity for students to exercise agency in the learning process (King, 1992; Scardamalia & Bereiter, 1992) by pursuing knowledge that holds individual relevance, and may thereby potentially support the basic psychological need for autonomy (Deci & Ryan, 2000). Furthermore, self-generated questions can deepen, reorganize or otherwise enrich a learner’s understanding. This may lead to the accumulation of knowledge and value for knowing in a domain, and thereby increase the likelihood of re-engaging with a topic (Renninger, 2000).

From a cognitive perspective, student-generated questions are important because they both constitute and accompany important cognitive processes. Question generation is a critical component in the active construction of knowledge (Bruner, 1960; Papert, 1980; Piaget, 1952; Wittrock, 1992), problem-solving (Zoller, 1987) and creativity (Sternberg, 1987; Albergaria-Almeida, 2011). The process of formulating questions has been proposed as a central mechanism for cognitive development, one which helps the learner target and integrate information at the moment when the learner is ready, engaged and motivated (Chin & Chia, 2004; Chouinard, 2007; Piaget, 1952). Self-generated questions have been theorized to support comprehension because they involve active processing of information, can support explanatory reasoning, and can guide learners to identify and make connections between key components and with prior knowledge (Chi, Bassok, Lewis, Reimann & Glaser, 1989; Collins, Brown, & Newman, 1990; Graesser, Baggett & Williams, 1996; Palincsar & Brown, 1984). Consequently, student-generated questions have been shown to support comprehension, memory, and retention of key ideas (Bugg & McDaniel, 2012; King, 1994).

From a metacognitive perspective, student-generated questions are important in self-regulating learning. The questions a learner generates can result from or trigger any of the principal self-regulatory phases of forethought, performance control, and evaluation (Zimmerman & Lebuhn, 2012). In the forethought phase, self-generated questions are part of organizing and directing learning (e.g., Scardamalia & Bereiter, 1992). During performance control, learners may use self-generated questions to monitor comprehension and identify and articulate uncertainties, discrepancies, and gaps in understanding (Palincsar & Brown, 1984; Rosenshine, Meister, & Chapman, 1996; Wong, 1985). In the evaluation phase, learners evaluate their questions in light of learning goals (Flammer, 1981) and plan actions accordingly.

From a social perspective, student-generated questions are important because they can lead to adaptive social interactions that support learning and development in classrooms. Asking a question that cannot be answered on one's own is an example of instrumental help seeking (Nelson Le-Gall, 1981; 1985), an adaptive form of self-regulation in academic settings (Newman, 1994). When asking a question in a classroom setting, a student influences the nature and direction of classroom discourse (Lemke, 1990; Nystrand, Wu, Gamoran, Zeiser, & Long, 2003) thereby engaging in a social act that may support his agency and autonomy (Ryan & Deci, 2000). Student-generated questions can facilitate discussion and debate in classrooms and direct and drive instruction in ways that support students' interests (Chin & Osborne, 2008). Peers' questions can scaffold metacognition by assisting learners as they monitor comprehension (Palincsar & Brown, 1984). In addition, when a student asks a question in class, classmates can learn both from witnessing a peer's reasoning process (Craig, Sullins, Witherspoon & Gholson, 2006) and from the instructional exchange that results (Mills, Danovitch, Grant, & Elashi, 2012). Student-generated questions can also provide important feedback about student learning to teachers by communicating misconceptions and gaps in understanding and levels engagement and interest (Dillon, 1998; van der Meij, 1994).

In sum, the questions students generate during learning are a potential resource for teaching and learning (Chin & Osborne, 2008). Student-generated questions can motivate information seeking and engagement, support comprehension and self-regulatory behaviors, and can lead to adaptive social interactions that support individual and group learning.

Gaps in the Student-Generated Questioning Literature

The literature on student-generated questions has several gaps that the current study aimed to address. These gaps include a description of student-generated questions as they process information in common instructional settings such as lectures, students' evaluations or self-appraisals of the questions they generate, and empirical studies examining the relationships among student-generated questions, motivational constructs, and achievement in classroom settings.

Descriptions of Student-Generated Questions during Science Lectures

As I describe in greater detail in Chapter 2, researchers have yet to document basic descriptive patterns about the frequency and qualities of student-generated questions during whole-group learning. The majority of empirical studies of students' questions during classroom learning have focused on the questions students ask aloud (e.g., Dillon 1988). Such research has described classrooms in which students' questions are scarce, a pattern that appears with surprisingly regularity across age groups, subjects and countries (Bergey, Horvat, & Cromley, 2012; Buseri, 1988; Dillon, 1988; Good, Slavings, Harel, & Emerson, 1987; Pedrosa de Jesus, Teixeira-Dias & Watts, 2003; Susskind, 1969; van der Meij, 1988). The methods employed by many of these studies examined only the act of question *asking* and thus little is known about the number and nature of the questions students generate *but choose not to ask* during classroom learning.

While this is a gap in the student-generated questioning literature in general, it is specifically the case for learning during undergraduate science lecture courses. The nature of the learning task and domain and the social context undoubtedly shape the questions students generate and ask. The current study examined the questions students generated and asked during the course of an 8-lecture unit in an introductory undergraduate chemistry course. I examined student-generated questions in this context for several reasons: introductory science lecture courses are gateway courses to careers in science, technology, engineering and math (STEM). The majority of undergraduate introductory science courses still rely on information dissemination models such as lectures (National Research Council, 2003). Achievement in such courses is an important factor in persistence in STEM majors (Maltese & Tai, 2011). Thus, understanding how students learn course material in these settings may inform best practices for instruction in similar contexts. In addition, introductory STEM courses have large-enrollments and therefore provided the opportunity to maximize the number of participants in a single instructional context.

Self-Appraisals of Student-Generated Questions

Several theories of question asking and help seeking (described in detail in Chapter 2) suggest that students evaluate their questions in light of learning goals and the social context. This component of questioning, however, is not well understood. Much of the literature on student questioning has examined barriers to question asking (Ryan, Pintrich, & Midgley, 2001). If a question is asked only when the benefits outweigh the costs, as many researchers assume (van der Meij, 1990; Karabenick, 1994), how students appraise their questions may be an important aspect of decision making regarding their questions. As I describe in greater detail in Chapter 2, the current study examined two values that may be relevant to students' appraisals of their questions; both are grounded in theoretical and empirical research. First, questions may be perceived to have great or little value for the individual in light of his or her learning goals (Flammer, 1981; Scardamalia & Bereiter, 1992). Second, a student may evaluate a question in terms of its value for others. The second self-appraisal builds on a central finding from my prior research on student questioning: when deciding to ask or withhold a questions, a salient consideration for students was the perceived usefulness of their question for others (Bergey et al., 2012). A better understanding of how students appraise their questions may illuminate the complex decision making process that precedes a question being asked (Flammer 1981).

Relations Among Student-Generated Questions, Motivation Constructs and Achievement

Another gap in the literature on student-generated questions is the lack of empirical studies that have examined the relationships among students' questions, motivational constructs, and achievement in classroom environments. Anderman and Wolters (2006) concluded their review of motivational literature in educational psychology noting that a particularly ripe area of research is the relationship between motivational constructs and specific aspects of information processing. Preliminary research in this vein suggests that different motivation orientations may be associated with different cognitive processes (e.g., Huber, Beckmann & Herrmann, 2004). In the current study, I examined the relations among several motivational self-beliefs—situational and personal interest, personal goal orientations, and self-efficacy—and one contextual motivational construct: perceived teacher emphasis on mastery goals. As I describe in Chapter 2, I selected these motivational constructs because of their theoretical connections to question generation and asking and because of their prominence in motivational literature. An understanding of the connections between these constructs and questioning behavior has the potential to extend the existing literature on the process of generating questions, which has previously focused nearly exclusively on cognitive variables.

As I noted above, the act of generating and pursuing the answer to one's questions has been associated with adaptive cognitive, metacognitive, and social processes. Yet, the relationship between students' questions and achievement may not necessarily be positive, since questions may also indicate comprehension difficulties. Findings from interventional studies have indicated that self-questioning can be an effective learning strategy that leads to achievement (Rosenshine, Meister, & Chapman, 1996), yet prior research has not examined the relations among students' questions and achievement with more naturalistic methods.

Defining the Phenomenon and Terms

Students' questions can serve many functions (Dillon, 1982; van Zee, 2000). Graesser, Person, and Huber (1992) identified four distinct functions of questions in educational settings: correcting a knowledge deficit (e.g., a glitch in an explanation of an event), monitoring common ground (e.g., a questioner's assessment of the answerer's knowledge), social action coordination (e.g., asking permission), and controlling conversation and attention (e.g., rhetorical question). This study focuses on questions that fall in the first category: correcting a knowledge deficit. Wisher & Graesser (2007), van der Meij (1987) and others have termed such questions *sincere* or *authentic information-seeking questions.* Such questions are born of an awareness of not knowing and a desire to learn. Van der Meij (1997; as cited in Graesser & Person, 1994) articulated several criteria for questions to be genuinely information seeking:

> (a) the questioner does not know the information asked for in the question, (b) the questioner believes that the presuppositions of the question are true, (c) the questioner believes that an answer exists, (d) the questioner wants to know the answer, (e) the questioner can assess whether a reply constitutes an answer, (f) the questioner believes that the answerer knows the answer, (g) the questioner believes that the answerer will not give an answer in absence of the question, (h) the questioner believes the answerer will supply the answer, and (i) the questioner

> believes that the benefits of asking the question exceed the costs. A question fails to be an information-seeking question to the extent that these assumptions are violated (p. 109).

While sincere-information seeking questions can occur in the minds of students in any learning context (e.g., while reading, writing a paper, completing an assignment, studying for an exam) and can be communicated to others in many ways (e.g., email, study groups, office hours), the current study focused on the questions that students generated or asked during classroom lectures. As I described above, lectures constitute a significant portion of how undergraduate are exposed to information, a pattern that is especially true of first- and second-year students in STEM fields.

This study addressed various steps in the process of generating and asking questions, which I describe in detail in Chapter 2. For purposes of clarity, I use the verbs *to ask* and *to pose* to refer to the act of verbalizing a question in a public setting (e.g., asking a teacher or a classmate). By contrast, I use the verbs *to generate, to formulate, to conceive,* and *to have* to refer to the process of articulating a question in one's mind and/or recording a question on paper. Such questions may or may not be publically asked.

I use the terms *self-appraisal of questions* to refer to the process of self-reported evaluations of self-generated questions (see Chapter 2 and 3). For expository purposes, I use the term *motivational self-beliefs* to refer to the motivational constructs of *interest, personal goal orientations* and *self-efficacy* to avoid long lists of constructs. For the same reason, I use the term *help seeking beliefs* to refer to the constructs of *executive help* *seeking, instrumental help seeking, help seeking threat, help seeking avoidance, and preference for a formal help seeking target.*

I use the terms question *qualities* and *characteristics* to refer to various features of students' questions. As I describe in Chapter 2 and 3, *qualities* and *characteristics* include the number and type questions, question syntax, and whether questions were resolved during class meetings.

The Current Study

I examined the questions students generated and asked during lectures over the course of a 3-week chemistry unit, and examined the relations among these questions and their characteristics and personal and situational interest, goal orientations, self-efficacy, perceived teacher emphasis on mastery goals, and achievement. As I detail in Chapter 3, I collected student-generated questions via a daily question log and classroom observations. Motivational constructs were measured before, during, and after the target unit using self-report questionnaires, and achievement was measured by course exams immediately before and after the target unit. The present study examined 4 primary research questions:

(RQ1) What are the relations among the quantity, qualities, and self-appraisals of student-generated questions? RQ1 aims to describe students' questions, their qualities, and self-appraisals of questions including possible patterns or profiles among question characteristics. This question contributes to the literature on student questioning by describing the contents and inter-relations of question characteristics, aspects that are often kept private in classroom settings and rarely researched.

(RQ2) What are the relations among motivational self-beliefs and perceived teacher emphasis on mastery goals and the quantity, qualities, and self-appraisals of student-generated questions? This research question addresses the gap in the literature regarding students' questions and self-appraisals as they relate to motivation constructs. As I describe in greater detail in Chapter 2, little is known about the relations among the characteristics of students' questions, self-appraisals of questions, and a range of motivational constructs, despite theorized connections to questioning behaviors.

(RQ3a) Can exam scores be predicted by the quantity and qualities of student-generated questions? (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and Exam 2 scores? RQ3a and RQ3b examine the relations among question-generation processes and achievement. As I describe in greater detail in Chapter 2, questions can indicate both cognitive processes as well as knowledge deficits, and therefore relations between student-generated questions and achievement are unclear, including possible indirect relations via perceptions of self-efficacy and comprehension.

(RQ4) Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, help seeking beliefs, achievement, and the quantity, qualities, and self-appraisals of generated questions? This question seeks to bridge the gap between literatures on question generation and question asking. RQ4 examines how features and variables associated with question generation may be similar or different from question asking.

CHAPTER 2

LITERATURE REVIEW

The nature, function, and importance of questions in human thought have attracted the attention of scholars from diverse disciplines for millennia (Dillon, 1982; Graesser, 1985). Within the fields of education and psychology, researchers have historically focused on the questions that teachers pose to students to direct and support their learning (e.g., Bloom, 1956; Carlsen, 1991; Wilen & Clegg, 1986). Over the last three decades, researchers have increasingly turned their attention toward students' questions, investigating the role self-generated questions play in the learning (Beck, 1997; Collins, Brown, & Newman, 1989; Edelson, Gordon, & Pea, 1999; Graesser & Person, 1994; King, 1992; Palincsar & Brown, 1984; Piaget, 1952; Scardamalia & Bereiter, 1992; Zimmerman, 1989). As I described in Chapter 1, researchers have found that the act of generating and asking questions are important components in a broad set of cognitive and social processes that support learning. Consequently, the phenomenon of student questioning has received the attention of researchers from a variety of domains, including research on learning during computer and human tutoring (e.g., Graesser, McMahen, & Johnson, 1994), reading comprehension (e.g., Taboada & Guthrie, 2006), help seeking in classroom environments (e.g., Karabenick & Newman, 2006), inquiry-based learning (e.g., Chin & Osborne, 2006), and discourse analysis (e.g., Nystrand, Wu, Gamoran, Zeiser & Long, 2003).

Despite the increasing and diverse interest in student-generated questions and its reach into diverse literatures, much remains unknown about this important aspect of learning. In particular, little is known about the questions students generate as they process instruction in classroom environments and how students evaluate the questions they generate. Furthermore, little is known about the relations among student-generated questions, motivations, and achievement.

In this chapter, I begin by describing the theorized components involved in the process of generating and asking questions. In doing so, I highlight the under-theorized components of self-editing and evaluating questions. I describe a theoretical framework that locates questioning processes at the intersection of individual characteristics, the learning tasks, and the social context. I follow with a review of literature on the frequency and types of questioning in classrooms as well as the relationship between students' questions, motivation and achievement. I conclude this chapter by describing the research questions and hypotheses of the study.

Theoretical Perspectives on the Process of Generating and Asking Questions

When investigating the phenomenon of students' questions, researchers have brought to bear cognitive, social-cognitive, and social-cultural theoretical perspectives. Much early work on student questioning and on-going research on question generation during reading has largely assumed a cognitive perspective (e.g., Mikaya & Norman, 1979; Otero & Graesser, 2001; Taboada & Guthrie, 2006). Questioning research from a cognitive perspective often draws on theories of active processing, schema theory, and metacognition (Yang, 2006). Such research has focused on cognitive characteristics of individuals (e.g., prior knowledge, metacognitive skills) and/or text characteristic (e.g., coherence) as they relate to the questions learners generate. It is important to note that studies with a cognitive perspective have mainly focused on the ability or propensity of an individual to *generate* questions—rather than *ask* them in public settings. Such studies are typically conducted in laboratory settings (e.g., Baker, 1979; Graesser, Lu, Olde, Cooper-Pye, & Whitten, 2005; Mikaya & Norman, 1979; Olney, Graesser, & Person, 2012). Studies with a cognitive focus rarely investigate the role of motivation. As a result, nearly all studies that have examined the questions that students *generate* (versus ask) have ignored the role of motivational variables.

Other researchers have focused on question asking in classrooms and have brought to bear a social-cognitive perspective (e.g., Karabenick & Sharma, 1994; Ryan & Pintrich, 1997; van der Meij, 1990). Such research assumes that questions involve both cognitions about learning tasks (e.g., evaluation of need for help) as well as cognitions and motivations in a social context (e.g., the role of social goals). Studies with a socio-cognitive perspective have investigated how question asking is influenced—and often thwarted—by social and psychosocial factors such as key socializers (e.g., Newman, 1990) or perceptions of one's self, others, and the environment (e.g., Ryan, Pintrich & Midgley, 2001). Such studies have focused on the factors influencing the social interactions that accompany *asking* questions—as opposed to *generating* questions. Thus, the majority of what is known about the role of motivation in questioning processes has framed student-generated questions as a form of help seeking with a focus on factors that inhibit students from seeking the help they need. As I discuss in greater detail below, social-cognitive motivational variables, such as interest, goal orientation, and self-efficacy have theoretical relations with the questions students generate as well as their decisions to ask questions or not.

Still other researchers have approached student-generated questions with a social-cultural or socio-linguistic lens. From this perspective, questions are viewed as dialectally co-constructed among individuals and shaped by the social, cultural, linguistic, and historical contexts (e.g., Carlsen, 1991; Yang, 2006). From a social-cultural perspective, a student's question echoes the preceding discourse and shapes proceeding interactions (Wertsch, 1991). For example Nystrand and colleagues (2003) examined how the frequency of students' questions that were asked in class triggered additional students' questions, leading to dialogic spells.

Each theoretical position has illuminated different aspects of the phenomenon of student questioning. The present study was informed by cognitive and socio-cultural perspectives but is framed primarily from a social-cognitive perspective. From this lens, the process of generating a question, as with cognition and metacognition more generally, occurs at the intersection of three overlapping and interacting influences: (a) individual characteristics, (b) the nature of the learning task and domain, and (c) the social-cultural context. Figure 2.1 illustrates this conceptual framework.

As I describe over the course of this chapter, several individual characteristics may be relevant to questioning processes, including prior knowledge and achievement, motivational self-perceptions, and perceptions of the environment. The primary focus of the present study was how individual characteristics related to questioning. The process of generating questions is also shaped by the specific nature of the learning task, the academic domain, and the socio-cultural context. A student is likely to exhibit different questioning behaviors while listening to a chemistry lecture compared to when he or she is conducting a chemistry experiment in a lab for the same class. In the same vein, differences in the nature of knowledge and knowing in a given domain—chemistry compared to literature, for example—will likely shape the kinds of questions that student are likely generate. At the same time, contextual factors shape questioning behaviors, including the immediate instructional and social context of the classroom as well as broader institutional and societal influences.. Furthermore, individual characteristics, the nature of the learning task and domain, and the social-cultural context are themselves inextricable intertwined. For example, an individual's motivational self-beliefs such as self-efficacy reflect perceptions of and experiences with the learning tasks he or she is faced with (e.g., the nature of the material, type of assessment) and the context in which the learning takes place (e.g., opportunities for asking questions in- and out-of-class, social norms). Therefore, while this study focuses on the relations among individual characteristics and questioning processes, these processes and relations must be understood in light of their deep connectedness to the learning task, domain, and context.

As Figure 2.1 illustrates, questioning processes have a reciprocal relationship with individual characteristics, the nature of the task, and the social context, and these contextual factors influence and are influenced by questioning processes. For example, students' questions may lead to behaviors and experiences that change students' individual characteristics, such as knowledge and motivations. Questions that are made public may be taken up by a teacher, who then modifies the learning task. Students decisions to ask or withhold questions shape behavioral norms for the context and influence the behaviors of others, such as peers or the teacher, which in turn, can influence later questioning processes.

I now turn to examine the theorized cognitive components and processes involved question generation.

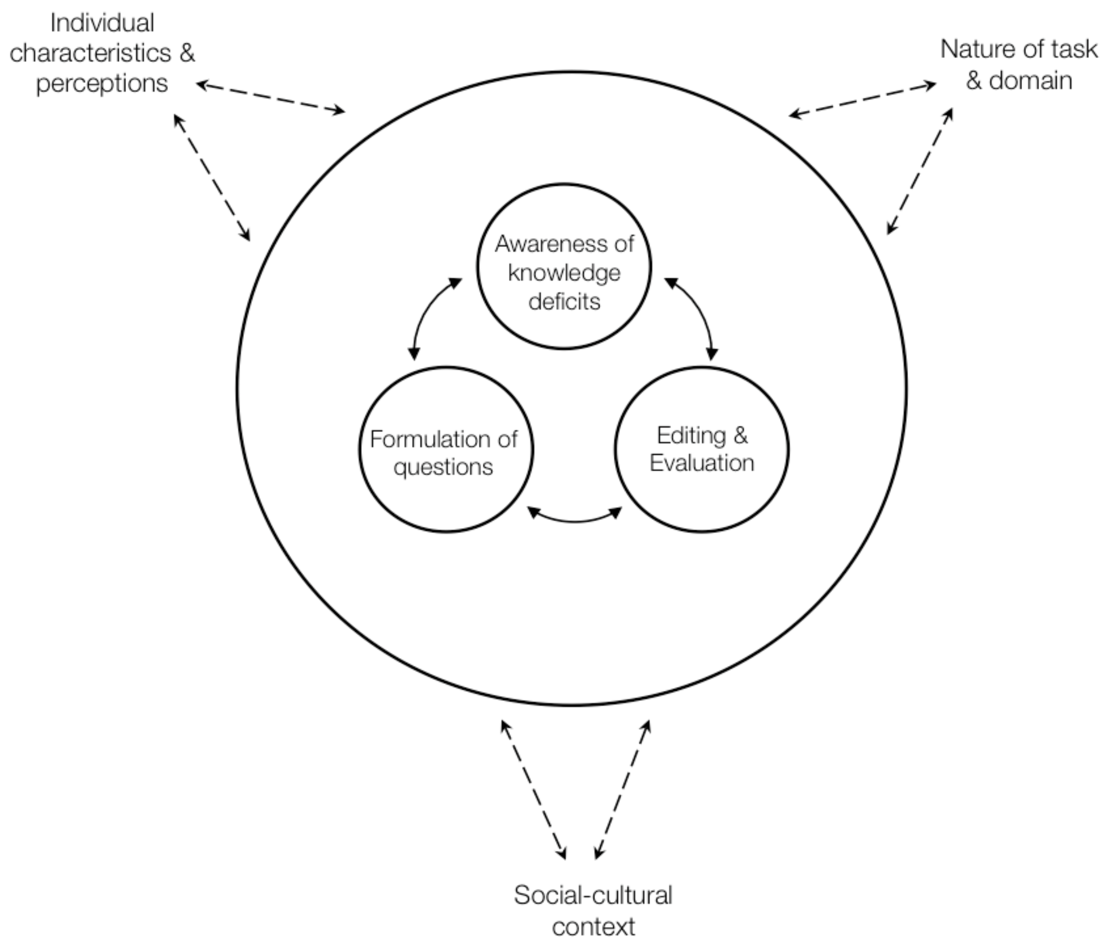

Figure 2.1. Contextual Model of Question Generation and Evaluation.

Components and Processes of Questioning Generation

What are the cognitive components involved in generating question? Researchers have described the cognitive processes involved in generating and asking questions in different way (Dillon, 1988, 1990, 1998; Flammer, 1981; Nelson Le-Gall, 1981; Otero & Graesser, 2001; van der Meij, 1990, 1994; cf. Otero & Graesser, 2001). These various descriptions and models differ in what aspects of questioning are highlighted, how components and phases are organized, and the extent to which cognitions about social and contextual factors are explicitly addressed. In his early work on questioning, Graesser (e.g., Graesser & McMahen, 1993) identified and found empirical support for three components of information-seeking questioning: (1) an awareness of a knowledge deficit, (2) verbal formulation of a question, and (3) social editing. These components are represented in some form in most descriptions of the steps or decisions in question generation and asking, and thus provide a parsimonious model of key cognitive elements in the process of generating questions. I expand Graesser and McMahen's notion of social editing—referring to it as Evaluation and Editing Processes—to include a broader range of cognitions that shape this component. In doing so, I highlight ways in which this component has been under-theorized and under-researched. In the section below, I present these components in a sequence in which they may occur, although as I discuss later, the components do not necessarily follow the order in which I present them.

Component 1: Awareness of a Knowledge Deficit and its Triggers

There is strong agreement among researchers that sincere information-seeking questions emerge from an awareness of a knowledge deficit (Dillon, 1998; Graesser & McMahen, 1993; Flammer, 1981; Otero & Graesser, 2001; Piaget, 1952; van der Meij, 1990). This awareness is a metacognitive state of consciously knowing that one does not know.

This state of being aware of a knowledge deficit can be triggered in many ways. Graesser, Person, and Huber (1992) identified five categories of cognitive events that trigger awareness of knowledge deficits. First, awareness of a knowledge deficit may be triggered while trying to comprehend a written or spoken message. For example, a student listening to a lecture on assessment instruments may become aware she does not understand what the instructor means by the term *standardized*. Second, a learner observing a phenomenon may encounter a glitch in his or her ability to explain the event. For example, after witnessing a lightning strike and hearing thunder, an individual becomes aware that he cannot explain *why* lightning causes thunder. Third, a learner may become aware of a knowledge deficit when he or she encounters an obstacle to a plan or difficulty solving a problem. For example, a student tasked with generating descriptives statistics for only female participants may think, "How do I create analyses that compare groups?" At the same time, she may wonder, "Why I am asked to look at gender differences anyway?" Fourth, a contradiction between two or more knowledge structures can trigger awareness of a knowledge deficit. For example, a student's prior knowledge about temperature may conflict with his teacher's insistence that from a scientific perspective there is no such thing as cold, just lack of heat. Finally, an awareness of a knowledge deficit can arise from deciding between equally attractive options. For example, a student believes that an ANOVA and a multiple regression could answer a research question equally well may ask his advisor, "Which is better?"

Importantly, van der Meij (1990) noted that the events that trigger awareness of a knowledge deficit may come from internal or external sources. Internal triggers refer to mental events that expose the gap in knowledge. For example, one might become aware of a knowledge gap while attempting to explain something to herself. External triggers refer to events outside of the individual that prompt the awareness a knowledge deficit. For example, a student's question in class may make others aware that they also do not understand something, or a teacher may task students with writing down their questions, which triggers the search for a knowledge deficit. Regardless of how it is triggered, at the heart of a sincere information-seeking question is an awareness of a knowledge deficit.

Before turning to the second component of question generation, it is worth noting that the terms *deficit* and *gap* convey the somewhat negative notion that a question is a deficiency that requires remediation. Alternatively, questions might be framed as growth potentials or opportunities to deepen, strengthen, expand, or transform understandings, conceptions that are more in line with research on questions that arise from interest, curiosity, and wonderment (Renninger, 2000; Scardamalia & Bereiter, 1992; Texiera-Dias et al., 2003). Nevertheless, I use the terms *knowledge deficit* and *gap in understanding* in the current study to be consistent with much of the psychological literature on question generation.

Component 2: Verbal Formulation of the Question

A second component in generating a question is expressing the knowledge deficit a verbal form (Dillon, 1988; Graesser & McMahen, 1993; van der Meij, 1994). Sometimes awareness of a knowledge gap and the act of verbally expressing it occur simultaneously. In other instances, a question might be formulated and reformulated in one's mind, as one clarifies the exact nature of the knowledge deficit or considers how to best express the question to another person. In formulating a question, a student selects the words and structures to express the knowledge deficit. Because the act of articulating a question is, in part, a linguistic task it interacts with developmental stages and verbal abilities (Otero & Graesser, 2001; van der Meij, 1994).

Component 3: Evaluation and Editing Processes

When a learner becomes aware of a knowledge deficit, the deficit (or the resulting question) can become an object for editing and evaluation (Graesser & McMahen, 1993). Research examining student questioning, including literature on the related act of help seeking, have described the process of editing or evaluating whether one's lack of understanding is important and worth pursuing (e.g., Dillon, 1998; Graesser & McMahen; Karabenick & Dembo, 2011; Nelson Le-Gall, 1981, 1985; Newman, 1994, 1998; van der Meij, 1994). When a student becomes aware of a knowledge deficit, he or she may engage in metacognitive suppression strategies that that resolve the deficit itself, such as inferring the answer or using fix-up strategies (Baker, 1979) that reconcile, ignore, or otherwise resolve contradictions, discrepancies, or gaps in understanding (Chinn & Brewer, 1993; Otero & Campanario, 1997; Otero & Graesser, 2001). In the case that the knowledge deficit is unresolved, an individual may evaluate whether the question is important in light of learning goals (Flammer, 1981), appropriate in light of the context (Bergey et al., 2012; Ryan, Pintrich & Midgley, 2001), and whether the potential benefits of pursing an answer to the question outweigh the costs (Dillon; van der Meij; Karabenick, 1994).

This editing and evaluation process may lead to many different outcomes, including, in many cases, keeping one's question private. Presumably, many questions that are raised in the minds of learners are never asked (Dillon, 1988; Flammer, 1981; van der Meij, 1994). Dillon conjectured that as much as

> 95 percent of the questions that we have in mind to ask we never go on to utter…we may think the better of it and follow one of the numerous other paths

> available. These include keeping quiet and giving off that we know and understand (p. 20).

While awareness of a knowledge deficit may be a necessary condition for asking a sincere information-seeking question, it is not a sufficient condition for question asking. A learner's *awareness* of a gap in information must become a *desire* for information if it is likely to motivate question asking (Flammer, 1981). In their componential model, Dillon (1988) and van der Meij (1990; 1994) posited that perplexity was the mental state that motivated information seeking. The authors described perplexity as both cognitive and emotional engagement in response to cognitive disequilibrium regarding the knowledge deficit. Yet awareness of a knowledge deficit may not necessarily be accompanied by perplexity. Readers can probably recall a time when they were aware of not knowing something but were not perplexed, perhaps deeming the lack of understanding as unimportant or peripheral to the task at hand.

What determines whether a question in the mind of a learner will go on to be asked? According to Flammer (1981), this is the most difficult and complex issue of question-asking theory: "the description and prediction of the development of an actual need for specific information and the process leading up to the utterance of a question" (p. 414).

Much of the research on students' decision making regarding their questions has focused on costs. Researchers have identified many individual and contextual factors that inhibit students from asking their questions during class. These factors include individual variables such as threats to academic competence (Ryan & Pintrich, 1997) and social competence (Ryan, Gheen, & Midgley, 1998) and contextual factors, such as perceptions about peers and teachers (Karabenick & Sharma, 1994), in- and out-of-class asking opportunities (Bergey et al., 2012), and teachers' response to questions (Nystrand et al., 2003; Ropp, 2002). A learner is more likely to withhold her question when one or several of these factors are salient. For example, in a large lecture hall, a student may have very few opportunities to ask a question, even when he deems his question important for his learning. In situations where the costs are high or when opportunities to ask questions are few, students may choose to tolerate not asking a question that aims to remediate a knowledge deficit or to postpone it to a later time (Flammer, 1981).

Van der Meij (1990) presumed that a question posed in a public setting reveals what the asker thinks is *worth asking*. While potential costs may act to inhibit students from asking their questions, the extent to which a student perceived *value* in his or her question presumably acts to promote the pursuit of the answer. More broadly, values for a task are an important motivational resource that shape decision making and choice (Eccles & Wigfield, 1995). Students are more likely to engage in tasks that they deem interesting, useful or important. For example, it is logical to presume that, all things held constant, a student is more likely to pursue a question—asking someone or otherwise seeking information—that is deemed very useful compared to a question that is perceived as trivial or peripheral to learning goals.

Flammer (1981) theorized that the perceived utility of one's question was critical in determining whether awareness of a knowledge deficit translated into desire for information. Whether awareness of a knowledge deficit becomes a desire for knowledge depends in large part on the extent to which the knowledge deficit impedes the learners' goals and subgoals, both conscious and unconscious. Flammer theorized that the utility value of one's question in serving one's goals might range from being substantial to irrelevant. A knowledge deficit may constitute an impasse that must be resolved for achieving one's goal or it may be perceived to have only a minor effect. Flammer articulated some of the notable factors that may influence perceived utility, including

> the importance of the question, whether an a priori likelihood of some possible answer already exists, the presumed ability of the partner [i.e., answerer] to answer, the likelihood of understanding an answer from a given partner, and the likelihood that information would be available later without a question (p. 416).

Researchers examining the role of values in decision-making have traditionally examined value *for the individual* (Wigfield & Eccles, 2000). That is, an individual evaluates a task in terms of values to him- or herself. However, it is important to bear in mind that values may be assigned not only in terms of benefits and costs to the individual. Concerns for and identification with others may shape an individual's values and influence the actions they choose to take. Students may take into account the perceived value of their question for others' learning because asking a question during class makes demands on others. When a student asks a question during class they exert an influence on their environment that is different than other forms of participation, such as responding to a teacher's question. A student's sincere information-seeking question solicits a response, which in turn makes a demand on others, such as the teacher, to respond to the question. In asking a question, a student can influence the direction of classroom discourse and in doing so make a demand on instructional time, a valuable resource shared by the class. Even when teachers ask for students' questions, a student who takes up such a request makes a claim on the allotted time for questions.

The importance of a student's perception that his or her question has value for others builds on a central finding of my prior research. Bergey et al. (2012) interviewed 18 graduate students after class meetings of an introductory statistics course. Students were asked to recollect their experience regarding their decision to withhold questions. Many students reported that they were likely to choose to withhold a question when they viewed the question as having limited value for others. Students reported being concerned that their question would unduly take up class time for what may be an individual knowledge deficit. A question's value for others was one of the most salient factors students considered as they explained their decision-making process, often trumping their own interest in the question or its perceived value for themselves as an individual learner.

It is important to note that while students' appraisals of their questions may be central to decision making about whether a question is asked or not, evaluations may also shed light on other aspects of learners. For example, different purposes that drive achievement (i.e., achievement goal orientations) may influence how or why students value questions for themselves or others. Similarly, appraisals (either alone or in combination) may illuminate how students view their knowledge gaps relative to how they perceive others are learning the same material.

The Sequence of Components of Questioning

In the preceding section, I described the components theorized to be central to the process of generating and evaluating questions. These common components include an awareness of a knowledge deficit, verbal articulation of questions, and processes of editing and evaluating questions. It may seem logical that these components proceed one after another: An individual experiences an awareness of a knowledge deficit, which leads to the verbal formulation of a question, which is then evaluated or otherwise resolved, leading to a decision to ask or withhold the question. Indeed, we certainly can and often do experience these components of the questioning process in this order. It is important to note, however, that the components may function in an interactional—rather than a sequential—manner (Graesser & McMahen, 1993). For example, one may decide to ask a question and consequently seek a knowledge deficit, or after formulating an initial question, one might refine it and deem the new question as being more valuable.

Graesser and McMahen (1993) provided some preliminary evidence for the non-sequential nature of the components in the questioning process. The authors conducted a sequence of five studies—three of which I focus on here—with undergraduate psychology students, who were assigned to reading conditions in which a set of word problems or short texts contained varying degrees of deleted, conflicting or irrelevant information. In one experiment (A), students were asked to generate as many questions as they could after reading and again while working through the set of word problems. In another experiment (B), a new group of participants was encouraged to ask questions to an assistant in a private setting in an adjacent room while solving the same set of word problems. In third experiment (C), a new group of participants was asked to make an anomaly detection rating but not generate questions for the same set of word problems. The authors observed that students in the anomaly conditions in experiment A were more likely to generate a question that addressed an anomaly than their counterparts in Experiment C were likely to detect an anomaly. The authors interpreted this as evidence that the task of question asking can trigger awareness of anomalous information and knowledge deficits. In addition, the authors observed that when students were required to ask their questions in a social setting (Experiment B)—even one in which some social barriers were minimized—students were less likely to ask questions compared with those who wrote their questions, but their questions were significantly more likely to be related to the anomalies. The authors interpreted this as evidence that questions that survive the social editing process may be more specific to task-related learning. It is important to note that the series of experiments were conducted with manipulated—rather than naturally occurring—social conditions and texts, and therefore may have limited external validity for classroom environments. Nevertheless, the studies provide empirical evidence for the somewhat counterintuitive notion that components of questioning may not follow a set sequence.

In sum, several theories of question asking suggest that students evaluate their questions in light of learning goals. However, this component in the questioning process is under-researched. Researchers have assumed that the benefits of asking a question must outweigh the costs, yet much more is known about the perceived barriers to question asking than about how students appraise the questions they generate. The perceived value of a question to further an individual's understanding may be an important value in offsetting costs. In addition, given the extent to which students' questions can shape the classroom environment and make demands on shared resources, it may be important to attend to how students evaluate the worth of their questions for others. Thus, theoretical and empirical literature—including my prior research—suggests both a question's value for the individual and value for others may be important self-appraisals.

Next, I review literature on the type and frequency of questions in classrooms, followed by a review of individual characteristic variables associated with question generation and asking.

Frequency and Types of Student-Generated Questions in Classrooms

In his seminal work on student questioning and teaching, Dillon (1988) opined, "Student questions enjoy generous place in educational theory but occupy small room in educational practice" (p. 8). Despite the potential advantages of students' questions to support learning, classrooms practices often foster student passivity (Good et al., 1987; Wood & Wood, 1988). The infrequence of student questioning in classrooms appears to be a surprisingly universal phenomenon. In countries such as the former Czechoslovakia, Germany, Netherlands, Nigeria and the United States, researchers have found that students' questions constitute a small fraction of classroom discourse (Buseri, 1988; Good, Slavings, Harel, & Emerson, 1987; Susskind, 1969; van der Meij, 1988). Student-generated questions are found to be relatively rare in classrooms across the age spectrum, from elementary and secondary classrooms (e.g., Good et al., 1997), university (e.g., Pedrosa de Jesus et al., 2003), and graduate classrooms (e.g., Bergey et al., 2012).

For example, Dillon (1988) observed 27 high school social studies classrooms in six schools, randomly sampling a 10-minute section of classroom discourse. He observed a total of 11 student-generated questions. Dillon calculated that while teacher-generated questions accounted for approximately 60% of teacher talk, student-generated questions accounted for only 1% of student talk. Based on Dillon's results, Graesser and Person (1994) noted that on average an individual student asks .11 questions per hour, contrasted with a dramatically high mean of 69 teacher-generated questions during the same amount of time.

Good et al. (1987) examined developmental patterns of the frequency of student questioning. The authors purposively selected high- med- and low-achieving students in classrooms from kindergarten and grades 1, 3, 6, 7, 9 and 12. The authors sampled 12 50-minute periods in multiple classrooms in each of the grade levels, recording the frequency of students' questions by type. Collapsing across achievement groups, the authors observed that older students asked more conceptual and explanatory questions than younger students. For example, a 9th grader student was 10 times more likely to ask an explanatory question than a kindergartner. An interesting developmental picture emerged with regard to the frequency of student questioning across grade levels and by achievement groups. In kindergarten, low-achieving students asked more academic questions (e.g., relating to explanations, conceptual information) than did medium- and high-achieving students. By contrast, for most secondary grade levels, low-achieving students asked fewer such questions compared with med- and high-achieving students. Good et al. concluded low-achieving students become increasing intellectually passive via a vicious cycle in which low-achievement is perpetuated and exacerbated by passivity. It is important to note that the standard deviations around the mean scores were very high; in many cases the standard deviation far exceeded the mean. For example, med-achieving 12th grade students asked a mean of 1.31 academic questions per 50-minute observation, which had a standard deviation of 2.82. These high standard deviations reflected substantial variation of question between observations. Good et al. interpreted this variation to reflect the important role of context, especially teachers, in shaping the frequency of students' academic questions. It may also have suggested high variability between individuals within an achievement level.

Researchers have found low frequencies of student questions at the university level as well. West and Pearson (1993) examined the frequency and type of undergraduate's questions by sampling one hour of classroom discourse in 30 medium-sized (25-40 student) communication classes. In these 30 hours of instruction, students asked a total of 108 questions, an average of 3.6 questions per hour per class. Nearly a third of these questions addressed logistical issues. The authors characterized the majority of the remaining questions as general knowledge and clarification questions. In a prior study of 15 communications classrooms using similar sampling methods, the authors (Pearson & West, 1991) found similar rates of questioning: approximately 3 student-generated questions each hour per class.

Researchers who have investigated the frequency patterns of questions over time for a group of learners have found substantial variability (Pedrosa de Jesus et al., 1997; Teixeira-Dias, Pedrosa de Jesus, Neri de Souza, and Watt, 2005). For example, Bergey and Cromley (2012) recorded all of the questions asked by 24 graduate students during a 1-semester introductory statistics course, collecting students' questions across multiple contexts, including during class, office hours and via email. The authors found the frequency of questions to be highly variable across time with question frequency dramatically spiking before exams, a pattern that appeared for questions asked both in- and out-of-class.

In sum, the number of students' questions that make it into public discourse in classrooms is regrettably rare, yet little is known about the extent to which students *generate* questions that are ultimately kept private. Presumably, students withhold a great majority of their questions (Dillon, 1988). The many psychological barriers (Karabenick, 1994; Ropp, 2001; Ryan, Pintrich, & Midgley, 2001) and social pressures (Bergey et al., 2012; Ropp, 2002; Ryan, Pintrich, & Midgley) that act to encourage students to keep their questions to themselves presents a challenge to researchers who seek to understand the extent to which students generate questions in authentic, public settings like classrooms.

A European team of researchers conducted a series of investigations of questioning patterns with undergraduate chemistry students, using research methods that provided students with both public and private means of asking questions. Employing longitudinal descriptive methods, Pedrosa de Jesus et al. (1997) examined the questions students generated during a foundational undergraduate chemistry course. The authors followed a pilot group of 32 Portuguese students, who were part of a larger chemistry course, which included multiple instructional components: large-group lectures (approx. 130 students) and smaller-group seminar and laboratory sections (approx. 16 and 32 students, respectively). Students in the pilot group were incentivized to generate questions (e.g., earning points towards final grade) and could pose questions through four question collection mechanisms: (1) an intranet software system, which students could access on- and off-campus (e.g., a digital drop box) (2) e-mail, (3) a physical drop box to collect questions on slips of paper, and (4) project notebooks. In addition to other course components, students in the pilot group (1) met weekly with the lecturer to discuss the questions they had generated from the week and (2) were interviewed about their experience with the intervention. The entire group of 32 students generated on average about 3 questions per week, a rate that remained relatively constant over the semester. Occasionally, the number of questions spiked, the reasons for which are discussed later. The rate of student-generated questions was surprisingly low given the unique circumstances of the study: (1) students were incentivized to generate questions, (2) they received individualized attention and feedback from the lecturer regarding the contents of their questions, and unlike the majority of the studies of question frequency reviewed above, (3) students could ask questions privately. Despite these presumably question-friendly practices, the rate of question generation remained consistent with observational studies of question asking.

Teixeira-Dias et al. (2005) scaled up the pilot study described above, examining the questioning patterns of 100 Portuguese undergraduate students during 1-semester foundational chemistry course. The course supplemented traditional lectures with several supplementary course components designed to foster motivation and interest in the domain, including mini-projects, topical lectures derived from students' questions, practical laboratory sessions, and case-study seminars. As with the Pedrosa de Jesus et al. (1997) study described above, students could communicate questions in several ways: asking them during learning sessions, via email or digital drop box, leaving a written question in a collection box and by recording questions in assignment notebooks. The authors categorized questions using a two-category taxonomy: (1) confirmatory questions, which were characterized by questions that clarified ideas, sought detail or facts, or addressed issues of specificity, exemplification or definition, and (2) transformational questions, which provided evidence of restructuring of understanding, were hypothetic-deductive, or extended what was known. Over the course the semester, the percentage of transformational questions increased from 13% to 48% of the total number of questions. Question frequency did not increase linearly, with the most questions posed during the second and fourth (of four) month. Questions peaked during supplementary sessions, such as topical lectures and case-study seminars. The authors concluded that as students' knowledge deepened, students' questions reflected deeper processing of the information. In addition, students asked more questions in learning contexts that focused on topics of interest and which allowed greater student interactions compared to large-group lecture settings. Interestingly, different settings did not differ by question type.

Chin, Brown and Bruce (2002) found similar patterns with younger students. The authors analyzed the questions 8^{th} grade science students asked during lab experiments using a similar two-level taxonomy: basic information questions (e.g., requesting factual or procedural information) and wonderment questions (e.g., comprehension, planning, predicting, applying). The authors found that 86% of students' questions addressed procedural or basic information.

Not surprising, the frequency with which students ask questions is shaped by the social context and the learning task. Therefore, the number of questions students generate can be quite high when conditions support question asking. In one-on-one tutoring sessions, for example, students tend to ask many more questions than they do in class (Fishbein et al., 1990; Graesser & Person, 1994). Graesser and Person (1994) investigated undergraduate psychology students' questioning during two 1-hour tutoring sessions and observed that the frequency of questioning was much higher in one-on-one tutoring environments than had been found in classroom environments. The authors reported that a student in their tutoring study was 240 times more likely to ask a question in 1 hour of tutoring compared to the rates Dillon (1988) found for question asking in class. Students' questions, although more frequent, were likely to request basic information. Ninety-two percent of the questions students generated corresponded to the lowest level of Bloom's taxonomy (i.e., factual knowledge).

Graesser et al. (1996) demonstrated the point that students can ask questions at dramatically higher rates with learning aids such as a menu of questions. The authors investigated student questioning in an individualized computer environment in which students explored content about musical instruments exclusively by selecting from a palette of questions. In such conditions, the average student generated 135 questions in an hour, 800 times the average rate of questioning Dillon (1988) found in classroom environments.

Summary

Taken together, the literature reviewed in the preceding section demonstrates that the number and type of questions can be important characteristics when studying student questioning in classroom setting and have a demonstrated history as variables of interest in the student questioning literature. There are several general tendencies in this literature with regard to the type and frequency of student-generated questions in classroom settings. Student-generated questions are often found to be infrequent during classroom instruction, a pattern that appears across a range of developmental levels and domains. This pattern has also been found in introductory undergraduate chemistry courses, even when students could ask their questions via private channels of communication. Studies that have collected undergraduate STEM students' questions via a written medium found frequencies similar to studies that investigate question asking in public settings.

Researchers have observed variability with regard to question frequency. Studies with designs that sample single periods across many classrooms demonstrated that questioning frequency varies considerably by classroom, suggesting the importance of contextual factors such as the classroom environment. Studies with designs that examine questioning over time demonstrated considerable differences in the same students' questioning patterns from class period to class period. Question frequencies can change dramatically over time, perhaps in response to assessment cycles and differences in instructional activities and the learning content.

Regarding the frequency with which students ask different types of questions, the different taxonomies (described below) used by researchers to categorize questions presents a significant challenge to synthesizing the findings across studies. A request for an example may be characterized as a lower level question by one set of researchers and as elaborative one by another team. Most studies do not publish the contents of the original questions, making it very difficult to undertake a cross-study analysis with a common typography. Nevertheless, the evidence from multiple studies suggests the following trend: students ask questions that clarify or confirm ideas or request basic information much more frequently than they ask questions that suggest deeper processing of information, such as questions that demonstrate explanatory and hypothetic-deductive reasoning or curiosity and wonderment. Questions that suggest deeper and more integrative processing of information are rare, although such questions may increase over time as students' knowledge base and perhaps motivation increases.

Several issues regarding the frequency and types of students' questions remain unknown. The methods employed by many of these studies confounded the frequency of *generating* questions with the act of *asking* them. It is unclear if the frequency of students' questions is relatively low because students do not naturally have many questions in their minds during classroom instruction or whether students generate many questions in their minds during learning and filter most of them. A better understanding of the relationship between question generation and question asking has the potential to inform interventions that seek to increase or improve student-generated questions. For example, a foundational question is the extent to which interventions should aim to stimulate question generation or focus on lowering barriers to question asking? Researchers have yet to document basic descriptive patterns about the frequency and type of questions that undergraduates generate during classroom activities and their evaluations of these questions.

A second key gap in the existing literature on student-generated question is the lack of knowledge about the role of some individual characteristics. The studies reviewed above have examined the mean frequencies and types of questions for groups of learners irrespective of individual differences on variables that have theoretical reasons for influencing questioning. To the extent that researchers have investigated individual differences, the exclusive focus has been on differences in prior knowledge, which I describe below in greater detail. Other individual characteristics, such as students' motivations, however, may play an important role in shaping the frequency and types of questions students generate and how they evaluate their questions for themselves and others.

In the preceding section, I reviewed theoretical literature on key components in questioning models and the empirical literature on the types and frequency of questions in classrooms. I now turn to examine the relationship between student questioning and various individual and context factors.

Individual Factors that Influence Question Generation and Asking

Researchers have identified cognitive and motivational variables that are related to formulating and asking questions in classroom environments. These include prior knowledge and achievement, interest, achievement goals orientations and goal structures, and self-efficacy. I describe the theoretical rationale for why these variables may influence question generation and asking along with a review of relevant empirical studies.

The Role of Prior Knowledge and Achievement

Prior Knowledge

Prior knowledge has been identified as a key variable for shaping student-generated questions. As described above, the notion that sincere information-seeking questions are triggered from awareness of knowledge deficits is broadly accepted. Prior knowledge can be conceptualized as an organized and elaborated network of relationships about a concept or situation (Ausebel, 1963; Bruner, Goodnow, & Austin, 1956; Fenker, 1975; Olney, Graesser & Person, 2012). Questions are generated based on the awareness of gaps in an individual's schema, such as the nature of components in the schema, their relations, or missing or poorly defined aspects of a conceptual network. Theoretically, schema might influence the quantity and qualities of students' questions in disparate ways. Less elaborated or poorer organized schema may lead an individual to be aware of a greater number of gaps in knowledge or more substantial gaps. In such a case, lower prior knowledge may result in more questions and more of these questions may seek basic orienting information. Alternatively, more elaborated and better-organized schema may provide more opportunities to become aware of gaps (e.g., the greater number of possible relations between nodes in a schema) and these gaps may reflect qualitatively different types of deficits.

In their seminal study of question asking, Miyake and Norman (1979) hypothesized that prior knowledge would have an inverse, U-shaped relationship with the number of questions an individual was likely to generate. That is, learners with low prior knowledge were likely to ask few questions because they lacked the necessary structures to become aware of their knowledge deficits. Students with high prior knowledge were likely to ask fewer questions because the more elaborated nature of their knowledge would yield fewer deficits. Student with a medium amount of prior knowledge would ask more questions because they had sufficient knowledge to generate questions and insufficient knowledge to have already mastered the material. The authors studied the questions introductory psychology undergraduates generated in a think aloud while reading either an easy (i.e., non-technical) or a hard (i.e., technical) manual about a software program. Students were assigned to a control condition or a prior knowledge condition, in which prior knowledge was experimentally manipulated by receiving a brief training on a similar technical device prior to the tasks. The authors found a condition by task interaction. Compared to student in the control condition, students in the prior knowledge condition asked fewer questions with the easier manual and more questions with the harder condition. The reverse pattern was found for subjects in the control condition. The authors interpreted the findings in support of the inverse, curvilinear relationship between prior knowledge and question asking frequency.

This hypothesis has been widely cited (e.g., Flammer, 1981; Loewenstein, 1994; Neumann, 2000; Taboada & Guthrie, 2006; Tan & Seah, 2011) despite the limits to the external validity of the study and some empirical evidence to the contrary. Van der Meij (1990) noted Miyake and Norman's findings were replicated only once and that the majority of the empirical evidence supported a linear and negative relationship between prior knowledge and question generation. Van der Meij noted that presumably learners with low prior knowledge activate general schema that enable them to generate questions.

Studies that examine only the relationship between prior knowledge and the number of questions ignored other factors that likely complicate the straightforward association suggested by the studies described above. The negative linear or negative curvilinear hypotheses focused on the knowledge deficit component of the questioning process. However, prior knowledge may also influence the editing and evaluative component of questioning. A student with low-prior knowledge may feel more self-conscious about his or her knowledge deficit and therefore be less willing to expose it in a public setting. Good et al.'s (1987) study of questioning patterns for students of different achievement levels, which I described previously, provided evidence that this can be the case: low achieving students asked fewer questions than med- and high-achieving students at older, but not younger, developmental stages. Newman and Goldin (1990) observed a decline in student-generated questions in classrooms around late-elementary and middle school, which corresponded to the increasing developmental concerns about peers' perceptions.

From the perspective of schema theory, prior knowledge influences the *type* of questions in addition to the quantity of questions. Dillon (1988) and van der Meij (1990; 1994) have described how a question contains presumptions about what the asker believes to be true and on which a question rests. For example, the question *Who is the prime minister of India?* presumes that the country of India has a prime minister. Similarly, we might expect students to ask about basic facts and definitions before they ask about their relations. To build on the previous example regarding the Indian prime minister, we might expect a student to ask about what a prime minister is or what *India* means before we would expect a question that applies these concepts. Therefore, we would logically expect for prior knowledge to influence the types of questions students ask. Presumably, prior knowledge would be associated with questions that reflect more elaborated knowledge.

There is some empirical support for this notion. Graesser, Lu, Olde, Cooper-Pye, and Whitten (2005) investigated the questions and eye moments of 40 undergraduate psychology students viewing a diagram as they tried to understand the mechanical failure of a lock and key apparatus. The authors found that students who had a deep understanding of how the apparatus worked asked higher quality questions and fixated their gaze on key components compared to those with shallow comprehension. After studying the illustration of the device, the comprehension scores on how the device worked were significantly correlated with questions that addressed plausible causes for the device failure, but were not significantly correlated with the total number of questions. Prior knowledge was associated with higher quality questions (though not higher quantity of questions) and these questions were in turn associated with comprehension.

Scardamalia and Bereiter (1992) investigated whether grade 5 and 6 students asked qualitatively different questions with different level of prior knowledge. One class had recently completed a unit on fossil fuels and thus was presumed to have more prior knowledge. A similar class had not yet studied fossil fuels and thus was presumed to have low prior knowledge. Both groups were allotted 25 minutes to generate written questions about the topic. Interestingly, students in both groups generated the same number of questions: approximately 5.5 questions, a finding that supported neither of the hypotheses described above regarding the relationship between prior knowledge and the quantity of questions. The authors found counterintuitive results. Compared to the students who had completed the unit on fossil fuels, the students who had *not* studied fossil fuels generated more questions that were rated by experimenters as having greater potential to promote advances in knowledge. That is, students with less prior knowledge asked questions that were deemed to have greater potential for learning. The authors argued that the findings suggested that little prior knowledge was necessary for students to generate questions which, if pursued, would lead to productive knowledge advances. The students who had recently studied the topic tended to ask narrower and more detail-oriented questions, which by the standards used by the authors—who were interested in productive questions for inquiry-driven learning—were deemed less productive.

In sum, regarding the quality of questions, there is a small amount of evidence that tentatively suggests that higher prior knowledge may be associated with deeper processing, such as explanatory reasoning. At the same time, it appears students even at a young age and with limited prior knowledge can generate questions that could productively guide future study. Regarding the quantity of questions, prior knowledge influences the number of questions students generate in ways that may be difficulty to predict. Researchers have found that prior knowledge has been associated with more, less, and the same number of questions and with linear, curvilinear, and no relationships. Importantly, these relationships have not been explored when motivational variables have been considered. Consequently, the relationship between prior knowledge and the quantity or quality of questions may have been confounded with motivational variables.

Prior Achievement

There are various reasons why prior achievement may be associated with the frequency and type of questions students generate or ask. Questions are means and products of processing of information, active learning, or comprehension monitoring. To the extent that prior achievement successfully captures the tendency of students to process and comprehend deeply, we might expect that students with higher prior achievement would engage in different questioning patterns compared with students with lower prior achievement. Prior achievement may also be a proxy for other relevant variables, such as prior education, study habits, abilities, and motivation, each of which may singly and in combination influence the quantity and qualities of student-generated questions.

Findings from the limited number of studies that have reported the relationship between achievement and student questioning patterns suggested that prior achievement may be positively associated with the quantity and quality of questions. As described previously, Good et al. (1987) found that lower achieving middle and high school students asked less questions than their higher achieving peers. Graesser and Person (1994) found that the prior achievement of undergraduate psychology students was inversely related to the number of questions students generated during tutoring but positively correlated with the quality of questions (e.g., questions that corresponded to higher levels of cognition according to Bloom's taxonomy).

In sum, prior achievement appears to be positively associated with the quality of questions and may be negatively associated with the quantity of questions. The theoretical explanation for these relationships is difficult to determine due to the fact that many cognitive, motivational and contextual variables influence achievement.

Interest

Interest is a motivational construct that may influence question generation and question asking. Hidi and Renninger (2006) defined interest as "the psychological state of engaging or the disposition to re-engage with a particular class of objects, events, or ideas over time" (p. 112). Researchers conceptualize interest differently (for reviews see Renninger & Hidi, 2011; Schiefele, 2009). There is general agreement that interest includes both affective and cognitive components (Linnenbrink-Garcia et al., 2010; Schiefele, 1991; Renninger & Hidi, 2011). The affective component typically refers to the positive emotions experienced during engagement. The cognitive component typically refers to the activation of perceptions, representations, and values that occur during engagement (Hidi & Renninger, 2006). The affective and cognitive components that comprise interest are the product of an interaction between an individual and a specific context.

Researchers typically differentiate between two types of interest: situational and personal interest. Situational interest refers to the state of interest that results in response to specific stimuli or context (Renninger, Hidi & Krapp, 1992). Situational interest has been theorized to be comprised of multiple components, including a triggering or "catch" phase, when one initially experiences novelty, surprise, complexity or incongruity (e.g., Berlyne, 1960) and a "hold" phase (Hidi & Baird, 1986; Mitchell, 1993), when the state of interest is maintained. Linnenbrink-Garcia and colleagues (2011) have recently argued that the latter hold stage consists of both affective and valuing component.

Personal interest, a second type of interest, is characterized by the disposition to re-engage with a domain over time (Schiefele, 1991; Renninger, 2000). The disposition to re-engage in a domain is theorized to be driven by accrued knowledge for a domain and accrued value for knowing in a domain (Alexander, 1997; Hidi & Renninger, 2006). Prominent conceptualizations of individual interest identify both value and affective components (Krapp & Prenzel, 2011; Renninger & Hidi, 2011; Schiefele, 2001). Individuals with a personal interest in a topic or domain value the domain (e.g., because it is useful or important to who they are) and because engaging with the topic or domain is generally enjoyable.

Situational and personal interest are related in models of interest development (Hidi & Renninger, 2006). Personal interest can develop from situational interest. Positive experiences re-engaging with a domain can lead to increased knowledge and value for knowing in a domain, which can further increase the likelihood of reengagement. In this way, situational interest can develop over time into personal interest. Hidi and Renninger conceptualize four phases of interest: (1) triggered and (2) maintained situational interest and (3) emerging and (4) well-developed personal interest. Different phases of interest may be associated with different types of questions.

How might situational and personal interest be related to the questions students generate? Interest has been found to shape attention and strategy use during learning, which independently and in combination may influence the questions students generate. Several researchers have theorized that attention mediates the influence of interest on learning, such that interest directs attention, which is a necessary condition for learning (Berlyne, 1960; Dewey, 1913; Hidi, 1995; Hidi, Renninger & Krapp, 2004; James, 1890; Renninger, 1990; Thorndike, 1935). While there are disagreements among interest researcher regarding the construct of situational interest (for a review, see Renninger & Hidi, 2011), Ainley and Patrick (2006) contended that "when triggered by achievement tasks the state of interest involves focused energy, attention, concentration and positive affect" (p. 179). The focused attention and concentration associated with situational interest could affect the questions students generate in different and potentially opposing ways. On the one hand, energy, attention, concentration, and positive affect could lead to more or higher quality questions as student monitor their understanding closely, process information deeply, and may have the situational disposition to entertain their questions. On the other hand, this concentration and focus may be associated with better comprehension and therefore, fewer questions or questions that do not reflect core knowledge deficits.

In addition to directing attention, the values for knowing in a domain that characterize personal interest may promote deeper processing of information. Deeper processing may, in turn, generate different patterns of questioning. Renninger (2000) argued that values for knowing in a domain activate deep processing strategies during learning which facilitate the assimilation and accommodation processes through which knowledge is constructed (Piaget, 1952). That is, individuals who have accumulated knowledge and value for a domain—more personal interest—may be more likely to effortfully elaborate, deepen, reorganize, or otherwise enrich his or her understanding compared with individuals with low interest. Renninger (2000) referred to *curiosity questions*, which “are not simply answered” but “involve some reorganization of what has been understood to date” and which “result in enhanced value of the subject content” (p. 377). Curiosity questions are similar to what Scardamalia and Bereiter (1992) termed *wonderment questions*, which “reflect curiosity, puzzlement, skepticism, or a knowledge-based speculation, in contrast to a groping for basic orienting information” (p. 188). Students who have more developed interests may generate more wonderment or curiosity questions during classroom experiences.

Schiefele (1991) described a series of studies that provide evidence for the role of interest in promoting adaptive learning strategies. In a sample of university students, Schiefele found that personal interest was most highly correlated (r = .44-.47) with learning strategies that promoted deeper processing of material, such as elaboration and information-seeking, compared to strategies that required more surface-level processing (e.g., rehearsal).

Interest is also related to students’ goals for learning and their purposes for engaging in such goals. Hidi and Renninger (2006) argued that compared with students with less developed interest, students with more developed interest set different goals for learning. For example, Lipstein and Renninger (2006) reported that students aged 12-15 with less developed personal interest set minimal goals for writing, such as merely completing the writing task, while students with more developed interest set ambitious writing goals for themselves. Goal may, in turn, influence the questions students generate during learning. Renninger (2000) suggested that learners' questions are grounded in their learning goals. For example, "a child working with multiplication might ask, 'Why isn't six times seven the same as five times eight?' The ability to pose questions that are rooted both in what is known and in what still needs to be figured out is the basis for a person's developing knowledge about what he or she could do or might eventually be able to undertake in pursuing particular subject content" (p. 376).

Some empirical literature has examined the relations among personal interest and questions. For example, Baram-Tsarbari and Yarden (2007, 2009) have examined the questions that people submit into *ask-a-scientist* websites, and have analyzed questions to argue for the presence of interest profiles and developmental trends. However, little is known about the relations among personal interest and question generation in typical classroom settings. As I noted above, higher levels of prior knowledge have been associated with both more (Miyake & Norman, 1979) and less (van der Meij, 1990) questions. These discrepant findings may be related to differing levels of personal interest. For example, two individuals with comparable levels of prior knowledge but different levels of personal interest may generate different patterns of questions. An individual with high interest may be more likely to seek out questions or engage in deep processing of information that leads to questions whereas the individual with low interest may be unlikely to go beyond the presented material.

It is important to keep in mind that personal interest may influence behaviors outside of the classroom that affect the questions students generate during classroom activities. For example, students with more developed personal interest may be more likely to prepare for class meetings, such as completing required readings and bringing supplementary study materials to class, which could influence the questions they generate and asking during class.

Furthermore, personal interest may also influence how a learner perceives and appraises the questions she generates. Personal interest may influence how students evaluate their questions since interest is based on accrued knowledge and value in a domain. Compared with students with low interest, learners with high personal interest may be more likely to view their questions as interesting, important, and valuable for their learning and the learning of others. In addition, in order for questioning to support learning, a question presumably must drive actions that result in acquiring knowledge, furthering understanding, and/or future re-engagement with materials. These actions can involve internal cognitions, such as making an inference based on prior knowledge, or external actions, such as seeking out information from a teacher, classmate, text or other source of information. Since personal interest is characterized by value for knowing in a domain, individuals with high personal interest may be more likely value the act of filling a gap in knowledge.

In sum, there are several theoretical reasons to suggest that personal interest may related to the quantity or qualities of student-generated questions. The accrued knowledge and value that characterize personal interest may shape attention and learning goal, which may influence patterns of questions. The attention, concentration, and positive affect associated with both situational and personal interest may be related in question generation and asking in complex ways. Despite compelling theoretical connections, I was unable to locate any empirical studies that have explored these relations.

Achievement Goals and Help Seeking

Achievement Goals

Achievement goal orientations may influence the questions that students generate and how students perceive and respond to their questions. Achievement goal theory is a prevailing theoretical framework for investigating motivation in educational environments (Kaplan & Maehr, 1997). Achievement goals (Ames, 1992; Dweck, 1986; Nicholls, 1984) refer to the purpose of why an individual engages in a task (Anderman & Maehr, 1994). Researchers generally conceptualize two types of achievement goal orientation: *Mastery goals*, which are characterized by an individual's focus on *developing* competence and *performance goals*, which are characterized by *demonstrating* competence. Most researchers further conceptualize performance goals as having either a positive (approach) or negative (avoidance) valance. Accordingly, performance-approach goal orientation refer to the goal to demonstrate normative competence while a performance-avoidance goal orientation refers to the goal to avoiding appearing incompetent by normative standards. Some researchers also distinguish valances for mastery orientations. Elliot & McGregor (2001) proposed a 2 x 2 achievement goal framework, which posited analogous approach and avoidance valences for mastery goals as have been recognized with performance goals. Accordingly, mastery-approach goal orientation refers to the goal of developing competence according to task-based on intrapersonal standards while mastery-avoidance goal orientation refers to the goal of avoiding task-based or intrapersonal incompetence. The distinction between mastery approach and mastery avoidance is relatively recent and, relative to the other goal orientations, has received considerably less attention by researchers (for a recent meta-analysis, see Hulleman et al., 2010).

Goal orientations were initially conceptualized as situated in a specific context; however, researchers have also conceptualized them as more stable tendencies across achievement situations (Kaplan & Maehr, 2007). An extensive empirical literature suggests that mastery goal orientations are associated with adaptive forms of engagement in educational settings and performance goals, particularly performance-avoidance goals, are associated with maladaptive patterns of motivation (for reviews see Anderman & Wolters, 2006; Kaplan & Maehr, 2007).

Achievement goal orientations may directly influence the questions students generate. Given that mastery orientation has been positively associated with the use of effective and deep processing strategies, one may expect that students with a strong mastery orientation would engage in strategies that resulted in higher quality questions compared with students with low mastery orientation. Given that a performance-avoidance is associated with less effective and more surface processing strategies, one may expect that a performance-avoidance orientation would generate lower quality questions. The relationship between question generation and performance-approach orientation is less clear. Students with a performance-approach orientation would likely seek to demonstrate competence in the normative ways competence is demonstrated for a context. It may be that the drive to demonstrate competence would lead to effective learning strategies and therefore high quality questions. Alternatively, concerns with demonstrating competence may influence students to only generate questions that relate to assessments. There are not clear reasons to explain why goal orientations would systematically lead to greater or lesser numbers of questions. I have been unable to find any empirical studies that examine the relationship between goal orientations and question generation.

Achievement goal orientations may also influence whether students are willing to seek help or ask their questions during class. To the extent that question asking is an instantiation of help seeking, the literature on the relationship between achievement goals and help seeking—which I will describe below—can inform this issue. It is important to recall, however, that question asking in classrooms can also differ from typical conceptualizations of help seeking in ways that may influence the relationship between goal orientations and questioning behavior. Unlike many instances of help seeking, the act of asking a question can present an opportunity to publically demonstrate knowledge, ability and interest. A student with a performance-approach orientation may be motivated to ask a so-called smart question to publically display prior knowledge or to demonstrate intellectual prowess. Good et al. (1987) observed this type of question asking behavior and applied the termed “attention on-task questions” for those inquiries that were believed to be motivated to bring positive attention to the asker. By contrast, a mastery-oriented student may be motivated by different reasons, such as to learn the material fully or to pursue a genuine interest or curiosity about the topic.

Goal orientations may be related to how students evaluate or appraise their questions. In particular, the meaning of a question’s value for oneself or others may differ by goal orientations. For example, students who are mastery oriented may perceive value in their questions for supporting others, who like themselves, are striving to develop competence. By contrast, students who performance oriented may perceive their question as having value for others as they compete to demonstrate normative or relative competence.

Help Seeking

Researchers have identified a set of help seeking constructs, which are related to question generation and asking. Nelson Le-Gall's (1981; 1985) distinguished between two types of help-seeking: instrumental and executive. *Instrumental help seeking* refers to help seeking behavior that aims at mastering material and to help that is sought only when it is necessary and to the extent to which it supports understanding (e.g., asking for a hint). By contrast, *executive help seeking* (also called *expedient help seeking*) refers to help seeking behavior that seeks to minimize time or effort on a task regardless of learning (e.g., asking for an answer). Researchers have also distinguished between two types of targets for those seeking help: formal and informal. *Formal help seeking* refers to help sought from formal sources such as teachers. *Informal help seeking* refers to help sought from informal sources such as peers. Two additional constructs are often used in research on academic help seeking. *Help seeking threat* refers to the extent to which an individual's perceived the act of seeking help as threating to himself. Researchers have identified different sources for threats including threats to cognitive competence (Karabenick & Knapp, 1991; Newman, 1990; Newman & Goldin, 1990; Ryan, Hicks & Midgley ,1997) and threats to social competence (Ryan & Pintrich, 1997). Much of the research on help seeking has used survey methods that assess intentions to seek or avoid help rather than observed behaviors. Thus, outcome variables of studies are assessed with *help seeking avoidance*, which taps the self-reported intentions to avoid seeking help (e.g., Karabenick, 2003).

Questions are often construed as a form of help seeking (e.g., Butler, 2007; Karabenick, 1998). It bears mention, however, that the act of pursuing an answer to questions—including ones that are sincere information-seeking questions—is not necessarily an act of help seeking. For example, discourse analysis conceptualizes students' questions in classroom discourse as "bids" or moves made, which act to potentially shift or direct the nature and direction of classroom talk. Furthermore, questioning can sometimes be more adequately described as information seeking rather than help seeking. As I describe above, Renninger (2000) argued that as students accrue knowledge and value for knowing in a domain they begin generating "curiosity questions," which tap deep, domain-specific issues.

While some information-seeking questions may not be instantiations of help seeking, in many instances the act of a student asking a question shares many similarities to the act of a student seeking help. For example, when a student is confused or reaches an impasse in solving a problem, asking a question to a teacher or peer is, in fact, a form of seeking help. As I describe above, researchers have found that "wonderment" questions are relatively rare in classroom discourse and students' questions that seek clarification or basic information are much more prevalent (e.g., Chin et al., 2002; Texiera-Dias et al., 1997).

Relations Between Achievement Goals, Help Seeking, and Questioning

Achievement goal orientations have been theorized to influence academic help seeking behaviors, such as asking a question in class. Mastery-oriented learners are more likely to perceive help seeking as beneficial, are less threatened by help seeking, and are more likely to engage in instrumental help seeking compared to performance-orientated students (Butler & Neuman, 1995; Newman, 1991, 1994; Ryan & Pintrich, 1997; Ryan, Gheen, & Midgley, 1998).

By contrast, performance-oriented students are theorized to be more likely to perceive help seeking as a threat to competence. Empirical studies have supported this claim. A performance goal orientation has been associated with help seeking threats and help seeking avoidance (Butler & Neuman, 1995; Newman, 1991; 1994: Ryan et al., 1997; Ryan & Pintrich, 1997). Although much help seeking literature has focused on students in elementary and secondary school settings, the relationships described above appear to present in undergraduate and adult learning environments as well. For example, Karabenick (2003) found that for students in large enrollment courses mastery orientation had a negative relationship with help seeking threat, a positive relationship with instrumental help seeking, and a positive relationship with students' preference to seek help from formal sources. Performance-approach and performance-avoidance orientations were associated with avoidance of help seeking.

Skaalvik and Skaalvik (2003) studied the relationship between goal orientation and help seeking in a sample of Norwegian adults who were returning to school for a high school diploma. The authors found that perceived threats of help seeking were negatively predicted by ego-enhancing (i.e., performance-approach) and ego-defeating (i.e., performance-avoid) goal orientations, with stronger path coefficients for the latter variable. Contrary to much of the extant literature, the authors did not find a relationship between task-focused (i.e., mastery) orientation and threats to help seeking.

Goal orientations may influence the questions students generate and ask. As I mentioned above, question asking in classroom environments can be an instantiation of help seeking and, as such, may be similarly influenced by competence threats and goal orientations. Questioning is a unique form of classroom communication in that a question asker reveals his or her prior knowledge, assumptions, interests, and ignorance to listeners (Dillon, 1988; van der Meij, 1994). One common inhibition to asking a question in public is the fear that one will ask a so-called *stupid question*. Such concerns are a prototypical example of how asking a question may be perceived as a threat to competence.

Karabenick (1994) demonstrated that goal orientations influence how an individual thinks about the hospitability of an environment for questioning and the meaning of questions in that environment. The author investigated the relationship between goal orientation, perceptions of the openness of the environment to questioning, and students' perceptions of the teachers' attribution for students' questions. Using survey methods in a undergraduate psychology course, Karabenick found that students with a mastery orientation were more likely to perceive the environment as supportive of questioning compared to performance-oriented peers. In addition, goal orientations were associated with different perceptions of teachers' attributions about why students asked questions. Students with a mastery orientation were more likely to perceive that the teacher attributed student questioning to motivation to learn while students with a performance orientation were more likely to assume that the teacher would make negative attributions, such as lack of ability.

In sum, goal orientations have been found to influence help seeking behavior in classrooms. In general, mastery orientation has been associated with instrumental help seeking where as performance-avoid orientations have been associated with avoidance of help seeking. Researchers have yet to examine the relationship between goal orientations and question asking in actual classrooms. Much of the extant literature on help seeking used survey methods which examined students' self-reported intentions to seek help. Such assessments tap self-perceptions regarding help seeking rather than actual classrooms behaviors, which may differ from perceptions. Furthermore, little is known about the relations between goal orientation and student-generated questions. Different goal orientations may lead students to engage in different learning strategies, which lead to different patterns of questions. Having discussed the goal orientations individuals adopt for themselves I now turn to address how students perceive the goal structures of their classroom environment.

Goal Structures in the Classroom

Achievement goal theory assumes that an individual's motives for achievement are influenced by perceptions of the achievement environment in addition to personal beliefs and dispositions (Ames, 1992; Maehr, 1984; Nicholls, 1989). *Goal structures* refer to perceptions of environmental features that support certain individual goal orientations (Ames, 1992). As with individual goal orientations, two principal types of achievement goal structures are recognized: mastery and performance. A classroom with a mastery goal structure is perceived to support improvement and growth. In such environments academic risk-taking is encouraged, mistakes are framed as learning opportunities, success is defined as self-referenced—as opposed to norm-referenced—improvement. A classroom with a performance goal structure emphasizes extrinsic rewards and defines success as out-performing others. In line with the valence distinctions for individual goal orientations, performance goal structures can be further divided into performance-approach and performance-avoid goal structures with the former emphasizing achieving better than others and the latter emphasizing not performing worse than others and avoiding mistakes. It is important to draw attention to the fact that goal structures do not refer to objective realities regarding the classroom environment but rather individual perceptions of the environment. Researchers have found perceptions of goal structures to vary significantly across individuals within the same instructional context (Kaplan, Gheen, & Midgley, 2002; Ryan, Gheen, & Midgley, 1998; Wolters, 2004). Researchers have examined students' perceptions regarding teacher emphasis on goals (i.e., teacher goals) as well as perceived goals for the classroom in general (i.e., classroom goal structures). Compared with classroom goal structure, teacher goals have received stronger validation evaluations (e.g., Koskey, Karabenick, Woolley, Bonney, & Dever, 2010), and therefore were used in the present study.

Mastery goal structure has been found to be associated with adaptive motivational, cognitive and affective outcomes. For example, mastery goal structure was associated with students adopting an individual mastery goal orientation, self-efficacy, use of effective learning and coping strategies, seeking challenge and help when it is necessary, effort, positive affect, coping strategies, and satisfaction during learning (e.g., Ames & Archer, 1988; Kaplan & Midgley, 1999; Karabenick, 1994; Moos & Azevedo, 2006; Murayama & Elliot, 2009; Nolen, 2003; Ryan & Pintrich, 1997; Wolters, 2004). By contrast, performance goal structures tend to be associated with individual performance goal orientations and maladaptive motivations, cognitions, and affect, such as avoiding help when it is necessary, low academic self-concept, disruptive behavior, self-handicapping, cheating, and negative affect (Kaplan, Gheen, & Midgley, 2002; Moos & Azevedo, 2006; Middleton, Kaplan, & Midgley, 2004; Murayama & Elliot, 2009; Murdoch, Miller & Kohlhardt, 2004; Ryan & Pintrich, 1997).

A mastery goal structure tends to exert stronger influence on students' beliefs, actions, and achievement than does a performance goal structure (Kaplan & Maehr, 2007), a pattern that may be the result of the mediating influence of competence (Middleton, Kaplan & Midgley, 2004). Whereas students of all ability levels benefit from an environment that focuses on improvement, only relatively high performing students benefit from an environment that fosters competition.

What relationship might perceived mastery goal structures have with question generation and question asking? To the extent that questions are a form of help seeking, evidence suggests that a mastery goal structure might be associated with less avoidance of help seeking (Ryan, Gheen, & Midgley, 1998; Karabenick, 2004). Karabenick (2004) investigated whether perceived goal structures were associated with intentions to seek or avoid help for undergraduate students in large-enrollment courses. After controlling for individual goal orientations, perceived mastery goal structure was positively associated with intentions to seek needed help and negatively associated with avoidance of help seeking. A performance-avoid goal structure predicted the reverse pattern.

As I noted previously, researchers investigating the *generation* of questions (not question asking, per se) have largely examined cognitive rather than motivational antecedents. Therefore, there is no direct evidence to inform hypotheses about how goal structures may affect patterns of student-generated questions. Nevertheless, there are theoretical reasons that suggest relationships. A mastery goal structure is characterized by learning for understanding and has been associated with students using effective learning strategies (Ames & Archer, 1988; Karabenick, 2004; Nolen, 2003). Consequently, questions that demonstrate deep-processing of information may be more likely to be associated with a mastery goal structure compared to a performance goal structure. In addition, in a mastery-oriented classroom, questions may be perceived as valuable for both asker and classmates.

In sum, goal structure has been found to affect students' motivations, cognitions and affect, with mastery goal structures associated with more adaptive beliefs, attitudes and behaviors compared to performance goal structures. A mastery goal structure may be associated with unique patterns of question generation and question asking based on the positive relationships between mastery goal structure, adaptive help seeking, and the use of effective learning strategies.

Self-Efficacy

Competency beliefs influence individuals' motivations, beliefs and actions in academic setting. Self-efficacy is a central variable representing competency beliefs (Schunk & Zimmerman, 2006). Self-efficacy refers to peoples' belief in their ability to learn or perform actions at designated levels (Bandura, 1997). Self-efficacy is considered task-specific. That is, a STEM major's self-efficacy for the task of succeeding in her biology class may differ from her self-efficacy for succeeding in her chemistry or history course. In the same way, a student may have different levels of efficacy for tasks within a class, such as learning the contents of different parts of a chapter. Research has demonstrated that self-efficacy is an important variable in predicting initiation of, effort and persistence during academic tasks as well as learning and achievement in academic settings (Schunk & Zimmerman, 2006).

How does self-efficacy relate to question generation and asking? To the extent that question asking is a form of help seeking, research on the relationship between self-efficacy and help seeking is relevant. Self-efficacy has been found to be related to adaptive help seeking behaviors. Ryan et al., (1998) examined the relations between efficacy beliefs and self-reported patterns of help seeking in 6^{th} grade students. Self-efficacy was positively associated with intentions to seek needed help and negative associated with avoidance of help seeking. Karabenick (2003) investigated perceptions regarding help seeking for undergraduate students in large-enrollment courses. In this context, self-efficacy was significantly and positively associated with instrumental help seeking (r = .21) and seeking help from formal sources such as the instructor (r = .15). The results of a cluster analysis indicated four cluster profiles: (1) Strategic/adaptive help-seeking from formal sources; (2) strategic/adaptive help seeking from informal sources; (3) non-strategic help seeking; and (4) avoidant help seeking. Clusters one and two were characterized by high self-reported scores of cognitive and metacognitive strategies and high self-efficacy and differed primarily in their preference for helping sources (i.e., formal vs. informal). Cluster 3 was characterized by low use of cognitive and metacognitive strategies, greater help seeking threat (compared to clusters 1 and 2) and low self-efficacy. Students in Cluster 4, characterized by avoidance of help seeking, had similar levels of self-efficacy as students in cluster 3. While these findings demonstrated associations between self-efficacy and help seeking profiles, it is important to note that differences in mean levels of self-efficacy across clusters were modest (i.e., ranging from 3.2 to 3.6 on a 5-point scale).

Ryan and Shin (2011) argued that one reason why self-efficacy relates to help seeking implicates social motivational goals (Newman & Goldin, 1990; Ryan, et al., 1998). Students who believe they will be successful in school are more likely to view seeking needed help as beneficial to their learning goal. Their efficacy for the task emboldens them in the face of possible judgments from peers. By contrast, students with low self-efficacy are more likely to question the usefulness of seeking help and become more susceptible to fears about the judgments of others.

The theoretical relationship between self-efficacy and question *generation* is unclear. High self-efficacy may drive attention, effort and strategy use during lectures, which may result in more and higher quality questions. Alternatively, having many questions during lectures in a course (e.g., “feeling lost”) may be associated with low self-efficacy. Self-efficacy may also influence students’ behavior before attending class meetings, such initiating and persisting in homework tasks, which in turn influence the questions they have during class. It is important to note also that students’ questions may influence self-efficacy beliefs. The questions that students generate during learning may provide them with information about their level of mastery, for example, which in turn may lead to a change in self-efficacy beliefs.

In sum, self-efficacy is a key variable in predicting motivation, cognition and achievement, including in STEM contexts (Lau & Roeser, 2002). Self efficacy has been positively associated with using learning strategies and adaptive patterns of help seeking such as seeking help when it is necessary and seeking help from formal sources. There are no clear theoretical relationships between self-efficacy and question generation and no evidentiary base for generating hypotheses.

In the preceding section, I have reviewed theoretical and empirical literature on the relationships between student-generated questions and interest, goal orientations, goal structures, and self-efficacy. Below, I describe the relationship between these motivational variables, student-generated questions and achievement. I then summarize methods for studying student questioning in classrooms before concluding the chapter with research questions and hypotheses.

Relations Between Motivation and Achievement

A large body of empirical research has examined the effects of personal interests, goal orientations and self-efficacy on achievement. The empirical base for the relationships between these motivational variables and achievement is substantial. Since the direct effect of motivation on achievement is not central focus of this study, I describe the literature only briefly below.

Interest has a positive correlation with achievement (Schiefele, 2009). In a meta-analysis of the relationship between interest and achievement, Schiefele, Krapp, and Winteler (1992) observed that interest explains approximately 10% of the variance in achievement, which was relatively stable across grade levels and subject areas. In general, the association between achievement and personal interest appeared to be the strongest when achievement reflected deep, conceptual understanding since personal interest has been more associated with deep processing than with surface processing (Wigfield & Cambria, 2010).

The literature on goal orientations is vast and a full review of the relations between goal orientations and achievement is more complex than can be covered here. Nevertheless, I will briefly describe general patterns in the literature. As I noted previously, a mastery goal orientation has positive associations with many adaptive learning outcomes, yet empirical studies do not consistently find a significant positive effect of mastery orientation on instructor-given grades (Anderman & Wolters, 2006; Hulleman et al., 2010). With regard to findings on the effect of a performance-approach orientation on achievement, the pattern is also somewhat ambiguous: researchers sometimes—but not consistently—find significant positive effects of performance-approach on achievement (Anderman & Wolters, 2006; Hulleman et al.; Kaplan & Maehr, 2007). Several studies have found negative relationships between a performance-avoid orientation and achievement (Kaplan & Maehr, 2007).

Linnenbrink-Garcia, Tyson, and Patall (2008) reviewed 90 studies that examined the relationship between achievement goals and achievement. Mastery-approach goals were positively associated with achievement in about 40% of the studies; the same was true for performance-approach goals. The authors found some evidence that factors such as task difficult, ability level, and developmental factors may moderate the relationships between achievement goals and achievement. For example, a mastery goal orientation may have a stronger relationship with achievement for difficult rather than easy tasks or when individuals have relatively low rather than high ability levels. Performance-approach goals may be more associated with achievement with older rather than younger students. Thus, while the pattern between achievement goals and achievement is complicated, approach goals tend to have positive associations with achievement.

Self-efficacy is positively associated with achievement (Schunk & Zimmerman, 2006). In a meta-analysis of the effect of self-efficacy on achievement, Multon, Brown, and Lent (1991) found an average effect size of .35 for studies with college students. Of particular relevance to the proposed study, self-efficacy has been linked with achievement in the context of STEM undergraduate education (Besterfield-Scare, Atman, & Shuman, 1997; Hackett, Betz, Casas & Rocha-Singh, 1992; Lent, Brown & Larkin, 1987; Lent, Brown, Schmidt, Brenner, Lyons & Treistman, 2003).

Student-Generated Questions and Learning Outcomes

Students' questions can have a complex relationship with comprehension and learning outcomes, since questions can both indicate a lack of comprehension as well as facilitate processes that lead to increased comprehension. Since questions identify knowledge deficits, learners can use their questions as a source of information to monitor and make evaluations about their comprehension (Palincsar & Brown, 1984). As I noted previously, the desire to resolve knowledge deficits has been theorized to motivate cognitive, motivational, and social processes that can lead to improved comprehension (Berlyne, 1960; Chinn & Brewer, 1993; Festinger, 1962; Flammer, 1981; Loewenstein, 1994). For example, the presence of a question can trigger students to actively process information, initiate explanatory reasoning, and guide learners to identify and make connections between ideas and with prior knowledge (Chi et al., 1989; Collins, Brown, & Newman, 1989; Graesser, Baggett, & Williams, 1996).

At the same time, students' question represent gaps in understanding and therefore may indicate the extent to which the student generally lacks understands of the material. In settings such as lectures, students' questions—or some aspect of their questions such as the quantity or type—may indicate low comprehension. If students' questions do not trigger successful information seeking and processing, either during the lecture or afterwards, or are an indication of more general comprehension difficulties, then students' questions or question features may be predict low comprehension or achievement.

Findings from interventional studies strongly suggest that with training self-questioning processes can lead to measurable gains in learning (Bugg & McDaniel, 2012; King, 1992; Rosenshine, et al., 1996; Wong, 1985). Rosenshine, Meister, and Chapman conducted the most recent meta-analysis of the effects of classroom interventions that trained students to generate—and usually answer—questions on learning. The 26 studies included in the analyses involved students ranging from grade 3 to college. Interventions varied by the instructional methods used to train students, including having students generate questions with signal words (e.g., Who? What? When? Where? How?), with generic stems (e.g., How are *x* and *y* similar and different?), for conceptual categories (e.g., searching for the main idea or narrative components), and at different units of analysis in the target text (e.g., questions whose answers are found in a sentence or across multiple sentences). The average effect sizes were .36 and .86 on standardized and researcher-generated questions, respectively. In sum, the literature has demonstrated the potential for supporting student learning and achievement via interventions that train students to generate and answer their questions.

It is not clear how findings from intervention studies should be interpreted in light of the research design of the current study. Many of the interventions in the meta-analysis described above explicitly or implicitly involved answer seeking as part of training. Furthermore, the gains of the experimental groups were compared to control groups, who presumably were generating questions naturalistically. Therefore, it is difficult to extrapolate the findings to the effects of student-generated questions without training or to contexts in which question-generation is not framed as a question-and-answer-seeking learning strategy (e.g., Bugg & McDaniel, 2012; King, 1992).

In the preceding section, I reviewed the relations between selected motivational constructs, student questioning, and achievement. Before turning to the proposed research questions and hypotheses, I review effective methods for studying and categorizing student-generated questions.

Methods for Investigating Student-Generated Questioning

Researchers investigating student-generated questioning have examined the phenomenon in a range of settings—from highly controlled laboratory settings to interventions in "messy" classroom—and have used a variety of methods, including questionnaires, interviews, written logs, and naturalistic observations. Each setting and method provided a set of affordances and limitations that shaped the data that was collected and potentially influenced the behavior of questioners in unique ways. In the section below, I will highlight some of the more common research methods.

Many researchers have studied learners' questions in controlled laboratory settings (e.g., Allwood & Elliasson, 1988; Fishbein, Eckart, Lauver, Van Leeuwen, & Langmeyer, 1990; Graesser & McMahen, 1993; Graesser, Lu, Old, Cooper-Pye, &

Whitten, 2003; Karabenick, 1996; van der Meij, 1990). In laboratory conditions, researchers have a high degree of control over stimuli and can often control or randomly distribute extraneous variables. These conditions often lead to high internal validity. However, findings from highly controlled conditions may have limited external validity. Such studies investigated learning in the context of an experiment—often with psychology undergraduate students who participate for credit—rather than conditions that more closely resemble the process of learning in typical formal educational settings such as classrooms.

Within the context of classrooms, researchers have investigated student questioning using a variety of research methods. Observational studies investigating the type (Good et al., 1987) or frequency (Dillon, 1988; Good et al., 1987; Person & West, 1991 West & Person, 1994) of questions have sampled classroom discourse for periods of time. These studies provided a glimpse into generalizable patterns of students' questions across classrooms, domains, and topics. However, the sampling methods of these studies have ignored classroom contextual factors and usually did not examine questions over time periods that were meaningful in terms of conceptual learning (e.g., a curricular unit).

Survey methods have been a common means of collecting data on students' and teachers' perceptions relative to questioning during class and their reported intentions to seek or avoid help (e.g., Karabenick, 1994; Karabenick & Sharma, 1994; Ryan & Pintrich, 1997). Questionnaires have been helpful to researchers in quantifying perceptions and intentions. However, one limitation of studies that rely exclusively on data from questionnaires was that no data were collected about the actual questions that students generated while learning and their decision-making process about these specific questions. Investigating the actual questions students generate has the potential to extend the existing literature on the quantity and qualities of students' questions and their appraisals of the values of these questions.

Recent work in self-regulated learning has focused on the use of methods that allow researchers to examine the cognitive and metacognitive processes as they occur *during* learning activities. Zimmerman (2008) reviewed the history of self-regulated learning research methods, noting that recent theoretical advances stemmed from innovative *online measures* of learners' processing. Boekaerts and Corno (2005) described several developments in research methods that capture students' self-regulatory processes—such as question generation—during learning in classroom settings, including situated self-report questionnaires, observations, and learning diaries. Several studies have successfully captured students' online self-regulating processes using a variety of methods such as think-aloud studies (e.g., Chi et al., 1989; Azevedo & Cromley, 2004; Azevedo, Cromley & Siebert, 2004), stimulated recall (Kaplan, Lichtinger, & Margulis, 2011; Xu & Corno, 1998), and analyzing traces in computer-enhanced learning environments (Bernacki, Byrnes, & Cromley, 2011; Winne & Perry, 2000).

Learning journals or diaries are one promising new research method for capturing students' *online* cognitive and metacognitive processes (Boekaerts & Corno, 2005; Zimmerman, 2008). Several researchers have used learning journals as a follow-up activity to course work in which students write down the main idea of the lesson, articulate topics for further study, and plan for using learning strategies (Berthold, Nückles, & Renkl, 2007; Glogger, Schwonke, Holzäpfel, Nückles, & Renkl, 2009; 2012; Nückles, Hübner, & Renkl, 2009). Although such methods are still in the early stages, initial findings have suggested that students' use of cognitive and metacognitive strategies as manifest in learning journals predict learning and achievement.

Researchers have used written logs to record the questions students generated during learning activity in classroom contexts. For example, Graesser and McMahen (1993) instructed undergraduate students during a single class to generate questions in writing after reading a problem and then again while solving problems. As I described previously, Pedrosa de Jesus and colleagues tracked students' questions using a range of written media such as digital and paper drop boxes and notebooks. Furthermore, research on self-regulation using learning journals by Bethhold, Glogger, Nückles, Renkl, and colleagues included having students generate questions as a remediation strategy for gaps in understanding. For example, Glogger et al. (2012) used learning journals to study students' metacognitive monitoring and use of self-regulatory strategies after classroom instruction in 9^{th} grade math and biology classes. The authors found that both the quantity and quality of cognitive and metacognitive strategies used in learning journals predicted achievement. Interestingly, the authors noted that nearly a third of students did not use negative metacognitive monitoring (e.g., noting a gap in understanding, formulating an open question) during the six-week intervention.

Learning journals are an emerging method to capture students' questions in naturalistic context. To my knowledge, no studies have used written logs of questions during lectures. However, previous uses of learning journals provide a precedent for this promising method of studying questions as they occur during classroom learning.

Taxonomies of Questions

Another methodological challenge in the literature on student-generated questions is the process of categorizing questions. Several attempts have been made to categorize types of questions (for reviews see Chin & Osborne, 2008; Pomerantz, 2005). In the following section, I briefly describe the range question taxonomies and then provide a description and justification for using one such taxonomy.

A question's content can reveal the types of cognition that produced it or which are required to answer it. For example, Bloom's taxonomy (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956; Krathwohl, 2002) is an example of a heuristic widely used by educators for categorizing questions according to the level or depth of cognition that the question requires of the answerer. Researchers have created a variety of question taxonomies that attempt to group questions by the associated type of cognition. Questions have been characterized according to whether the primary source and referent is text-based or knowledge-based (e.g., Chin et al., 2002; Scardamalia & Bereiter, 1992), whether questions clarify or transform understanding (Pedrosa de Jesus, et al., 2003), the extent to which questions indicate factual or conceptual integration (Taboada & Guthrie, 2006), or by the semantic stem of the question, such as *Wh*- questions (e.g., Robinson & Rackstraw, 1972).

One comprehensive question taxonomy was initial developed by Lehnert (1978) and then expanded by Graesser and colleagues (Graesser, McMahen, & Johnson, 1994; Graesser, Person, & Huber, 1992; Graesser & Person, 1994). Similar to Bloom's taxonomy, the Graesser-Person-Huber (GPH) taxonomy classifies questions according to the type of answer(s) that a question logically solicits. Lehnert (1978) developed the foundation for this taxonomy in her story-telling understanding system QUALM. Graesser, Person, and Huber (GPH, 1992) added additional conceptual categories for a total of 18 classes of questions. The GPH question types are: (1) Verification, (2) Disjunctive, (3) Concept completion, (4) Feature specification, (5) Quantification, (6) Definition, (7) Example, (8) Comparison, (9) Interpretation, (10) Causal antecedent, (11) Causal consequence, (12) Goal orientation, (13) Instrumental/procedural, (14) Enablement, (15) Expectational, (16) Judgmental, (17) Assertion and (18) Request/directive. The Graesser-Person-Huber taxonomy reached its most elaborated form in Graesser, McMahen and Johnson (1994), in which categories of questions were grouped into those requiring short and long answers. This distinction has limited use for the proposed study and therefore, I will refer to the GPH as presented in Graesser et al. (1992).

Unlike Bloom's taxonomy, the GPH taxonomy was not designed to be strictly hierarchical with regard to the level of cognition required by a given class of questions. Nevertheless, Graesser and his colleagues have argued that some classes of questions are associated with deeper processing than other classes. For example, questions from the classes *causal antecedents* (10), *causal consequences* (11), *goal orientation* (12), *enablement* (13), *instrumental/procedural* (14), and *exceptional questions* (15) tend to be related to logical and explanatory reasoning or the application of ideas to contexts and are correlated with deeper levels of cognitive processing, as presented in Bloom's taxonomy (Graesser & Person, 1994). In addition, some question categories tend to indicate knowledge elaboration, comparison between concepts, or drawing inferences, such as the *examples* (7), *comparison* (8), and *interpretation* (9) categories.

It is important to note that question classes are not rigidly tied to depth of processing or type of cognition. While recognizing the tendency of some questions types to support deeper thinking, Graesser and colleagues did not define classes of question as qualitatively better or worse—higher-order or lower-order—than others. Yang (2006) cautioned against assuming cognitive levels can be easily deduced from decontextualized hierarchical taxonomies. The same question can demonstrate different levels of cognition for two different students. Yang provided as an example the question, *What are the causes of the U.S. Civil War?* One student may be performing the shallow-processing task of clarifying the bullet points from a preceding presentation slide, while a second student may be engaging in a higher-order task of understanding a causal explanation. The example highlights some of the problems of assuming that taxonomical categories can be strictly interpreted as evidence of higher-order thinking.

The classes of questions in the GPH taxonomy can be uses as either a *polythetic* or *monothetic* scheme. Polythetic schemes allow questions to be categorized as more than one class. Stokal (1974) argued that most adequate categorization schemes in the social sciences are polythetic. Questions often belong to more than one category. For example, the question *Was the economy a cause of the Civil War?* relates to a *causal antecedent* [10] while also being a *verification* question [1]). A question can solicit several logical answers and therefore be categorized into several classes. Graesser, McMahen, and Johnson (1994) provided the example question, *Why are you selling your car?* which can be logically and truthfully answered by referring to causal antecedents (e.g., “It has been leaking oil”) or goal orientation (e.g., “to save money for college”), with the former might be said to a friend and the latter said to a potential buyer. By providing a way to catalogue questions in multiple ways, the GPH taxonomy allows researchers to capture the actual complexity of questioning in social situations.

The GPH can also be used as a monothetic scheme. To do so, rules must be established for breaking "ties" when more than one class is relevant. For example, Graesser, Person and Huber (1992) used the taxonomy in this way by giving causal questions prominence over taxonomical questions.

The GPH taxonomy has several advantages over other taxonomies. Graesser and colleagues reported developing the taxonomy from an existing exhaustive taxonomy of speech acts that had the highest degree of reliability (Graesser, Person & Huber, 1992). The GPH avoids shortcomings of lexical taxonomies (e.g., *Wh*-questions) such as how different lexical arrangements can address the same contents (e.g., *What* is the length of the string? and *How* long is the string?). Compared with other taxonomies, the GPH taxonomy requires less inference on the part of the observer. For example, some taxonomies group questions by alleged cognitive function, such as whether a question clarifies or extends the asker's understanding (e.g., Watts, Gould & Alsop, 1997). Making this judgment requires information about the individual's purpose and present conceptualization, information that is often unavailable and must be inferred by the observer. Finally, the GPH taxonomy has been used widely to study questioning, including reading from science texts (Costa, Caldeira, Gallástegui, & Otero, 2000), computer-based and one-on-one tutoring environments (Graesser & Person, 1994), learning from concept-maps (Olney, Graesser & Person, 2012), physicians questions and database queries (Keyes, 1996; Staveri, 1996), watching TV (Graesser, Lang & Horgan, 1988) and librarian reference desk queries (White, 1998). Although GPH has not yet been used for investigating student questioning in classroom environments, its use in a variety of environments in which questions naturally occur suggests that it may be well suited for investigating student questions in classrooms.

Research Questions

Based on the theoretical and empirical literature reviewed above, I have developed a set of research that seek to extend the literature on student-generated questioning. I ask these questions in a particular context: undergraduates who are attending lectures in an introductory chemistry course. Given the exploratory nature of this study, I pose research questions rather than test hypotheses.

The first primary research question (RQ1) is: What are the relations among the quantity, qualities, and self-appraisals of student-generated questions? As I describe in the literature review, there is little empirical evidence regarding the quantity and qualities of questions during lectures (e.g., Pedrosa de Jesus et al., 2001). Therefore, I do not make specific hypotheses with regard to this question.

The second primary research question (RQ2) is: What are the relations among motivational self-beliefs and perceived teacher emphasis on mastery goals and the quantity, qualities, and self-appraisals of student-generated questions? Specifically, I examine the motivational self-beliefs of personal interest in chemistry, goal orientations and self-efficacy. As I described in the preceding literature review, I was unable to locate any empirical studies that investigated the relations between motivational self-beliefs and questioning behavior. There are theoretical reasons to suggest, if tentatively, that personal interest, mastery-approach goal orientation and perceived teacher emphasis on mastery goals may be positively associated with high quality questions and high self-appraisals of questions. As I describe in Chapter 3, motivational self-beliefs and perceived teacher emphasis on mastery goals are likely to be correlated with each other and, thus, may have overlapping relationships with the quantity, qualities, and self-appraisals of questions.

I ask a third primary research question examines the relations among achievement, questions, and motivations and is posed in two parts: (RQ3a) Can exam scores be predicted by the quantity and qualities of questions? and (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and Exam 2 scores? Regarding the RQ3a, there is some evidence to suggest that self-questioning may be positively associated with achievement. Yet as I described previously, students' questions may have a complex relationship with learning outcomes since questions may indicate both knowledge deficits as well as potential for the remediation of such deficits. Regarding RQ3b, characteristics of students' questions may be associated with end-of-class perceptions, such as self-efficacy beliefs or perceived comprehension of the material, which may in turn predict achievement.

A fourth primary research question (RQ4) asks: Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, help seeking beliefs, achievement, and the quantity, quality and self-appraisals of generated questions? RQ4 has the potential to extend the existing literature by examining (1) help seeking *behavior* instead of self-reported *intentions* to seek help and (2) the relationship between generated and posed questions in terms of their quantity, quality and self-appraisals. There are theoretical reasons to suggest that students who ask questions may have higher scores on personal interest, mastery goal orientation, perceived teacher emphasis on mastery goals, instrumental help seeking and formal help seeking target measures and lower scores on performance-avoid goal orientation, help seeking avoidance, and help seeking threat measures compared with students who do not ask questions in class.

CHAPTER 3

METHODS

This study was designed to examine the questions that students generated during chemistry lectures and their relations with motivations, in-class question asking, and achievement. The study proposed four central research questions. (RQ1) What are the relations among the quantity, qualities, and self-appraisals of student-generated questions? (RQ2) What are the relations among motivational self-beliefs, perceived teacher emphasis on mastery goals, and the quantity, qualities, and self-appraisals of student-generated questions? (RQ3a) Can exam scores be predicted by the quantity and qualities of questions? (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and Exam 2 scores? (RQ4) Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, achievement, and the quantity, qualities, and self-appraisals of generated questions?

Research Design

The study used multiple data sources to investigate the research questions, including two waves of questionnaires, question logs in which students recorded their questions during class meetings, course exams, and field observations. The design of the study intended to stimulate students' awareness of their questions during a typical learning activity and to make students' questions observable by minimizing barriers to asking or sharing them. The methods used in this study fall somewhere between a naturalistic study and a conventional intervention study. The methods were not strictly naturalistic because one of the principal data collection instruments—question logs—

likely influence questioning behaviors themselves. The act of asking students to record their questions during lectures *intervened* on the classroom environment and students' experiences while learning, presumably influencing aspects of student questioning. Yet the methods are different from much interventional educational research because the study did not include a control or comparison group. The purpose of the study was not to test the effectiveness of an intervention in which students recorded their questions. Rather, the purpose was to examine questioning in a context that encouraged students to attend to, record and communicate their questions to the instructor in a way that minimized logistical, psychological, and social barriers to asking them during class. As described in the literature review, the instructional practice of having students focus on their questions during learning is an effective learning and instructional strategy (e.g., King, 1994; Meister & Chapman, 1996). Accordingly, the study was designed to examine student questioning via research tools that are themselves pedagogically justifiable. The results are discussed and should be interpreted in light of how the data were collected.

The Learning Task and Instructional Context

General Chemistry I was a 3-credit hour large-enrollment introductory chemistry course for STEM majors offered by College of Science and Technology at Temple University. General Chemistry I met Mondays, Wednesdays and Fridays from 8:00-8:50 AM during the 15-week spring semester. Enrollment in a college level math course was a co-requisite for General Chemistry I. Passing the course was a pre-requisite for intermediate and higher-level courses in STEM majors (e.g., General Chemistry II, Organic Chemistry, etc.). The course had a reputation of being academically rigorous and challenging. Relative to norms across the University, General Chemistry I had a high drop/withdraw/failure rate (over 20%).

Students enrolled in the General Chemistry I lecture course were also enrolled in a recitation section, which met once each week for 50 min and had a maximum class size of 26 students. The recitation sections, which were taught by graduate teaching assistants, were designed to reinforce and clarify material presented during the lecture course.

The instructor of the lecture course had a Ph.D in chemistry and was a fulltime teaching faculty member in the College of Science and Technology. In addition to designing and delivering the chemistry lectures, the instructor supervised the teaching assistants who taught the recitation sections. The instructor, a middle-aged White male, presented himself casually during lectures, commonly wearing a flannel shirt, jeans, and a pair of hiking boots. In his interactions with students, the instructor was friendly and causal, peppering his lectures with his quirky sense of humor (e.g., referring to the *molar mass* equation as the *Mickey Mouse* equation and impersonating a befuddled Bart Simpson stumped by a problem he was solving).

Instructional time consisted nearly exclusively of the instructor presenting information. Student occasionally participated by shouting out or mumbling answers to leading questions posed by the instructor (e.g., *The state of this element is what?...Solid, right?*). The instructor conveyed a sense of urgency to cover the instructional materials and keep on the schedule laid out in the syllabus. The instructor began class meetings promptly at the start of the hour and often continued talking a minute or two past the bell at the end of class. The 50-min lectures were fast-paced with the instructor moving briskly between topics, examples, and problems.

Most commonly, class meetings consisted of the instructor solving problems and writing key points on the large white boards at the front of the lecture hall. The instructor also occasionally presented information using commercial PowerPoint slides that accompanied the textbook; students had access to electronic copies of these slides and the instructor encouraged students to bring them to class meetings. During one class meeting, the instructor taught from a photocopied handout, which he provided in hardcopy to students and which was projected in the front of the lecture hall.

Instruction emphasized the application of conceptual and procedural knowledge to solve well-defined chemistry problems with objective answers. The 8 lectures covered a wide range of concepts, included limiting reactants, theoretical and percent yield, molarity, stoichiometry, aqueous solutions, precipitation reactions, acid-base reactions, pressure and partial pressures, properties of gases and gas laws, kinetic-molecular theory, diffusion and effusion, energy, the First Law of Thermodynamics, and heat and work. The majority of the instructional time (66%) consisted of the instructor modeling how to solve chemistry problems which applied these concepts and which were taken directly from the students' textbooks. For example, a typical question that was solved during class time was: *The reaction of 167g* Fe_2O_3 *with 85.8g CO produces 72.3g Fe. Find the limiting reagent, theoretical yield and percent yield*. Following the schema used in the textbook, the instructor commonly progressed through the following sequence as he modeled problem solving: 1) Identifying what the objective of the problem was (e.g., in the problem above, calculating the theoretical and percent yield and identifying the limiting reactant); 2) creating a problem-solving plan (e.g., [step 1] calculate the mass of the limiting reactant by converting the mass of each reactant to moles; [step 2] use stoichiometry to convert moles of reactants to moles of products; [step 3] divide this theoretical yield by the actual yield to calculate the percent yield); and 3) executing the plan in a series of steps. The instructional focus on solving problem was consistent with homework, quizzes, and exams that assessed the application of conceptual and procedural knowledge to solve chemistry word problems.

Participants

Participants were 103 students enrolled in General Chemistry I during the spring 2013 semester. Initially, 105 students enrolled in the study; however, 2 students did not complete a single question log and were therefore dropped from analyses. Demographics are presented in Table 3.1. Participants constituted 57% of the total number of student enrolled at the start of the semester.

Characteristics of the Sample

Age, Sex, and Race/Ethnicity

Participants had a mean age of 20 years (*SD* = 2.47) and were 55% female. The sample was racially diverse; participants were 43% White, 18% Asian, 12% Black, 6% Indian subcontinent, 5% Latino, 4% Middle Eastern, and 13% identified as a different race or more than one race (see Table 3.1).

Parental Education

Parental education level was collected as a proxy for socio-economic status. Students' reported the highest degree attained for their mother and father. The modal level of highest education for mothers and fathers was a high school diploma; participants reported that 21% of mothers and 24% of fathers had achieved a Bachelor's degree or higher (see Table 3.1).

School Year

Most students in the sample were in their first year of college (55%), followed by third-year (26.5%), second-year (13.7%), post-baccalaureate (4%), and fourth-year (1%). It is unclear why there were more juniors than sophomores in the study since students tend to take General Chemistry I towards the beginning of their college experience. The relatively large number of juniors may be the result of the atypical course offering of General Chemistry in the spring (a point discussed in the Course Sequence section below).

Academic Preparation

Regarding students' math preparation, students were asked to report all of the math courses they took in high school. Math preparation is relevant for General Chemistry I because much of the instructional time and assessments focused on students' ability to solve quantitative chemistry problems, which require mathematical (especially algebraic) skills. A large majority of students (89%) reported having taken Trigonometry in high school; 43% of students reported having taken a calculus class in high school.

Educational Aspirations

Students were asked to report the highest educational degree to which they aspired. In this sample, students reported high educational aspirations, with 16% aspiring to a master's degree and 78% aspiring to a professional (e.g., MD, DPharm, DDS) or doctoral degree (see Table 3.1).

Course Sequence

The typical course sequence for science majors is to take General Chemistry I in the fall followed by General Chemistry II in the spring. Students who take General Chemistry I in the spring—the semester this study was conducted—may be enrolled in the atypical course sequence for several reasons. A passing grade in General Chemistry I is a pre-requisite for students to enroll in General Chemistry II; students who did not pass General Chemistry I in the fall may re-enroll in the course in the spring. In this sample, 25% percent of students reported they had previously enrolled in the course but had not passed or completed it (i.e., dropped, withdrew, or failed). Students may have also taken the course in an atypical sequence in order to first fulfill math pre-requisites for General Chemistry I. Other students may be enrolled in General Chemistry I in the spring semester for a range of other scheduling issues.

Table 3.1

Participants' Sex, School Year, Parent Education, Educational Preparation and Aspirations

Category	*n*	Percentage of sample
Gender		
Male	46	44.7
Female	57	55.3
No response	0	0.0
Academic year		
First year	56	54.9
Second year	14	13.7
Third year	27	26.5
Fourth year	1	1.0
Post-baccalaureate	4	3.9
No response	1	1.0
Race/ethnicity		
Asian	18	17.6
Black	12	11.8
White	44	43.1
Latino	5	4.9
Indian Subcontinent	6	5.9
Middle Eastern	4	3.9
No response	1	0.9

(Continued Table 3.1)

Category	*n*	Percentage of sample
Mother's education		
Some high school	7	6.9
High school diploma	33	32.4
Some college	23	22.5
Bachelor's degree	21	20.6
Masters or higher	18	17.6
No response	1	0.9
Father's education		
Some high school	7	6.9
High school diploma	30	29.7
Some college	18	17.8
Bachelor's degree	24	23.8
Masters or higher	22	21.8
No response	2	
High school math preparation		
Less than trigonometry	11	10.9
At least trigonometry	90	89.1
Calculus	43	42.6
No response	2	1.9

(Continued Table 3.1)

Category	*n*	Percentage of sample
Educational aspirations		
Some college	3	2.9
Bachelor's degree	4	3.9
Master's	16	15.7
Professional degree	39	38.2
Doctoral degree	40	39.2
No response	1	0.9

Prior Achievement

Table 3.2 presents students' academic achievement prior to taking General Chemistry I as measured by self-reported Grade Point Average (GPA) on a 4-point scale. Students were asked to report their cumulative college GPA from the most recent semester preceding the study. The mean GPA for the sample was 2.99 (*SD* = .60).

Students were also asked to report their scores from the Verbal and Quantitative sections of the Standard Achievement Test (SAT). The mean SAT-Quantitative score for the sample was 590.93 (*SD* = 73.48) and the mean SAT-Verbal score was 555.35 (*SD* = 93.92). Note, however, that a large number of students (58.3%) reported that they did not recall their SAT scores or left the field blank. Due to the large amount of missingness, SAT scores were not included in analyses and the descriptive statistics may not represent the sample as a whole.

Table 3.2

Prior Academic Achievement for the Sample

Measure	*n*	% of sample	*M*	*SD*	Range
GPA	98	95.1%	2.99	0.60	0.89 - 3.93
SAT–Quantitative	43	41.7%	590.93	73.48	440 - 740
SAT–Verbal	43	41.7%	555.35	93.92	340 - 780

Note. GPA = Grade Point Average with the following scale: A = 4.00, A- = 3.67, B+ = 3.33, B = 3.0, B- = 2.67, C+ = 2.33, C = 2.00, C- = 1.67; D+ = 1.33; D = 1.00; D- 0.67.

Measures, Materials, and Data Sources

I assessed several motivational constructs, question characteristics, and achievement. Regarding motivational constructs, I assessed personal interest in chemistry; four individual achievement goal orientations; perceived teacher emphasis on mastery goals, and five help seeking beliefs. Given the scarcity of empirical studies examining the relations among motivations and student-generated questions in classroom contexts, I intentionally selected a broad range of motivational constructs. I selected these particular constructs because of their theoretical relations with question-generation processes, as I described in Chapter 2, and due to their prominence in current research on motivation. As I describe below, all motivational measures had been previously validated and demonstrate good reliability with undergraduate samples. In addition, I assessed students' self-reported situational interest, self-efficacy, and perceived comprehension at the end of class meetings. Regarding behavioral variables, I assessed various characteristics of student-generated questions, which students recorded using question logs. In addition, I conducted classroom observations with the aim of documenting the learning environment and question-asking behavior. Achievement variables consisted of raw (i.e., non-curved) grades for Exams 1 and 2 in General Chemistry I. Measures, sources, number of items, and estimated times for completing questionnaire scales are presented in Table 3.3. Below I describe in detail the nature, scoring, and validity and reliability of these measures.

Table 3.3

Information on Measures

Scale	Source	No. items	Approx. Min.
Wave 1			
Informed consent	Researcher-generated	n/a	3
Demographics questionnaire	Researcher-generated	14	5
Personal interest in chemistry	Linnenbrink-Garcia et al, 2010	8	4
Mastery-approach G.O.	Elliot & Murayama, 2008	3	1
Mastery-avoidance G.O.	Elliot & Murayama, 2008	3	1
Performance-approach G.O.	Elliot & Murayama, 2008	3	1
Performance-avoidance G.O.	Elliot & Murayama, 2008	3	1
Teacher Emph. Mastery Goals	Midgley et al., 2000	5	3
Self-efficacy in chemistry	Midgley et al., 2000	5	2
Instrumental help seeking	Karabenick, 2003	5	3
Executive help seeking	Karabenick, 2003	3	1
Help seeking threat	Karabenick, 2003	3	1
Help seeking avoidance	Karabenick, 2003	3	1
Help seeking target	Karabenick, 2003	3	1
Intention to persist in STEM	Perez et al., 2013	6	3
Training on question log	Researcher-generated	n/a	10
Total T1 (Completed in computer lab)		57	40

Note. G.O. = Goal orientation; Emph = Emphasis

(Continued Table 3.3)

Scale	Source	No. items	Approx. Min.
Wave 2			
Personal interest in chemistry	Linnenbrink-Garcia et al, 2010	8	4
Intention to persist in STEM	Perez et al., 2013	6	3
Self-efficacy in chemistry	Midgley et al., 2000	5	2
Mastery-approach G.O.	Elliot & Murayama, 2008	3	1
Mastery-avoidance G.O.	Elliot & Murayama, 2008	3	1
Performance-approach G.O.	Elliot & Murayama, 2008	3	1
Performance-avoidance G.O.	Elliot & Murayama, 2008	3	1
Teacher Emph. Mastery Goals	Midgley et al., 2000	6	3
Total T2 (Completed via e-mail)		37	16
Total T1 and T2		105	57

Note. Approximate times based on research with similar population (Cromley et al., 2011; Perez, 2012)

Personal Interest in Chemistry

I measured personal interest in chemistry using the Individual Interest scale developed by Linnenbrink-Garcia et al. (2010). Items were originally developed to assess interest in math; therefore, I adapted items to refer to interest in chemistry. For example, the original items *Math helps me in my daily life outside of school* was changed to *Chemistry helps me in my daily life outside of school.* Adapted items are presented in Appendix A, as they appeared in the Wave 1 Questionnaire. Personal interest can be measured in many different ways and researchers disagree on the nature of the construct (for a review, see Renninger & Hidi, 2011). I selected Linnenbrink-Garcia et al.'s measure because it taps affective (e.g., enjoyment) and value (i.e., utility, importance) components, both of which may be related to questioning and evaluation processes.

Scoring

Consistent with Linnenbrink-Garcia et al. (2010), participants responded to items using a scale that ranges from 1 ("Not at all true") to 5 ("Very true"). Thus, higher scores indicated greater individual interest for chemistry. Raw scores for each item within a scale were used to create a mean score.

Validity and Reliability

The Linnenbrink-Garcia et al. (2010) measure of individual interest was an adaptation of the task-value scale of the Motivated Strategies for Learning Questionnaire (MSQL; Pintrich & DeGroot, 1990). The scale was developed in line with Schiefele's conceptualization of interest as consisting of both value and affect components. Hence, the measure has face validity as items—as evident in the examples above—tap both value (i.e. utility and importance) and affect. Individual interest was related to but distinct from related motivational constructs such as situational interest. Linnenbrink-Garcia et al. (2010) reported a strong correlation between individual and situational interest (r = .61). I was unable to locate studies that reported the relationship between this measure of individual interest and self-efficacy or goal orientations. Linnenbrink-Garcia et al. reported strong internal consistency (Cronbach's alpha = .90) with a sample of 284 secondary students.

Achievement Goal Orientations

I measured individual goal orientations using four scales from the Achievement Goal Questionnaire (Revised) (AGQ-R; Elliot & Murayama, 2008): mastery-approach (MAP), mastery-avoidance (MAV), performance-approach (PAP) and performance-avoidance (PAV) scales. Each scale consisted of 3 items. An example of a MAP item was *My goal is to learn as much as possible.* An example of a MAV item was *My aim is to avoid learning less than I possibly could.* An example of a PAP item was *My aim is to perform well relative to other students.* An example of a PAV item was *My aim is to avoid doing worse than other students*. Complete scales for MAP, MAV, PAP, and PAV are presented in Appendices B, C, D and E, respectively. The instructions to the AGQ-R subscales asked participants to consider the General Chemistry I course as the context of the question.

Scoring

Consistent with Elliot and Murayama (2008), participants responded to items using a scale that ranged from 1 (“strongly disagree) to 5 (“strongly agree”). Thus, higher scores indicated greater endorsement for the construct. For example, the higher the numeric response on a MAP item, the more the respondent has endorsed a mastery orientation. Raw scores for each item within a scale were used to create a mean score.

Validity and Reliability

Elliot and Murayama (2008) argued that the AGQ-R constituted an improvement over previous measures of achievement goals because of several features. First, all items in all subscales explicitly referred to an aim or goal. The intention was to differentiate the goal orientation from the value driving it (e.g., value for mastery vs. goal for mastery), a modification that the authors argued is in line with original conceptualizations of goal orientations. Second, AGQ-R items minimized the inadvertent measurement of relative competence by avoiding extreme qualifiers such as *most* or *least* in favor of more general qualifiers such as *more* or *less.* Third, items in the AGQ-R did not explicitly refer to affect. While affect was related to achievement goals via investment of the individual in a future outcome, Elliot and Murayama argued it is not the primary focus of achievement goals. Fourth, AGQ-R avoided language that pits performance and mastery goals against each other, thereby assuming that mastery and performance goal orientations may be positively correlated with each other. Fifth, AGQ-R avoided language that could be applicable to both mastery and performance goals.

With regard to the proposed study, the AGQ-R has been normed with undergraduate populations whereas other frequently used measure of goal orientations such as PALS (Midgley et al., 2000) were designed and primarily tested with younger learners. Elliot and Murayama (2008) investigated the reliability and factor structure of the AGQ-R with a sample of 229 undergraduate students enrolled in an introduction to psychology course. The authors reported that the scales of the AGQ-R had high internal consistency with Cronbach's alpha of .84, .88, .92 and .94 for MAP, MAV, PAP and PAV scales, respectively. Confirmatory factor analyses suggested that the scales loaded on different factors, reflecting the theorized 2 x 2 model of achievement goals (Elliot & McGregor, 2001; Pintrich, 2000). Furthermore, the authors argued that the AGQ-R had predictive utility for theorized antecedents (i.e., need for achievement, fear of failure) and consequences (e.g., intrinsic motivation, grades; see Elliot & Murayama, 2008 for a complete review). Note that males and non-white ethnicities were under-represented in this sample compared with the sample of the proposed study. With regard to gender, a subsequent study with Omani undergraduates in a college of education found no evidence of measurement invariance between males and females (Alkharusi & Aldhafri, 2010). The authors reported similar Cronbach's alphas for AGQ-R scales.

Perceived Teacher Emphasis on Mastery Goals

I assessed teacher emphasis on mastery goals using the Teacher Mastery Goals scale from PALS (Midgley et al., 2000). The complete scale is presented in Appendix A as they appeared in the Wave 1 Questionnaire. The TMG scale included 6 items. I adapted items to refer specifically to the chemistry lecture course, congruent with adaptations of the PALS for secondary school students (Midgley et al., 2000). The term *chemistry lecture* will be used instead of the more general term *chemistry* to focus participants' thinking about the specific emphasis of the teacher on mastery goals *during lectures* since other chemistry contexts such as recitation and lab meetings may be perceived to have different goal structures. For example, the item with the originally wording *In our class, trying hard is very important* was changed to *In my chemistry lecture course, trying hard is very important.* I selected the PALS scale of teacher goal structure because it is a commonly used measure of this construct and has been validated using cognitive interviewing (Koskey et al., 2010).

Scoring

I used the original response range and anchors suggested by Midgley et al., (2000). Participants selected responses on a 5-point Likert-like scale with anchors at 1 ("Not at all true"), 3 ("Somewhat true") and 5 ("Very true"). Thus, higher scores indicated greater endorsement of the measured construct. Raw scores for each item within a scale were used to create a mean score.

Validity and Reliability

The measure of perceived teacher emphasis on mastery goals was designed to tap individuals' perceptions of their environment in terms of its support for developing mastery. As such, items had face validity because they focus attention on the context in relation to support for a mastery goal orientation. For example, the item *In our class, it's OK to make mistakes as long as you are learning* focuses attention on the class in terms of its support for learning despite mistakes, a characteristic of mastery goal orientation. The teacher mastery goals scale has been largely used with research on elementary and secondary students. Midgley et al. (2000) reported a Cronbach's alpha of .70.

Teacher goals are theorized to be precursors to personal goal orientations (Anderman & Maehr, 1994; Meece, E. M. Anderman & L. H. Anderman, 2006). Thus, a perceived teacher emphasis on mastery goals is related to but distinct from a personal mastery goal orientation. Friedel, Cortina, Turner and Midgley (2007) used the PALS teacher goal measures to examine the relations between personal goal orientations and perceived teacher emphasis of goals for 1,021 grade 7 students. The authors reported medium-sized, significant correlations between personal goal orientations and analogous teacher goals. For example, mastery-approach goal orientation and perceived teacher emphasis on mastery goals were correlated at r = .36.

Self-Efficacy for Chemistry

I measured self-efficacy for chemistry using the PALS academic efficacy scale (Midgley et al., 2000). The PALS scale consisted of 5 items, which are presented in

Appendix A, as they appeared in the Wave 1 Questionnaire. The original wording of items were adapted to specifically refer to content taught in chemistry lectures. For example, the original wording of the item *I'm certain I can master the skills taught in my classes this semester* was changed to *I'm certain I can master the skills taught in my chemistry lecture course this semester*. As with the teacher mastery goal scale, the term *chemistry lecture course* was used to focus attention on the specific context of the proposed study.

Scoring

I used the original response range and anchors suggested by Midgley et al., (2000). Participants selected responses on a 5-point Likert-like scale with anchors at 1 (“Not at all true”), 3 (“Somewhat true”) and 5 (“Very true”). Thus, higher scores indicated greater academic self-efficacy. Raw scores were used to create a mean score.

Validity and Reliability

Academic self-efficacy refers to an individual’s belief that he or she will be able to be successful at task at designated levels. PALS items had face validity in that they address students’ ability to achieve, even in the face of difficulty or challenge. Designated levels for success were not included in these items; however, this is consistent with other widely-used questionnaires that portend to tap this motivational construct (e.g., Motivated Strategies for Learning Questionnaire; Pintrich & DeGroot, 1990; see Bong & Hocevar, 2002 for a discussion). The PALS scale of academic efficacy has demonstrated high reliability. In a sample very similar to the one used in the present study, Perez (2012) reported a Cronbach’s alpha of .93.

The construct of self-efficacy is a distinct from other motivational constructs included in the proposed study. In a sample of 240 eighth and tenth grade Norwegian students, Diseth, Danielsen, and Samdal (2012) reported significant correlations between self-efficacy (as measured by an adaptation of the PALS measure) and mastery goal orientation, performance goal orientations and achievement (i.e., self-reported grades). Self-efficacy was strongly correlated with achievement (r = .56) and mastery goal orientation (r = .61) and had a moderate to small correlation with performance goal orientation (r = .24).

Instrumental Help Seeking

I measured instrumental help seeking using 5 items. Karabenick (2003) developed two items assessing this construct. I adapted these two items to specifically refer to the chemistry lecture course (following the rationale for this change described in previous measures). For example, the item with original wording *Getting help would be one of the first things I would do if I were having trouble in my class* was changed to *Getting help would be one of the first things I would do if I were having trouble in my chemistry lecture course.* In order to improve reliability of this scale (discussed below), I created three new items. The items follow the same two-part structure as the existing items: they include (1) a subjunctive clause describing a situation in which one could need help paired with (2) a declarative statement about seeking needed help. For example, one new item is *If I were having difficulty learning the material in my chemistry lecture course I would find someone who could help me.* The adapted and new items appear in Appendix A, as they appeared in the Wave 1 Questionnaire.

Scoring

I used the original response range and anchors used by Karabenick (2003). Participants selected responses on a 5-point Likert-like scale with anchors at 1 ("Not at all true"), 3 ("Somewhat true") and 5 ("Very true"). Thus, higher scores indicated greater instrumental help seeking. Raw scores were used to create a mean score.

Validity and Reliability

This measure of instrumental help seeking followed Nelson Le-Gall's (1981; 1985) distinction between help seeking when it is needed (instrumental help seeking) and help seeking that seeks to minimize one's time and effort (executive help seeking). The new and adapted items had face validity for instrumental help seeking in that they asked about a participant's willingness to seek help when it is needed. Karabenick (2003) reported somewhat low reliability with the two-item scale: a Cronbach's alpha of .62 with an undergraduate sample; Cronbach's alpha values for scales with only two items, however, are hard to interpret. The addition of the 3 new items may result in a more reliable measure. Cronbach's alphas were calculated for the new scale, including changes in alphas if certain items are dropped (see Table 4.1).

The construct of instrumental help seeking is distinct from but related to other help seeking constructs. Karabenick (2003) reported significant negative correlations with help seeking threat (r = -.26), help seeking avoidance (r = -.39) and executive help seeking (r = -.16) and a small but significant positive correlation with seeking help from a formal target (r = .17).

Executive Help Seeking

I measured executive help seeking using 3 items developed by Karabenick (2003). I adapted these items to specifically refer to the chemistry lecture course following the rationale provided previously. An example item was *I seek help in my chemistry lecture course when it means I have to work less.* The adapted and new items appear in Appendix A, as they appeared in the Wave 1 Questionnaire.

Scoring

I used the original response range and anchors used by Karabenick (2003). Participants selected responses on a 5-point Likert-like scale with anchors at 1 ("Not at all true"), 3 ("Somewhat true") and 5 ("Very true"). Thus, higher scores indicated greater executive help seeking. Raw scores were used to create a mean score.

Validity and Reliability

As described in the preceding section on instrumental help seeking, executive help seeking describes help seeking behavior that aims to minimize one's time and effort. The adapted items had face validity in that they ask about a participant's willingness to seek help in order to lower the effort for the individual. Karabenick (2003) reported good reliability with a sample of undergraduate students (Cronbach's alpha = .78).

The construct of executive help seeking is distinct from but related to other help seeking constructs. Karabenick (2003) reported significant positive correlations with help seeking threat (r = .52) and help seeking avoidance (r = .54), a small but significant negative correlation with instrumental help seeking (r = -.16), and no significant relationship with seeking help from a formal target.

Help Seeking Threat

I measured help seeking threat using 3 items developed by Karabenick (2003). I adapted these items to specifically refer to the chemistry lecture course following the rationale provided above. An example item was *I would feel like a failure if I needed help in my chemistry lecture course.* The adapted items appear in Appendix A, as they appeared in the Wave 1 Questionnaire.

Scoring

I used the original response range and anchors used by Karabenick (2003). Participants selected responses on a 5-point Likert-like scale with anchors at 1 (“Not at all true”), 3 (“Somewhat true”) and 5 (“Very true”). Thus, higher scores indicated greater help seeking threat. Raw scores were used to create a mean score.

Validity and Reliability

Help seeking threat refers to the perceived undesirability of seeking help (Ryan et al., 2001). The adapted items had face validity in that they address the extent to which a participant perceives seeking help as threatening or undesirable. Karabenick (2003) reported good reliability with an undergraduate sample (Cronbach’s alpha = .81).

The construct of help seeking threat is distinct from but related to other help seeking constructs. Karabenick (2003) reported significant positive correlations with help seeking avoidance (r = .69) and executive help seeking (r = .52), a moderate negative correlation with instrumental help seeking (r = -.26), and no significant relationship with seeking help from a formal target.

Help Seeking Avoidance

I measured help seeking avoidance using 3 items developed by Karabenick (2003). I adapted these items to specifically refer to the chemistry lecture course following the rationale provided above. An example item was *Even if the work was too hard to do on my own, I wouldn't ask for help with my chemistry lecture course.* The adapted items appear in Appendix A, as they appeared in the Wave 1 Questionnaire.

Scoring

I used the original response range and anchors used by Karabenick (2003). Participants selected responses on a 5-point Likert-like scale with anchors at 1 ("Not at all true"), 3 ("Somewhat true") and 5 ("Very true"). Thus, higher scores indicated greater help seeking avoidance. Raw scores were used to create a mean score.

Validity and Reliability

Help seeking avoidance refers to the avoidance of seeking help from others in situations in which help is needed (Ryan & Pintrich, 1997; Ryan et al., 2001). The adapted items had face validity in that they address a participant's intent to avoid help in a scenario in which help is needed. Karabenick (2003) reported a Cronbach's alpha of .77 with an undergraduate sample.

The construct of help seeking avoidance is distinct from but related to other help seeking constructs. Karabenick (2003) reported significant and large positive correlations with help seeking threat (r = .69) and executive help seeking (r = .54), a moderate negative correlation with instrumental help seeking (r = -.39), and no significant relationship with seeking help from a formal target.

Preference for a Formal Help Seeking Target

I measured the preference for a formal help seeking target using 3 items developed by Karabenick (2003). I adapted these items to specifically refer to the chemistry lecture course following the rationale provided above. An example item was *If I were to seek help in my chemistry lecture course, I would ask the teacher rather than another student.* The adapted items appear in Appendix A, as they appeared in the Wave 1 Questionnaire.

Scoring

I used the original response range and anchors used by Karabenick (2003). Participants selected responses on a 5-point Likert-like scale with anchors at 1 (“Not at all true”), 3 (“Somewhat true”) and 5 (“Very true”). Thus, higher scores indicated a greater preference for seeking help from a formal source compared to an informal source. Raw scores were used to create a mean score.

Validity and Reliability

Help seeking target refers to a participant’s preference for seeking help from formal sources such as the teacher compared with informal sources such as peers. Karabenick (2003) argued for the theoretical distinction between willingness to seek help and the help seeking target, as each addresses a distinct decision point in help seeking behavior (Karabenick & Dembo, 2011; Newman, 1990; 1994; 1998). The adapted items had face validity in that they address a participant’s intention to seek help from the teacher over students. Karabenick (2003) reported a Cronbach’s alpha of .66 with an undergraduate sample.

The construct of seeking help from a formal target is distinct from other help seeking constructs. Karabenick (2003) reported a small but significant positive correlation between seeking help from a formal target and instrumental help seeking (r = .17) and no significant relationships with executive help seeking, avoidance of help seeking or threat to help seeking.

Intention to Persist in a STEM Major

I measured participants' intention to persist in a STEM major using a scale developed by Perez, Cromley, & Kaplan (2013). The scale consisted of 6 items that address a student's intention to persist in or switch out of a STEM major. An example item was: *At the present time, how likely are you to remain in (or select) a science major?* The complete scale is presented in Appendix A, as they appeared in the Wave 1 Questionnaire. This construct was included in the study at the recommendation of the dissertation proposal committee and because intentions for future course selection has been a variable of interest in motivational research in STEM (e.g., Perez et al., 2013; Watt, 2006).

Scoring

I used the original response anchors used by Perez et al. (2013). Participants selected responses on a 5-point Likert-like scale with anchors at 1 (e.g., "Not at all true of me") to 5 (e.g., "Very true of me"). I aggregated this scale to be a measure of intention to *persist* in a STEM major in contrast to Perez's focus on the intention to *leave* a STEM major. Thus, items 1, 3, 4 and 5 were reverse coded. Higher scores indicated a greater intention to *remain in* a STEM major. Raw scores were used to create a mean score.

Validity and Reliability

The measure of intention to persist in a STEM major was designed to tap a participant's desire to remain in or, in cases in which the student has not already declared a STEM major, switch into a STEM major. Perez (2012) reported that the scale demonstrated strong reliability (Cronbach's alpha = .93) with a similar sample at the same institution.

Intention to persist in a STEM major is related to but distinct from motivational variables. For example, Perez reported that intention to *leave* a STEM major at Time 2 (i.e., in the middle of the semester) was significantly correlated with self-efficacy for science (r = -.24) and interest in science (r = -.28) at Time 1 (at the beginning of the semester).

Situational Interest

I assessed situational interest at the end of each class meeting with a single-item. Students were asked to respond to the following statement: *I found today's class interesting.* Appendix D illustrates how the item appeared in the question log.

Scoring

Participants selected responses on a 10-point scale anchored at *Not true of me* (1) to *Very True of Me* (10). A higher score indicated a higher level of situational interest. Each post-class perception was aggregated by calculating the mean score across all question logs completed by the participant. Therefore, the situational interest score represented the mean level of situational interest across all class meetings. Mean scores could potentially range from 1 to 10.

Validity and Reliability

Single-item scales measuring situational interest are one way researchers have attempted to measure situational interest during learning (Ainley & Hidi, 2002; Ainley & Patrick, 2006; Shernoff, Csikszentmihalyi, Schneider, & Steele Shernoff, 2003; for a review, see Renninger & Hidi, 2011). Single-item scales are useful for measuring situated phenomenological experiences during authentic learning settings with minimal disruption (Ainley & Patrick, 2006). Determining the reliability of single-item scales that measure phenomenological experiences in specific situations—such as reported situational interest during a specific class—is difficult (Ainley & Patrick). Common measures of reliability such as internal consistency are not possible with one item and variation in scores across measurement points may indicate actual variations in situational interest from situation to situation rather than measurement error. Ainley and Patrick argued "when the purpose of the measure is clear and unambiguous for the respondent and has a clear experiential focus, [single-item measures] can provide valid indicators of the state being investigated" (p. 282).

Situational interest is theorized to be related to but distinct from personal interest (Schiefele, 2009; Hidi & Renninger, 2006) and thus, measures of situated interest tend to correlate positively with personal interest. For example, Linnenbrink-Garcia et al. (2010) reported a strong correlation between situational and individual interest (r = .61).

Post-Class Self-Efficacy

I assessed self-efficacy beliefs at the end of each class meetings using a single-item. Students responded to the following statement: *I am certain that I can master the* *material from today's lesson.* Appendix D illustrates how the item appeared in the question log.

Scoring

Participants selected responses on a 10-point scale anchored at *Not true of me* (1) to *Very True of Me* (10). A higher score indicated a higher level of efficacy for learning the instructed material. Each post-class self-efficacy rating was aggregated by calculating the mean score across all question logs completed by the participant. Therefore, the post-class self-efficacy score represented the mean level of self-efficacy after lectures across all class meetings. Mean scores could potentially range from 1 to 10.

Validity and Reliability

Self-efficacy is most commonly measured using multi-item scales with questionnaires (Usher & Pajares, 2008), but has been measured in other ways, including with one- or two-item scales and over multiple time periods or tasks (e.g., Rebok & Balcerak, 1989; Schmitz & Wiese, 2006). The post-class self-efficacy scale differed from the self-efficacy for chemistry scale (described previously) because the former measured efficacy beliefs about a more specific task (i.e., mastering the instructed material on a given day) while the latter measured more general efficacy beliefs about succeeding in the course (e.g., *I'm certain I can master the skills taught in my chemistry lecture course this semester*). Prior research has found that compared to task-general efficacy beliefs, task-specific efficacy beliefs are more predictive of achievement (Bong, 2004; Pajares, 1996).

Post-class self-efficacy perceptions are distinct from but related to other situated perceptions. For example, self-evaluations such as assessing one's level of comprehension (perceived comprehension, described below) are related to efficacy beliefs, since perceptions of mastery can lead to increases in efficacy beliefs (Schunk, 2003).

Perceived Comprehension

I assessed the students' evaluation of their comprehension during lectures with a single item after each class meeting. Students were asked to respond to the following statement: *Today, I really understood the content of the lecture.* Appendix D illustrates how the item appeared in the question log.

Scoring

Participants selected responses on a 10-point scale anchored at *Not true of me* (1) to *Very True of Me* (10). A higher score indicated a higher level of perceived comprehension of the instructed material. Each post-class perceived comprehension rating was aggregated by calculating the mean score across all question logs completed by the participant. Therefore, the perceived comprehension score represented the mean level of self-reported understanding of the instructed material across all class meetings. Mean scores could potentially range from 1 to 10.

Validity and Reliability

Researchers have used single items to have students self-evaluate their learning on specific tasks (e.g., Dunlosky & Rawson, 2012). While self-evaluations may not reflect accurate appraisals of actual learning (e.g., Dunlosky & Rawson), the perceived comprehension item has face validity in that asks students to report how well they thought they understood the lecture. Self-evaluations are related to but distinct from related constructs. Self-evaluations are related to self-efficacy beliefs, since positive self-evaluations (e.g., that one is learning, understanding, or improving) can produce increased self-efficacy beliefs (Schunk, 2003).

Recording Student-Generated Questions in Question Logs

Participating students were asked to complete a question log during each class meeting during the target unit. A double-sided question log (presented in Appendix D) was placed on each student's desk before the start of the lecture. The question log contained three sections that participants completed during the class meeting. In the first section, students wrote their Temple University Identification Number (TUID), their first name, and the date. Before the start of class and during in-class announcements, members of the research team reminded students to enter this identifying information. Upon collecting question logs at the end of class meetings, members of the research team confirmed that, at minimum, the student had recorded his or her TUID.

In the second section of the question log, students wrote their questions and completed several pieces of information relating to their questions. First, students were asked to record the contents of each question. Second, students were asked to record the time of their question according to the large clock that was placed at the front of the classroom. Third, students were asked to appraise each question on two dimensions: *Value for Self* and *Value for Others* (described below). Fourth, students were asked to check a box if a question that had been recorded was resolved or answered by the end of a class meeting. Students received these instructions in writing on the question log, via verbal announcements during class, and via instructions and feedback in a 15-min training session.

Self-Appraisals of Questions: Value for Self

I used an item adapted from Scardamalia and Bereiter (1992) to measure participants' appraisals of the extent to which a question had the potential to lead to an advance in understanding. Participants completed the 1-item scale for each question they generated. The item, which referred to a specific question that the participant had recorded, read as follows: *The answer would be useful for my learning*. Appendix D illustrates how the item appeared in the question log.

Scoring

Participants rated *Value for Self* according to the following response options: *Very little* (1); *A small amount* (2) *Somewhat* (3); *Quite a bit* (4); and *A great deal* (5). Thus, questions with higher *Value for Self* scores indicated a participant's perception of a question's higher importance for his or her learning. Raw scores across questions and across logs were used to create mean scores. Mean scores could potentially range from 1 to 5.

Validity and Reliability

Value for Self refers to an evaluation of how much understanding an individual might gain from the answer to one's question (Scardamalia & Bereiter, 1992), an idea that is consistent with Flammer's (1981) notion that individuals may evaluate their questions in light of their learning goals. The item has face validity in that it asked the participant to rate the gain in understanding for a specific question. In the training session (described below) students had the opportunity to ask about the meaning of the item. Students' responses after practicing with the log during the training session suggested that they were not confused by the item. Since *Value for Self* was measured with a single item, the scale cannot be tested for internal consistency.

Self-Appraisals of Questions: Value for Others

I developed a 1-item scale to assess a participant's appraisals of the value of their question for others (*Value for Others*). The item was designed to be a corollary to *Value for Self* item, which was adapted from Scardamalia & Bereiter (1992). Participants completed the 1-item scale for each question they generated. The item, which referred to a specific question that the participant had recorded, stated: *The answer would be useful for others' learning*. Appendix D illustrates how the item appeared in the question log.

Scoring

Participants rated *Value for Others* according to the following response options: *Very little* (1); *A small amount* (2) *Somewhat* (3); *Quite a bit* (4); and *A great deal* (5). Thus, questions with higher scores indicated a perception of a question having more value for others. Raw scores across questions and across logs were used to create mean scores. Mean scores could potentially range from 1 to 5.

Validity and Reliability

Value for Others referred to an evaluation of how much others' might benefit from the answer to one's question. The item did not specify the nature of the value and students might have interpreted the item in different ways. For example, a participant could assign value for others for different reasons, such as the type of value (e.g., interesting or useful to others), the degree to which a single classmate could benefit (e.g., a little or a lot), the number of classmates who would benefit (e.g., many or a few), or the value for serving different goals (e.g., others' academic or social goals). Possible interpretations are discussed further in Chapter 5. The single-item scale cannot differentiate these issues. Nevertheless, the *Value for Others* scale has face validity in that it directly addresses about how much others would benefit from learning the answer to one's question. Since *Value for Others* was measured with a single item, the scale cannot be tested for internal consistency.

Achievement in General Chemistry I

Measures of achievement before and after the target unit consisted of students' scores on Exam 1 and Exam 2 in the course. The course instructor released participants' raw exam scores (before adjustments such as a curve were applied). Students who were enrolled in General Chemistry I took a total of three exams in addition to a comprehensive final. Each exam covered 2 to 3 chapters in the course textbook and spanned 8-10 50-minute lectures. Exams were group administered and lasted approximately 50 min. Exam 2—the dependent variable for achievement in this study—primarily measured students' ability to efficiently solving problems that involved applying conceptual and procedural knowledge, including making mathematical calculations. The exam consisted of 25 items, only 2 of which solely assessed conceptual knowledge. The remaining 23 items required students to solve problems that involved applying procedural and conceptual knowledge.

Normative Change in Motivation Constructs

I measured normative change in motivation constructs from Wave 1 to Wave 2 with dichotomous categorical variables based on students' change on a given motivational construct relative to the sample as a whole. A normative change score was calculated for each motivational construct that was measured at both time points (personal interest, self-efficacy, individual goal orientations, teacher emphasis on mastery goals, and intention to remain in STEM).

I computed normative change in motivation scores in two steps. In the first step, I created a residualized gain scores for each variable measured at Wave 1 and Wave 2 by regressing Wave 2 scores on Wave 1 scores and saving the standardized residuals. For residualized gain scores, a positive value indicated that a student had increased their score on a construct from Wave 1 to Wave 2 beyond that which would be expected for the group as a whole (i.e., fit line in the regression). In the second step, I dichotomized the residualized gains scores according to whether students had positive residuals (increased over time relative to normative changes in the sample) or negative residuals (decreased over time relative to normative changes in the sample) on the construct. The *increased* or *decreased* label referred to an individual's change from Wave 1 to Wave 2 *relative to the change of the sample as a whole.* Note that a student could be in the *Increased* category even if he or she decreased slightly from Wave 1 to Wave 2 so long as his or her decline was less than the sample as a whole.

Procedure

The data were collected in three phases. Table 3.4 presents the goals, data sources, duration, timing, and location for each phase. The first phase began near the start of the spring 2013 semester (Feb 6, 2013) and ended after the first General Chemistry I exam (February 20, 2013). During the first phase, I recruited participants, conducted field observations, obtained informed consent, trained students on how to complete question logs, and collected questionnaire data (Wave 1). The second phase spanned 8 class meetings, starting after Exam 1 (February 22, 2013) and ending after the last period before Exam 2 (March 18, 2013). In the second phase, I conducted field observations of class meetings and students recorded their questions using question logs during each class meeting. The third and final phase began the first class meeting after Exam 2 (March 18, 2013) and ended on March 29. During the third phase, I continued to conduct field observations of class meetings and students completed a short follow-up questionnaire (Wave 2).

Recruitment

Students were recruited into the study via an in-class announcement in the General Chemistry I lecture in the first week of class. In the announcement, I explained: (1) the purpose of the study was to examine the questions students generate during chemistry lectures, (2) participation in the study required approximately 1 hour of out-of-class time in which students would complete questionnaires, (3) participation required students' active participation during 8 class meetings by recording their questions; (4) students who elected to participate in the study would receive extra credit towards their General Chemistry I lecture course grade in exchange for their participation; (5) participation was completely voluntary; (6) their decision to participate or not in the study would in no way adversely affect their course grade; (7) their response to questionnaires and the questions they generated will be kept confidential; and (8) students could elect to complete an alternative assignment for the same amount of extra credit if they did not wish to participate in the study.

For students who participated in the study, extra credit points—equating to approximately 1% of their final course grade—were added to their General Chemistry I course grade by instructor. I designed an alternative assignment requiring an equivalent amount of effort and time but no students expressed interest.

Questionnaire Apparatus

I took several steps to ensure that students who participated in the study remained anonymous and that questionnaire data were kept confidential. Participating students completed questionnaires using Temple University's Blackboard™ portal. Blackboard™ is a web-based course management software that can be used to securely administer questionnaires. I created a Blackboard course named "General Chemistry I Extra Credit Study" for the exclusive purpose of this study and designated myself as "instructor" with full administrative rights. Students who wished to participate in the study were enrolled in the course with the designation of "student." All communication and interactive functions of Blackboard™ were disabled. Only users with the "instructor" designation were able to view participants in the study and access questionnaire responses. Users with a "student" designation were unable to view other participants or access questionnaire responses. Only students who wished to participate in the study were enrolled in the course to prevent non-participating students from accessing questionnaires.

Table 3.4

Summary of Phases of the Study

	Phase 1		Phase 2		Phase 3	
Method	Survey	Observation	Question logs	Observation	Survey	Observation
Goal	• Obtain informed consent • Assess selected variables • Train Ss on question log	• Record baseline of SGQs before intervention • Describe instructional style and norms regarding Q asking	• Record SGQs, and post-class perceptions	• Record Qs asked during class meetings • Record fidelity for question logs • Describe instructional style during the use of Qlogs	• Assess selected variables	• Record Qs asked class meetings after intervention • Describe instructional style after the use of Qlogs

Note. Q = Question; SGQ = Student-generated question; Qlog = Question log.

(Continued Table 3.4)

	Phase 1		Phase 2		Phase 3	
Method	Survey	Observation	Question logs	Method	Survey	Observation
Data source	• Wave 1 Questionnaire	• Field notes	• Question log	• Field notes • Fidelity evaluation	• Wave 2 Question-naire	• Field notes
Duration	~ 40 min	2 weeks (6 class meetings)	4 weeks*	4 weeks* (8 class mtgs)	15-20 min	2 weeks (6 class mtgs)
Dates & relation to target unit	2/6-2/15 Recruit on Feb. 6	2/6-2/18 2 weeks prior to the target unit	2/22 – 3/20* Entirety of target unit	2/22 – 3/20* Entirety of target unit	3/22-3/29 2 weeks after target unit	3/22 – 3/29 2 weeks after target unit
Where	Computer lab	Beury 160	Beury 160	Beury 160	Via e-mail	Beury 160

Note. Mtgs = Meetings; *Although the intervention spanned 4 calendar weeks, the intervention consisted of 2.5 instructional weeks due to a 1 week spring break recess.

Following the class meeting in which I explained the study, students wishing to participate in the study were asked to sign-up for a 1-hour time slot in which to provide informed consent, complete questionnaires, and receive training on how to use the question logs. The signup took place in the public lobby outside of the lecture hall (Beury 160). Possible time slots were posted on large poster boards hung on the walls of the lobby. Time slots covered 10 weekdays and include a range of times of day. Each time slot allowed for a maximum of 10 participants. Students were asked to print their full name and official Temple University e-mail in an available time slot. After signing up for the time of their choice, students received a slip of paper with directions to the computer lab where the questionnaire was administered as well as a place to record the date and time of their appointment. One day prior to their appointment, students received an individual e-mail reminding them of the time of their appointment and directions to the computer lab. Signs were posted in and around the computer lab signaling directions for study participants. The informed consent and initial questionnaire were completed in the College of Education computer lab.

Wave 1 Questionnaire

When students arrive for their appointment, they were asked to read and sign an informed consent form for the study. After consenting, students were directed to log into Blackboard™ and clicked on the link to the “General Chemistry I Extra Credit Study.” Before accessing the questionnaires, students were asked to enter a password that was provided to them, thus ensuring students not participating in the study did not have inadvertent access to the questionnaires. Participants were provided a unique identification number that protected the confidentiality of participants' responses in data downloaded from Blackboard™.

Wave 2 Questionnaire

For the second wave of questionnaires, I sent an individual e-mail to each participant requesting that they access their personal Blackboard™ site and complete a set of questionnaires. The Wave 2 questionnaire had the same security precautions described above and were accessible for two weeks after the e-mail was sent. The Wave 2 questionnaire was expected to take approximately 15-20 minutes to complete and students completed it at a time and location that was convenient for them. Table 3.4 describes the questionnaires and estimated times for Wave 2 data collection.

Question Collection Apparatus: Question Log

Students participating in the study were asked to record their questions during lectures for each of the eight 50-min lectures between the first and second exam. Students recorded their questions in a question log (see Appendix D). During the session in the computer lab students were trained on how to use the question log. The training consisted of watching a 3-min informational video explaining the parts of the question log and how students should use it during class. Instructions stated:

> It is common to have questions occur to you while listening to a lecture or watching a presentation. In fact, generating questions can help you learn course material. One of the goals of this study is to examine the questions that students have during chemistry lectures.
>
> During the lectures between Exam 1 and Exam 2, you will be asked to record your questions in a question log. Please record all of the questions about

the material that occur to you. Your questions may seek to clarify your understanding, ask for a specific piece of information or wonder about something related to the topic. It is important that you try to write down all of the questions that occur to you. The only questions you do not need to record are those that you immediately figure out the answer to or those that are completely unrelated to the topic being presented.

After you have written down a question, please note the time. There will be clock in the lecture hall. Then, please rate your question on two dimensions. First, rate the extent to which you think the answer would advance your knowledge on the topic. Next to your question, you will see the label "Useful to my learning" and a scale ranging from *Very little (1)* to *A great deal (5)*. Please rate your question accordingly. For example, you may think of questions whose answers would really make a big difference in your understanding. In this case, you would circle *a great deal*. Other questions might seem minor or trivial. For these questions you would circle *very little*.

The second dimension is the extent to which you think your question would benefit your classmates. Responses range from *Very little (1)* to *A great deal (5)*. Please rate your question accordingly. For example, you may think of a question that you think your classmates would also greatly benefit hearing the answer to. In this case, you would circle *a great deal*. You might think other questions that you think would be useful only to you. For these questions you would circle *very little*.

> To summarize, for each question you generate, please complete the three tasks described above: 1) write the question down; 2) record the time; 3) rate the question on two dimensions. Once you do this a couple of times, I think that you will find that you can do it quickly. There is additional space to write your questions on the back of the sheet.
>
> Finally, if you asked a question aloud during a class meeting, please make sure that you have written it down and rated it. Please circle question that you ask in class.

In addition to the instructions above, each student was instructed to check a box on the question log to indicate those questions that were resolved or answered between the time the question was recorded and the end of the class meeting.

Students accessed the training video via the secure Blackboard™ site created for the study. Students practiced using the log with a 5-min video on a chemistry topic (limiting reactants). The video automatically paused for 1 min at two points during the video (after approximately 2.5 min and at the end of the video). When the video was paused, students were instructed to write down any questions they had or could think of about the content of the video. After completing the training exercise, students had the opportunity to ask questions about using the log and were given feedback on their use of the log during the training exercise (e.g., reminders to complete sections left blank).

Approximately 30 min prior to each class meeting, I placed a question log on each seat in the lecture hall. Students not participating in the study were asked to ignore the question log on their seats. On the first and second class meetings in which the logs were used, I gave a brief announcement reminding students how to complete the logs. At the end of each class meeting, two members of the research team stood at the exits and collected question logs from students as they exited the hall.

Question Log Use During Class Meetings

Table 3.5 illustrates the amount of time and the number of breaks allotted to question logs during the 8 class meetings. Each class meeting, the instructor paused 2-3 times for a total of approximately 3 min per class meeting for students to record their questions. These breaks in instruction usually occurred during natural transitions in the lectures, such as when the instructor paused to erase the white board or moved from PowerPoint slide to the white board or vice versa. At these times, the instructor asked students "to take a minute or two to write down your questions." During the time allotted to question logs, the lecture hall was relatively quiet, with most students recording their questions or reviewing their study materials. At the end of the allotted time, the instructor continued with his lecture.

Following each class meeting, I entered students' questions into a spreadsheet, removing personally identifiable information. I organized the list of questions by topic and sorted the list by whether questions had been resolved or not. Lists of students' questions were sent via email to the instructor on the same day as the class meeting in which the questions were generated. The instructor disseminated each list to the recitation instructors. Thus, the question logs constituted an authentic channel of communication between students and their instructors.

The instructor frequently endorsed the value of the question logs for him as an instructor. He began many class meeting by thanking students for their questions. On several class meetings, he reviewed topics that appeared repeatedly across students' questions from the previous period. The instructor reminded students that their questions were also sent to the recitation instructors, so that they could be addressed as the instructors deemed appropriate.

Question Asking Before and After the Use of Question Logs

I observed and wrote field notes before, during, and after the intervention. Observations began 6 class meetings prior to the use of the question logs, during all 8 class meetings during which the question logs were used, and continued for 6 class meetings after the use of the logs. According to my daily observations, teacher–student interactions before and after the use of question logs did not appear to differ dramatically compared with interactions during the 8 class meetings in which students recorded their questions. As noted above, the instructor allotted approximately 3 min each class for students to log their question and often made remarks at the beginning of class meetings that validated students' questions and provided some answers. Beyond these actions, however, class meetings before, during, and after the intervention were very similar. For example, students *asked* very few questions throughout the observations, including in the time before, during, and after the use of the logs (11, 7, and 5 questions, respectively). In general, the question logs appeared to make minor changes to the use of class time during the intervention but did not make a noticeable impact on teacher and student behaviors during class meetings.

Table 3.5

Allotted Time for Question Log Use During Class Meetings

Day	No. Qlog breaks	Time Allotted for Qlogs (sec)	*M* time per break (sec)
1	3	215	71.7
2	2	221	110.5
3	2	178	89.0
4	3	192	64.0
5	2	213	106.5
6	2	149	74.5
7	2	121	60.5
8	3	168	56.0
M (SD)	**2.4 (.52)**	**182 (35.24)**	**79.1 (20.76)**

Note. Qlog = Question log.

Coding Question Characteristics

The process of coding questions for relevant characteristics was conducted in two phases. First, I established coding categories. Second, I trained a second coder, established good inter-rater reliability, resolved differences between the coders, and assigned final codes. I describe each stage below.

Establishing Coding Categories

In the first phase, my goal was to identify codes that were suitable for the questions students generated in this study. I accomplished this through iterative cycles of coding that used both deductive and inductive methods. Working deductively, I used the GPH taxonomy (Graesser et al., 1992) as a starting point, applying GPH question types when they adequately captured questions in the dataset. Working inductively, I generated possible new question types when GPH types did not adequately capture aspects of students' questions. I applied new and existing GPH coding categories to small subsets of questions and edited accordingly (Miles & Huberman, 1994). This iterative process yielded two primary changes to the GPH scheme. First, I modified the scheme so that question syntax became an orthogonal dimension to question type. Second, I identified new question types that were not adequately captured in the original GPH scheme. I describe each change below.

Question Syntax

Each question was coded for the semantic syntax of the question. Questions were assigned one of three syntax codes: *Closed, Open,* or *Assertion.* Questions with closed syntax (*Closed Questions*) were questions whose syntax constrained the solicited response to the extent that the question could be answered with a single word. For example, questions with closed syntax include questions that sought verification (i.e., Is x the case?) or presented a disjunction (i.e., Is *x* or *y* the case?). By contrast, questions that used open syntax (*Open Questions*) were phrased such that the question allowed for an array of responses. For example, questions with open syntax often started with interrogative words, such as *who, what, when, where,* or *why*. Assertions, the third syntax code, referred to statements that were not syntactically constructed as questions but which articulated a knowledge gap. Assertions included statements expressing surprise (e.g., But I thought *x* was the case?!) or describing a lack of understanding (e.g., I don't understand *x*).

Coding for question syntax constituted a modification of the GPH taxonomy. Three question types in the original GPH taxonomy were characterized by the syntax of the question: *verification questions, disjunctive questions*, and *assertions*. As noted previously, the GPH taxonomy can be applied as a monothetic scheme (a question is a assigned a single classification) or a polythetic scheme (questions can be assigned multiple classifications). When used polythetically, the GPH scheme would allow a question like *Does smoking cause cancer?* to be coded according to its syntax (e.g., verification) as well as other characteristics (e.g., causal antecedent). However, the original GPH taxonomy did not necessarily characterize questions by their syntax in cases where the syntax was neither disjunctive, verification, nor assertion, as with the question *What causes cancer?* For purposes of parsimony, I combined verification and disjunctive questions to constitute a single *closed* syntax code. In addition, I created a code for questions using *open* syntax, such that all questions could be coded according to their syntax. Therefore, in the current study, the modified GPH taxonomy is used as a polythetic scheme, in which each question is assigned exactly two codes: question syntax and question type. In cases where a question could be coded as two or more question types, a single question type was selected (see the description below for tie-breaking rules; descriptive results for question syntax are presented in Chapter 4).

Modified Question Types

GPH question types were modified in two ways: new question types were identified and original GPH categories were collapsed. Using the inductive methods described above, several new question types were identified, which were not part of the GPH taxonomy and which represent conceptually distinct types of questions in the dataset. New question types included question that addressed both procedural and conceptual knowledge. Question types, definitions, and examples are presented in Table 3.6 (Descriptive results for question types are presented in Chapter 4).

Additional types of procedural questions. Many of the questions students generated targeted information about executing procedures, reflecting the content of the lectures. Much of the instructional time focused on how to solve chemistry word problems, including how to identify relevant formulae, convert units, solve for specific values, and balance and reduce equations. Accordingly, students generated many types of questions about how and when to execute such procedures.

The diversity of procedural questions was not adequately captured by the single *procedural/instrumentality* code in the GPH scheme. I identified five distinct question types that addressed procedural knowledge: *Procedural General*, *Procedural Rationale*, *Procedural Conditional*, *Procedural Strategy Evaluation*, and *Procedural Enablement*. *Procedural General* referred to questions that sought general information about how to execute a procedure, including executing a set of procedures (e.g., *How do you calculate the depth of the snorkeler in problem 5.2?*) or steps in a process (e.g., *How do I remove spectator ions?*). *Procedural Rationale* referred to questions that sought a justification or rationale for a procedure. Examples of *Procedural Rationale* questions included *Why do we convert mmHg to atm?* and *Why do you have to balance the molecular equation before you balance* [the] *full ionic* [question]*?* *Procedural Conditional* referred to questions that sought information about the conditions when procedures were relevant, appropriate, or desirable. Examples of *Procedural Conditional* questions included *When do we have to use the solubility charts to solve the ionic equations?* and *In what case would you use D = PM/RT?* *Procedural Strategy Evaluation* referred to questions that sought a judgment regarding the appropriateness, desirability, or efficiency of a strategy to solve a problem. Examples of *Procedural Strategy Evaluation* included *Can we write the full ionic equation completely before figuring out* [the] *net ionic* [equation]*?* and *Can you complete the stochiometric process in any order*? Lastly, *Procedural Enablement* referred to questions that sought information about how one knows when to execute a procedure. Example of *Procedural Enablement* questions included *How do you know that* Cl^- *and* Na^+ *are spectator ions?* and *How can you tell if the equation is directly or inversely proportional?*

Some types of procedural question overlapped with existing GPH question categories. *Procedural Rationale* questions could be considered an instantiation of the GPH *Goal Orientation* category, which refers to the goals or motives behind an agent's actions. Likewise, *Procedural Enablement* could be considered an instantiation of the

GPH *Enablement* category. The new procedural question categories, however, facilitate analyses that distinguish between conceptual and procedural knowledge.

Additional types of conceptual questions. I added three conceptual question types that were not represented in the original GPH scheme: *Concept Extension*, *Concept Justification*, and *General*. *Concept Extension* referred to questions that sought information regarding the limits of a relationship or classification. *Concept Extension* questions differed from *Concept Completion* questions in that the former sought rules that defined limits or boundaries for concepts or relations (e.g., Is x always the case?), rather than simply completing or elaborating a concept (e.g., Who? What? When? Where?). Examples of *Concept Extension* questions included *In an acid-base type reaction is* CO_2 *and water* ***always*** *produced* ***every*** *time?; Is the precipitate* ***always*** *a solid?;* or *Will gas "A"* ***always*** *be a lighter gas?*

Concept Justification referred to questions that sought a justification for how it was possible that a concept or relation existed as it did. Examples of *Concept Justification* questions included *How can we not have a limiting reactant? Why will volume always be in liters? Why is* NO_3 *a spectator ion if it goes from* NO_3^- => $2NO_3^-$ *to balance [the] equation?* and *How can gas be a solute?*

General questions referred to questions that were vague or unspecified inquiries or expressed general confusion. Examples of *General* questions included *What is going on?; Net Ionic equations*? and *If only I knew what to ask.* Note that unless *General* questions explicitly addressed procedural knowledge, they were not coded as *Procedural General*, even though it is possible that the general confusion expressed by such statement may have referred to aspects of procedural knowledge.

In additional to the changes to the GPH scheme described above, I made the following modification. The GPH categories *causal antecedents*, *causal consequences*, and *expectational*—which all address causal relation—were treated as a single *Causal Relations* category for reasons of parsimony and because differentiating between causal antecedents and consequences can be difficult or impossible to infer in some cases (Graesser, McMahen, & Johnson, 1994).

Finally, the original and modified question types were grouped into two broad categories: conceptual questions and procedural questions. It is important to emphasize that *conceptual* questions addressed knowledge related to non-procedural concepts, including factual information (a point I discuss in further detail in Chapter 5).

Breaking Ties. Some questions fit the description of more than one question type. On such occurrences, the coders used information about whether the question was generated during instruction that addressed conceptual or procedural topics to inform their judgment. If this information did not suggest a clear decision, the coder assigned the question type associated with higher- rather than lower-order thinking, as described by Graesser and colleagues (Graesser et al., 2009).

Table 3.6

Description of Modified GPH Codes for Question Types with Examples Using Closed and Open Syntax

Question Type	Definition	Examples with Open and Closed Syntax
Conceptual Questions		
General/ Request*	Indicates vague, unspecified or general confusion or phrases a question in an unintelligible way.	(O) What is going on? If only I knew what to ask (C) n/a
Definition	Seeks information about the meaning of a word or phrase: What does *x* (term or phrase) mean?	(O) Do gases get changed during the reaction? (C) Is the molarity a ratio of solute to solvent or solute to solution?
Concept completion	Seeks factual information that clarifies or elaborates a concept or concepts. Who? What? When? Where?	(O) How was the percentage yield formula created? (C) Is the theoretical yield accurate?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Feature specification	Seeks information about the features or qualities of an entity. What qualitative characteristics does *x* have?	(O) How [do] we get carbonic acid? From where is the second H^+? (C) Is that P_c after P_a and P_b?
Relational/ Comparative	How are X and Y similar and different?	(O) What is the difference between a limiting reactant and a limiting reagent? (C) Is there a difference between a reagent and a reactant?
Concept extension*	Seeks information regarding the limits of a relationship or classification • Is X always the case?	(O) What does it mean if the law in use is not followed (i.e., Charles Law)? (C) Does the variable R always = 0.08206?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Justification of concepts or relations*	How can it be that *x* is the case? Why is it that *x* has a specified relation to *y*? Why are *x* and *y* different or similar?	(O) Why would salt be a (aq) and not (s)? (C) Is D = P(MM)/RT equals (Molar Mass/Molar volume) because the both equations are densities of gases?
Enablement (Non-procedural)	What resource, object, or information allows an agent to accomplish a goal? What knowledge (i.e., resource) allows an agent to know something (i.e., accomplish a goal)	(O) How do you know a substance is polar or nonpolar? (O) When do you know when a solution is (aq) compared to (l), (s), (g)?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code. † Analogous to GPH *Judgmental* category.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Causal relations	(Causal Antecedents): What led to an event or state? Why did an event occur? Why does a state exist? How did an event occur? How did a state come to exist? (Causal Consequences) What are the results or consequences of a state or event? What if X occurred? What if X did not occur? (Causal Expectations) Why did some expected event not occur? Why does some expected state not exist?	(O) Why do the spectator ions lose their numbers or amounts? (C) Do gases get changed during the reaction?
Purpose/Value†	What is the value of knowing X? What is the purpose of X? How is X used?	(O) What is a mole fraction used for? (C) The purpose of partial pressure is to find the pressure to plug into PV = nRT?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code. † Analogous to GPH *Judgmental* category.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Procedural Questions		
General Procedural*	Seeks information about how to accomplish a goal, such as: What plan or instrument allows an agent to accomplish a goal? How did agent do some action? How did agent do some action?	(O) How do you balance equations? (C) Do you convert the HNO_3 to NA_2CO_3?
Procedural Rationale*	Seeks information regarding the rationale behind Why did the agent execute (or not execute) *x* procedural steps? • Why is the agent not using *x* procedure?	(O) Why do you have to balance the molecular equation before you balance full ionic? (C) Is it really necessary to rearrange PV = nRT to fit what you're looking for?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Procedural Strategy evaluation*	Seeks an evaluation for the appropriateness or value of a procedure. • Is appropriate for X procedure to be executed? • Does the order of X and Y procedures matter? • Is Y an appropriate alternative to X?	(O) What is the best way to be sure if a compound is unstable? (C) Can I set Boyle's law up as a proportion?
Procedural Conditional*	Seeks information about the conditions under which X procedural acceptable or valuable?, such as What procedure would be called for in X conditions? Or What procedure would be called for in X conditions? Or What features below to a class of problems?	(O) When will we use the combined formula $(P_1V_1/T_1) = (P_2V_2/T_2)$? (C) Is that entire process included in every reaction?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code.

(Continued Table 3.6)

Question Type	Definition	Examples with Open and Closed Syntax
Enablement Procedural*	Seeks understanding of how • What object or resource allows an agent to accomplish a goal? What knowledge (i.e., resource) allows an agent to know something (i.e., accomplish a goal)	(O) How do you decide which ions are spectator ions? (O) How do you know that H_2SO_4 => $2H^+ + SO_4^{2-}$?

Note. (O) = Open syntax; (C) = Closed syntax; Codes based on question types and descriptions in Graesser et al., (1992); *New question code.

Inter-Rater Reliability for Question Coding

Once the question types were identified and defined, I trained a second coder using a randomly selected portion of the questions that was not be used to establish inter-rater reliability. After the two coders achieved strong inter-rater reliability on the practice portion of the questions, the two coders independently coded a set of questions that constituted approximately 35% of the total number of student-generated questions. Inter-rater reliability statistics are presented in Table 3.7. I calculated a Cohen's *kappa* for 1) Syntax codes; 2) All question types; 3) Procedural vs. conceptual question. Landis and Koch (1977) suggested the following benchmarks for interpreting *kappa* for observer agreement with categorical data: 0 = poor, .01–.20 = slight, .21–.40 = fair, .41–.60 = moderate, .61–.80 = substantial, and .81–1 = almost perfect. By these criteria, inter-rater reliability was excellent for question syntax (K = .90) and was good for question types at both the course-grain level (procedural vs. conceptual; K = .69) and the fine-grain level (all 15 question types; K = .65). Note that procedural and conceptual categories were not codes that were actually applied in the coding process. Questions were coded for question type and question types were then grouped to form broader conceptual and procedural categories.

Table 3.7

Inter-Rater Reliability for Syntax and Question Type Codes

Code	No. Coding Categories	K
Question syntax	3	.90
Question types		
All question types	15	.65
Conceptual vs. Procedural	2	.69

Note. K = Kappa.

Factor Analyses

I conducted Exploratory Factor Analyses (EFAs) on: (1) the four individual goal orientations and teacher emphasis on mastery goals scales at Wave 1; (2) the five help seeking scales at Wave 1; (3) personal interest in chemistry at Wave 1; (4) self-efficacy at Wave 1; and (5) intention to remain in STEM at Wave 1. I conducted factor analyses in order to determine the extent to which there was empirical support in this sample for theoretical distinctions between related constructs. I used the results from the factor analyses to make decisions about aggregating items from scales. I conducted EFAs rather than Confirmatory Factor Analyses because with 103 participants I lacked the statistical power to conduct the latter. All EFAs consisted of Principal Axis Factoring using an Oblimin rotation since variables were assumed to be correlated with each other (Keith, 2006). I conducted EFAs in iterative steps in which I considered the percent of variance explained by each factor, the factor loadings on the pattern matrix, an analysis of the scree plot, and theoretical considerations. Analyses were conducted as described below.

EFA on Wave 1 Goal Orientations Scales

I conducted Principal Axis Factoring on the four individual goal orientation scales—mastery approach, mastery avoidance, performance approach, and performance avoidance—and the teacher emphasis on mastery goals scale. In the initial step, the number of factors selected was based on eigenvalue greater than one. The analysis yielded a five-factor solution, explaining 57.4% of the variance. Three pieces of data indicated that the five-factor solution was not the most parsimonious. First, one of two items that loaded on factor five also loaded on other factors. Second, factor five explained a small amount of additional variance (3.9%). Third, the scree plot indicated only a marginal decrease in eigenvalues from a four- to a five-factor solution. In a second step, I conducted Principal Axis Factoring in which a four-factor solution was specified. The four-factor solution explained 51.8% of the total variance.

The factor loadings of the pattern matrix are presented in Table 3.8 with factor loadings less that .30 not shown to highlight the factor structure. The factor structure indicated that mastery approach, mastery avoidance, and teacher emphasis on mastery goals loaded on separate factors (Factor 4, 3, and 2, respectively). Performance approach and performance avoidance items loaded on the same factor, consistent with many empirical findings (Murayama, Elliot, & Yamagata, 2011; discussed in further detail in Chapter 5). Note that in the teacher emphasis on mastery goals scale, one item (*My chemistry lecture professor wants us to understand our work, not just memorize it*) had a loading below .30 on the factor associated with this scale. In the context of General

Chemistry I, students may have perceived that the instructor valued memorization (e.g., formulas, rules for solubility, etc.) as an important part of succeeding in the course; therefore, they may have responded to this items independent of how they perceived other aspects of the goal structure of the course. Despite the low factor loading, I decided to retain this item in the scale since it tapped an important aspect of the construct measured by the scale. In light of EFA results, mean scores were calculated for mastery approach goal orientation, mastery avoidance goal orientation, performance goal orientation, and teacher emphasis on mastery goals.

EFA on Help Seeking Scales

I conducted Principal Axis Factoring on the five help seeking scales: instrumental help seeking, executive help seeking, help seeking avoidance, help seeking threat, and preference for a formal help seeking target. In the initial step, the number of factors selected was based on eigenvalues greater than one. The analysis yielded a five-factor solution, explaining 51.8% of the variance. Three pieces of data indicated that the five-factor solution was not the most parsimonious. First, the three items that loaded on factor five also loaded on other factors. Second, factor five explained a small amount of additional variance (2.8%). Third, the scree plot suggested only a marginal decrease in eigenvalues from a four- to a five-factor solution. In a second step, I conducted Principal Axis Factoring in which a four-factor solution was specified. The four-factor solution explained 49.1% of the total variance. One item loaded on more than one factor. The analyses were repeated with this item removed.

Table 3.8

Pattern Matrix for EFA of Goal Orientation Variables

	Factor			
	1	2	3	4
MAP Item 1				-.761
MAP Item 2				-.628
MAP Item 3				-.551
MAV Item 1			.662	
MAV Item 2			.504	
MAV Item 3			.829	
PAP Item 1	.836			
PAP Item 2	.699			
PAP Item 3	.844			
PAV Item 1	.482			
PAV Item 2	.694			
PAV Item 3	.771			
TEMG Item 1				
TEMG Item 2		.589		
TEMG Item 3		.721		
TEMG Item 4		.516		
TEMG Item 5		.699		

Note. Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. MAP = Mastery approach goal; MAV = Mastery avoid goal orientation; PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals.

The factor loadings of the pattern matrix are presented in Table 3.9 with factor loadings less that .30 not shown to highlight the factor structure. Items from the Formal help seeking scale, help seeking threat scale, and the executive help seeking scale had strong loadings on Factors 2, 3, and 4, respectively. Items from the help seeking avoidance scale and instrumental help seeking scale loaded on the same factor. Avoidance items loaded negatively and instrumental items loaded positively on the factor, consistent with theoretical conceptualizations of these constructs (e.g., instrumental help seeking is characterized by not avoiding help when it is needed). I calculated a Cronbach's alpha for internal reliability for the help seeking avoidance and instrumental help seeking items, which loaded on Factor 1. Results indicated poor reliability when all items were grouped as a single scale ($\alpha = .56$); reliability improved when both avoidance items were dropped ($\alpha = .81$). Therefore, the two-item help seeking avoidance scale was dropped from future analyses. I calculated a mean score for instrumental help seeking items only as well as for executive help seeking, help seeking threat and preference for a formal help seeking target scales.

Table 3.9

Pattern Matrix for Exploratory Factor Analysis of Help Seeking Constructs

	Factor			
	1	2	3	4
HSAvoid Item 2	-.504			
HSAvoid Item 3	-.331			
HSExec Item 1				.609
HSExec Item 2				.584
HSExec Item 3				.368
HSInstr Item 1	.859			
HSInstr Item 2	.865			
HSInstr Item 3	.713			
HSInstr Item 4	.717			
HSInstr Item 5	.697			
HSThreat Item 1			.566	
HSThreat Item 2			.720	
HSThreat Item 3			.671	
HSFormal Item 1		.833		
HSFormal Item 2		.384		
HSFormal Item 3		.856		

Note. HSAvoid = Help seeking avoidance; HSExec = Executive help seeking; HSInstr = Instrumental help seeking; HSThreat = Help seeking threat; HSFormal = Preference for seeking help from a formal target.

Additional EFAs

I conducted three additional separate EFAs on Wave 1 items from the self-efficacy, personal interest, and intention to remain in STEM (StaySTEM) scales to evaluate whether these scales represented singular constructs. Personal interest items loaded on a single factor, explaining 61.4% of the variance. Self-efficacy items loaded on a single factor, explaining 54.2% of the variance. Accordingly, mean scores were calculated for Personal Interest and Self-efficacy items.

Principal Axis Factoring on StaySTEM generated communalities greater than one when attempting to extract two factors. The communalities table indicated the one item had substantially different loading than the others in the scale. Results from bivariate correlations and reliability analyses indicated that this item had lower correlations compared with other items in the scale. An examination of the wording of the item (*I am unlikely to switch to (or select) a non-science major before the end of this semester*) revealed that the item used possibly confusing language, such as a double negative. Therefore, the item was dropped. Principal Axis Factoring on the remaining five items indicated a single factor explaining 67.0% of the variance. Therefore, I calculated a mean score for the five remaining items.

Data Analysis

Descriptive Statistics

Means and standard deviations were calculated for all continuous and dichotomous variables. In addition, Pearson correlations were calculated among variables. Internal consistency was calculated for variables assessed with scales with 3 or more items using Cronbach's alpha (results are reported in Chapter 4).

Screening for Assumptions

I screen the data to evaluate the extent to which all variables met the required statistical assumptions of the analyses in which they were used. My analyses include univariate and multivariate analyses which require the following five assumptions to be met: 1) normality of distribution, 2) independence of observation, 3) homogeneity of variances, 4) no multicollinearity, and 5) linearity. I assessed each assumption for all variables. In addition, I check for the presence outliers, which can bias univariate and multivariate analyses.

Normality

For each variable, I calculated skewness and kurtosis statistics (presented in Chapter 4) and generated histograms and P-P plots for each variable, examining each for evidence of non-normal distribution.   In cases in which variables demonstrated evidence of non-normality (e.g., high kurtosis in a small number of variables), I took the extent to which the assumption was violated into consideration and made decisions about analyses accordingly (see discussion of normality in Chapter 4).

Independence of Observation

I assessed the extent to which self-appraisals were nested in clustering variables. Intra-class coefficients (ICCs) quantify the amount of variance in a variable that is attributable to a clustering variable (i.e., between group difference). ICC values less that .05 indicate that nestedness of variance within the clustering variable is minimal and multilevel modeling is unnecessary (Bickel, 2007). I calculated separate ICCs for *Value for Self* and *Value for Others* for two clustering variables: class meeting and question types (see Table 3.10). The small ICC values indicated that only a small amount of variance in self-appraisals was attributable to clustering variables. The low ICCs suggested that aggregating the self-appraisals across class meeting and across question type was appropriate.

Table 3.10

Intra-Class Coefficients for Self-Appraisals of Questions by Class Meeting and Question Type

	Value for Self	Value for Others
Class meeting	0.005	0.004
Question type	0.041	0.033

Homogeneity of Variances

For all pairs of predictor–outcome variables used in regressions and path analyses, I inspected a plot of the z-scores for residuals against the z-scores for the predicted regression scores. Plots did not suggest violations of homogeneity of variance. For the *t* tests described below, I examined the Levene's test for homogeneity of variance. None suggested a violation of the assumption of homogeneity of variance.

No Multicollinearity

I assessed multicollinearity for the multiple regression and path analyses by evaluating multicollinearity diagnostics. All analyses demonstrated acceptable levels of collinearity, as indicated by the Variance Inflation Factor (VIF), eigenvalues and variance proportions (Keith, 2006; Pedhazur, 1997).

I examined the multicollinearity for the regression of Exam 2 scores on *Value for Self* and *Value for Others* with particular care because the self-appraisals were relatively highly correlated (r = .68) and because the regression results revealed significant relationships that were not observed at the zero-order level. By evaluative criteria described by Pedhazur (1997), several pieces of evidence suggested that the significant regression results indicated the presence of a suppression effect rather than the result of entering multicollinear variables into the same regression. First, the direction of the estimated significant coefficients for each appraisal was consistent with the direction of the non-significant zero-order correlations (negative for *Value for Self* and positive for *Value for Others*). The VIF for each appraisal was sufficiently low even by conservative estimates (e.g., < 2.5). Moreover, multicollinearity leads to the overestimation of standard errors, which tends to result in non-significant rather than significant coefficients (Pedhazur).

I conducted two additional regression analyses to ensure that the pattern of significant results with regard to the self-appraisal was not simply the result of multicollinearity. First, I conducted a multiple regression of Exam 2 scores on self-appraisals, in which all variables were normalized. Multicollinearity statistics are influenced by the scales of variables; and therefore Pedhazur (1997) recommended normalizing variables—transforming variables so that each formed a 0-1 scale—when evaluating multicollinearity statistics. Multicollinearity statistics for the normalized and non-normalized variables were similar. Second, I regressed Exam 2 scores on the difference between the self-appraisals (*Value for Self* minus *Value for Others*). The regression results were similar to the regression with both appraisals entered. Taken together, these pieces of evidence suggested including both *Value for Self* and *Value for Others* in the same regression did not violate the assumption of non-multicollinearity.

Linearity

To test for linear relationships, I conducted a series of simple regressions between each endogenous variable and its predictors in the path models described below. In each simple regression, I compared the R^2 associated with linear and non-linear fit lines. Results suggested that the assumption of linearity was not violated.

Outliers

I screened for outliers in several ways. I examined the range of scores for evidence of univariate outliers. Two participants had exam scores of zero, which likely reflected students who did not attend class the day of the exam; the two exam scores were deleted. Regarding multivariate outliers, *Mplus* (version 7) calculates a set of diagnostic statistics to determine outliers: Cook's distance (Cook, 1977), the influence of each case on log likelihood, a log likelihood distance influence measure (Cook & Weisberg, 1982), and Mahalanobis distance (similar to leverage). Using these diagnostics, I identified potential outliers for the two path analyses.

For the model tested in Analysis A (described below), I compared results from analyses with and without potential outliers. The two sets of results were similar and therefore analyses with original scores are presented in Chapter 4. For the model tested in Analysis B (described below), I compared results from analyses with original and Winsorized scores (forcing all scores to fall within 2 *SD* of the mean). The results indicated minor differences. I inspected the specific cases and concluded that all scores represented plausible values (e.g., low endorsement of a construct, a low exam score) rather than errors. Therefore, I presented the analyses with the original scores and discuss the minor differences between the analyses with Winsorized and original scores.

Missing Data

Missing data on Wave 1 scales and question characteristics was minimal (< 3% missing). Missingness at Wave 2 was slightly higher (ranging from 8-12%). In the path models described below, I conducted analyses using Full Information Maximum Likelihood (FIML) estimator in *Mplus* (version 7.0). FIML maximizes statistical power by analyzing each unique pattern of missing data separately. Model parameters are then estimated using all known information, including from both complete and partially missing cases (Davey & Savla, 2010).

Alpha Levels and Corrections

I adopted an alpha level of $p < .05$ for all analyses. The large number of analyses and variables in the study, however, increases the possibility that significant findings occurred by chance (Type 1 error). Researchers sometimes account for multiple comparisons by dividing the alpha level by the number of analyses (e.g., Bonferroni correction) or related procedures (e.g., Hochberg's [1988] or Holm's [1979] sequential procedures). The standards for when alpha corrections are necessary (e.g., what constitutes a family of tests) is unclear (Miller, 1981; Rice, 1989). Furthermore, critiques of the appropriateness of alpha corrections in the natural sciences identify mathematical, logical, and practical concerns relevant to psychological and educational research (e.g., Cabin & Mitchell, 2000; Moran, 2003; Nakagawa, 2004). Bonferroni corrections present a relevant practical problem for the present study in that the more thorough and detailed the analysis of the data, the less likely significant findings are discovered (Moran, 2003). In the present study, I strived to balance Type I and Type II error. Given the relatively small sample size, I interpreted inferential statistics for all analyses as significant at $p < .05$. Where relevant, I also presented footnotes that indicate more conservative, Bonferroni-corrected alpha levels. I included Bonferroni-corrected alpha levels for instances in which more than one statistical test were used to test a null hypothesis (Rice, 1989). For example, the relations among question characteristics and Exam 2 scores were tested in four analyses—two path models, a multiple regression, and a *t* test on latent classes. Accordingly, a Bonferroni-corrected alpha level of $p = .05/4 = .0125$ is presented for more conservative interpretation of the results. Due to the different types of analyses (i.e., path models, multiple regression, *t* tests), I was unable to use less conservative alpha correction methods such as Hochberg's sequential procedures.

Data Analysis

Research Question 1

Research Question 1 (RQ1) asked: What are the relations among the quantity, qualities, and self-appraisals of student-generated questions? I answered RQ1 by conducting a descriptive analysis of question types, characteristics and their inter-relations and a latent class analysis (described in the Person-Centered Analyses section below).

Descriptive Analysis

Since little is known about the questions undergraduate STEM students generate during lectures, the aim of RQ1 was to describe the patterns and inter-relations of student-generated questions in terms of their quantity, qualities and self-appraisals. Questions are described in terms of the prevalence of question types and syntax.

Descriptive statistics and Pearson correlations were calculated for the number of procedural and conceptual questions, the percentage of questions that were unresolved at the end of class meetings, the percentage of closed questions, and students' appraisals of their questions (Results are presented in Chapter 4).

Research Question 2, 3a and 3b.

Research Question 2 (RQ2) asked: What are the relations among motivational self-beliefs and perceived teacher emphasis on mastery goals and the quantity, qualities, and self-appraisals of student-generated questions? Research Question 3a (RQ3a): Can exam scores be predicted by the quantity and qualities of questions? (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and Exam 2 scores?

I investigated RQ2 with several analyses, including testing two path models. The reason for testing path models was to simultaneously examine multiple relationships as well as to examine indirect relationships between variables. Due to the small sample size, I lacked the statistical power to test a single model that included all motivation variables, question characteristics, and Exam 2 scores. I examined bivariate correlations to determine which motivational constructs were significantly related to which question characteristics. Correlations indicated that Wave 1 interest and self-efficacy and post-class self-efficacy were related to the number of procedural and conceptual questions, the percentage of unanswered questions, and the percentage of closed questions while goal orientation variables tended to be significantly related with self-appraisals. Accordingly, primary analyses for RQ2, RQ3a, and RQ3b consisted of testing two path models (Analysis A and B). Analysis A (presented in Figure 3.1; described below) tested a model of relations among Wave 1 personal interest, Wave 1 self-efficacy, post-class self-efficacy, question characteristics (excluding self-appraisal), and exam scores. Analysis B (presented in Figure 3.2; described below) tested a model of relations among individual goal orientations and perceived teacher emphasis on mastery goals, self-appraisals of questions, and exam scores. Additional analyses to answer the above research questions included commonality analyses, a set of *t* tests comparing question characteristics for students who increased vs. decreased on motivational variables; calculating partial correlations; and a multiple regression in which Exam 2 scores were regressed on all question characteristics.

Analysis A

Analysis A (illustrated in Figure 3.1) tested a path model in which two exogenous variables—Wave 1 personal interest and Wave 1 self-efficacy—predicted five endogenous variables: the number of procedural question, the number of conceptual questions, the percentage of unanswered questions, the percentage of closed questions, and post-class self-efficacy. The model also specified that questions characteristics predicted post-class self-efficacy. In turn, question characteristics and post-class self-efficacy predicted Exam 2 scores.

The rationale for this model is that, as described in Chapter 2, self-efficacy and personal interest have been associated with strategy use and attention during learning and therefore may predict characteristics of students' questions while listening to lectures. As I described in the theoretical framework in Chapter 2, question characteristics may also influence individual perceptions, such as self-efficacy. Question characteristics may indicate aspects of students' knowledge and cognitive processes, which may have direct relations with exam performance. As I noted previously, I lacked statistical power to test a single model that included all motivation variables, question characteristics, and exam scores. I used bivariate correlations to identify these specific motivation variables and question characteristics.

Analytic strategy for evaluating model fit. I assessed model fit by evaluating a set of indicators including chi square value and significance test, Comparative Fit Index (CFI), Standardized Root Mean Square Residuals (SRMR), and Goodness of Fit Indicator (GFI). Root Mean Square Error of Approximation (RMSEA) is not recommended for sample sizes smaller than 250 (Hu & Bentler, 1999).

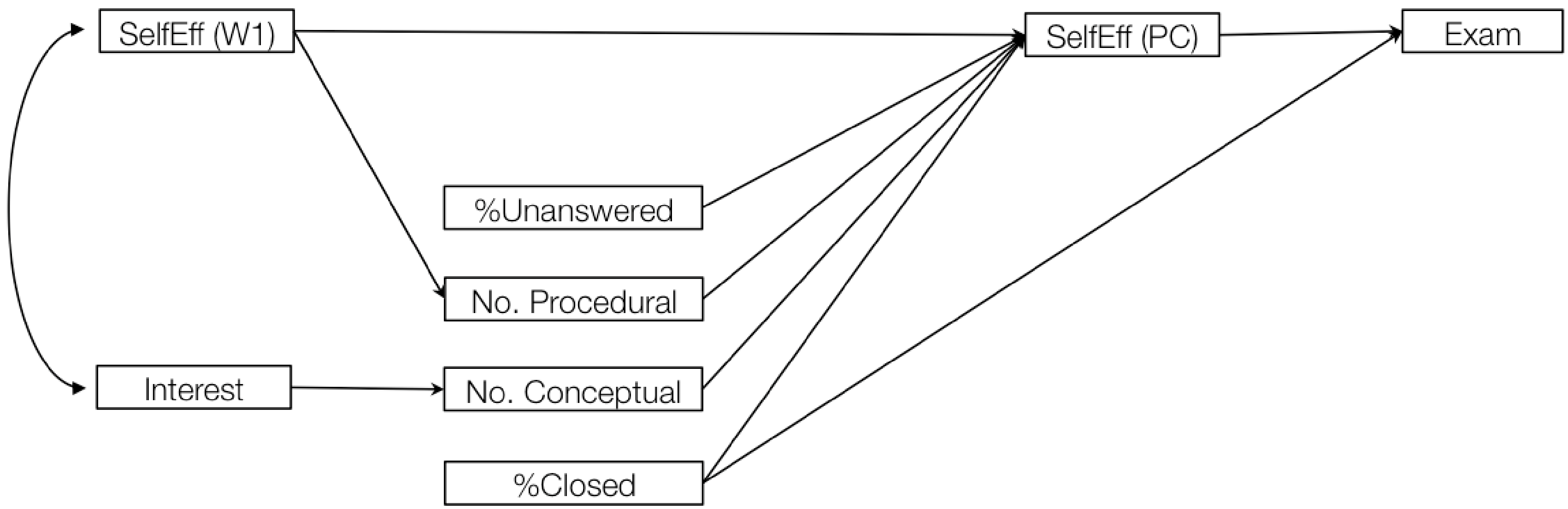

Figure 3.1. Path Model of Personal Interest, Self-Efficacy and Question Characteristics Predicting Exam Scores.

Note. Errors for endogenous variables not shown. W1 = Wave 1; PC = Post-Class; SelfEff = Self-Efficacy.

Model fit indices were interpreted according to evaluative criteria suggested by Kline (2010). A non-significant chi square indicates a non-significant difference between the variance and covariances predicted by the model and those in the data. Thus, a chi square that is non-significant indicates good fit. The chi square test of model fit was evaluated in light of other indicators. Hu and Bentler (1999) suggested that for small sample sizes (i.e., less than 250), researchers should privilege CFI and SRMR indicators to minimize both Type 1 and Type 2 errors. CFI compares the tested model to one in which all variables are uncorrelated. CFI values greater than .95 indicate good fit. SRMR assesses the difference between observed and estimated correlations. SRMR values greater than .09 may indicate a poorly fitting model (Hu & Bentler, 1999).

Good model fit indicates that parameter estimates may be interpreted with confidence. In the case of a good fitting model, I interpreted the model parameters to address the research question. The following parameters were of particular interest. First, the significance and relative strength of paths from exogenous variables (Wave 1 personal interest and self-efficacy) to endogenous variables (question characteristics) were important because these indicated the extent to which each motivational variable uniquely explained variance in features of students' questions. Second, the significance and relative strength of paths from question characteristics to post-class self-efficacy was important because these indicated the extent to which question characteristics explained variance in post-class self-efficacy perceptions after accounting for initial (Wave 1) self-efficacy. Third, the significance and relative strength of paths from question characteristics to exam scores was important because these indicated the extent to which question characteristics, along with post-class self-efficacy, explained variance in Exam 2 scores. Fourth, the amount of variance in endogenous variables explained by the model was important because these indicate the collective predictive power that motivation and/or question characteristics had on variables of interest. Fifth, the covariances among endogenous variables were important because they indicate the extent to which these variables were related. Finally, the indirect relations from motivation variables to exam scores via question characteristics and the indirect relations from question characteristics to exam scores via post-class self-efficacy were important because they indicated the extent to which motivation and question characteristics were indirectly related to exam scores.

Analysis B

Analysis B (illustrated in Figure 3.2) tested a path model in which four exogenous variables—Wave 1 mastery approach, mastery avoidance, and performance goal orientations and teacher emphasis on mastery goals—predicted two endogenous variables: *Value for Self* and *Value for Others*. The model also specified that *Value for Self* and *Value for Others* predicted Exam 2 scores.

The rationale for this model is that, as discussed in Chapter 2, the different purposes behind achievement strivings may be related to how students evaluate their questions. How students evaluate their questions may be related to the quality of students' questions or trigger information seeking behaviors that lead to improved understanding; therefore appraisals may be related to exam performance. As I noted previously, I lacked statistical power to test a single model that included all motivation variables, question characteristics, and exam scores. I used bivariate correlations to identify these specific motivation variables and question characteristics.

Analytic strategy for evaluating model fit. Analytic strategy for evaluating model fit in Analysis B followed the criteria described for Analysis A. In the case of a good fitting model, I interpreted the model parameters to address the research questions. The following parameters were of particular interest. First, the significance and relative strength of paths from exogenous variables (individual goal orientations and teacher emphasis on mastery goals) to endogenous variables (self-appraisals) were important because these indicated the extent to which each motivational variable uniquely explained variance in students' appraisals of their questions. Second, the significance and relative strength of paths from self-appraisal to exam scores were important because these indicated the extent to which each self-appraisal uniquely explained variance in exam scores. Third, the indirect relations from goal orientation variables to exam scores via self-appraisals were important because they indicated the extent to which goal orientation variables indirectly related to exam scores.

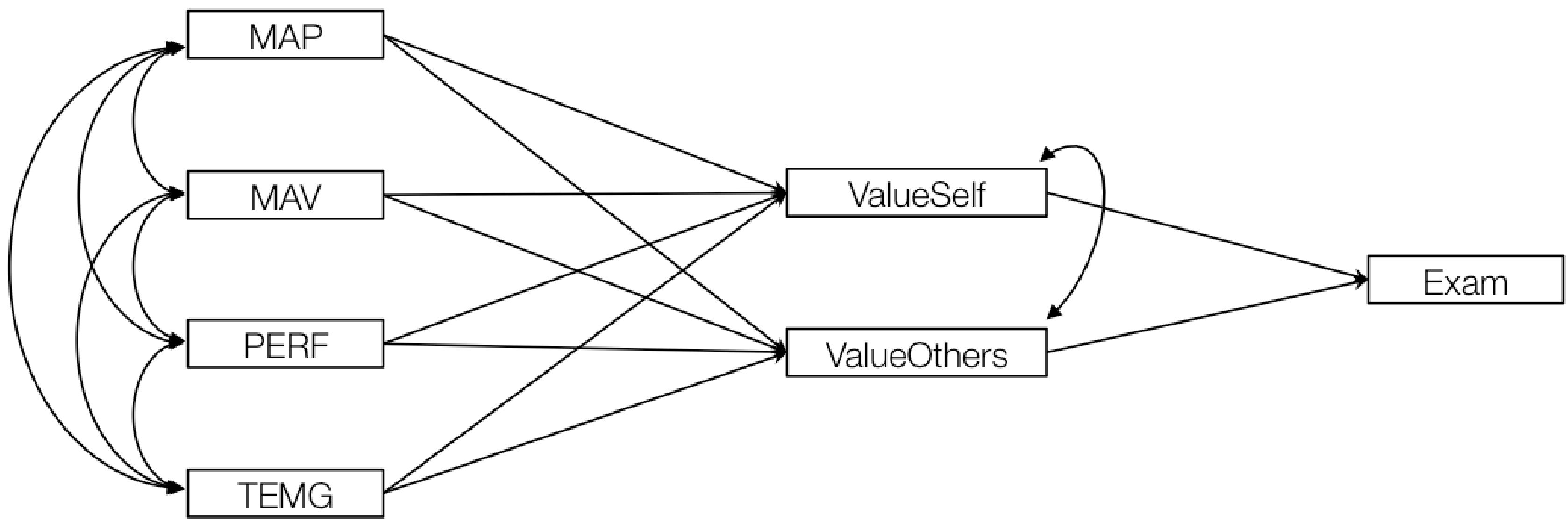

Figure 3.2. Path Model of Individual and Perceived Teacher Goal Orientations Predicting Self-Appraisals and Exam Scores.

Note. MAP = Mastery Approach Goal Orientation; MAV = Mastery Avoidance Goal Orientation; PERF = Performance Goal Orientation; TEMG = Teacher emphasis on mastery goals.

Statistical Power for Path Models

As a rule of thumb, path models require approximately 5-10 people per estimated parameter (Hancock & Mueller, 2010). In the model tested in Analysis A, there are 17 parameters estimated, requiring a minimum of 85-170 participants. In the model tested in Analysis B, there are 20 parameters estimated, requiring a minimum of 100-200 participants. Therefore, the sample size (N = 103) meets the minimum requirements for the number of participants relative to the number of estimated parameters. However, the path analyses may lack statistical power to detect small effects and may be vulnerable to idiosyncratic characteristics of the sample (points discussed in the Limitations section of Chapter 5).

Commonality Analyses

Commonality analysis calculates the unique and overlapping variances explained by predictor variables and can reveal the location and effect of collinearity and suppression effects (Zientek & Thompson, 2009). Therefore, commonality analysis provides a more complete accounting for the relationship in regressions compared to interpreting beta weights or structural coefficients alone (Nimon, 2010).

I conduct four commonality analyses as follow up analyses to further explore the unique and shared variance for variables that were significantly predicted by more than variable in the tested path models (Analysis A and B). Regarding Analysis A, I conducted commonality analyses examining the unique, shared, and total variance of a) post-class self-efficacy and b) exam scores. Regarding Analysis B, I conducted commonality analyses examining the unique, shared, and total variance of c) *Value for Others* and d) exam scores.

Regression of Exam 2 Scores on Question Characteristics

I conducted a multiple regression in which Exam 2 scores were regressed on all question characteristics. The rationale for this analysis is that the path models tested in Analysis A and Analysis B (described above) divided the question characteristics across two models. The multiple regression allowed me to examine the extent to which question characteristics collectively predicted exam scores.

Differences in Question Characteristics for Students with Normative Change in Motivation Constructs

To examine the relations among question characteristics and motivation constructs over time, I conducted independent-samples *t* tests that examined differences in question characteristics between students with normative growth versus decline on motivation variables from Wave 1 to Wave 2 (for how normative change categories were created, see the Normative Change in Motivation section above).

Research Question 4

Research question 4 asked: Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, achievement, and the quantity, qualities, and self-appraisals of generated questions?

I had originally intended to compared students who asked questions during class with students who did not ask questions during class in terms of their Wave I motivation and question characteristics. However, students *asked* very few questions during lectures. In the 8 class meetings observed for this study, six students asked a total of seven questions. This means that 97 students (94%) did not ask a question during class meetings. This substantial difference in the number of students who asked and did not ask questions prohibited the use of inferential statistics to examine differences between askers and non-askers. Thus, the proposed independent samples *t*-tests were not conducted. Instead, I presented descriptive statistics for motivational variables and question characteristics comparing *askers* and *non-askers.*

Person-Centered Analyses

Latent classes refer to subgroups or clusters of individuals within the sample who share similarities across variables. Analyses using latent classes—sometimes called person-centered or profile-centered analyses—have some advantages over variable-centered analyses since latent classes capture the overlapping and co-occurring nature of constructs and/or behaviors. Person-centered analyses that use cluster analytic techniques such as Latent Class Analysis (LCA) have been commonly used in motivational research, including research on achievement goals (e.g., Meece & Holt, 1993), values (e.g., Conley, 2012), and help seeking (e.g., Karabenick, 2003). For example, person-centered analyses have been useful in revealing common patterns of multiple goal orientations, such as when students endorse both mastery and performance goals (e.g., Tuominen-Soini, Salmela-Aro, & Niemivirta, 2011).

LCA is a model-based alternative to traditional forms of cluster analyses. Madgidson and Vermunt (2002) noted that LCA has several advantages over alternative cluster analytic techniques, including less arbitrary cluster criteria and rigorous statistical tests for competing models. In addition, the authors found that LCA was more accurate matching latent class membership to actual group membership in a simulation study, even when the simulated data was generated to be aligned with the assumptions of the tradition (i.e., K-means) clustering technique.

To evaluate the evidence regarding whether there were latent classes in the sample, I conducted three LCAs on Wave 1 motivational constructs alone (LCA1); question characteristics alone (LCA2), and Wave 1 motivational constructs and question characteristics (LCA3). For each LCA, I used the protocol and criteria described below to make decisions about the number of classes. In the event that there was evidence of more than one latent class (as was the case with LCA3, but not LCA1 and LCA2), both variable-centered and profile-centered analyses were used to answer RQ3a.

To determine the number of latent classes in a sample, I tested models in sequential order starting with 1 latent class (i.e., the entire sample), followed by 2 latent classes, followed by 3 latent classes, and so on (Muthén & Muthén, 2012). For each increase in the number of latent classes, I examined fit indices to determine if there was compelling evidence to accept a less parsimonious model (i.e., more latent classes) in favor of a better fitting model. In addition to examining model fit, I evaluated latent classes in terms of their practical and theoretical usefulness since latent classes should differ in ways that are of theoretical and/or practical interest to the focus of the study, such as having a meaningful number of cases in each latent class (Kline, 2011).
The argument for the number of latent classes was based on a constellation of fit statistics that indicated the adequacy of the model relative to more or less parsimonious models (i.e., with more or fewer classes). Following suggestions by Kline (2011), I used information criterion statistics, the Entropy index, the number of participants in each latent class, and the adjusted likelihood ratio test to evaluate the fit of the model relative

to other possible models. For the Akaike information criterion (AIC; Akaike, 1987), Bayesian information criterion (BIC; Schwarz, 1978), and the sample-size adjusted Bayesian information criterion (ABIC; Sclove, 1987), a smaller statistic indicates a better fitting model. The Entropy index indicates the ‘cleanness’ with which participants can be classified into one latent group or another. Higher entropy values indicate cleaner classifications of individuals into latent groups; entropy values over .80 represent adequately clean classification (Kline, 2011). As mentioned above, the number of participants assigned to a given latent class should be practically and theoretically meaningful. The addition of a latent class with minimal membership may not be practically or theoretically useful, despite a possibly improved model fit. The adjusted likelihood ratio test (LMR-LRT; Lo, Mendell, & Rubin, 2001) indicates whether a model with more latent classes has a significantly better fit compared to a more parsimonious model, above and beyond what is gained by fewer degrees of freedom (Nylund, Asparouhov, & Muthén, 2004).

CHAPTER 4

RESULTS

Descriptive Results

Motivation Variables

Descriptive statistics for motivation variables at Wave 1, after class meetings (post-class), and Wave 2 are presented in Table 4.1.

Motivations at Wave 1

At Wave 1, the following variables were relatively normally distributed: Interest; mastery approach goal orientation (MAP), mastery avoidance goal orientation (MAV), performance goal orientation (PERF); self-efficacy; teacher emphasis on mastery goals (TEMG); help-seeking avoidance (HSAvoid); executive help seeking (HSExec); instrumental help seeking (HSInstr); and preference for a formal target when seeking help (HSFormal). Many of the Wave 1 variables were slightly negatively skewed indicating that response tended to cluster toward the upper end of the scales. Two Wave 1 variables—help-seeking threat (HSThreat) and intention to remain in STEM (StaySTEM)—were relatively leptokurtic with high kurtosis statistics of 3.49 and 9.27, respectively. I do not use any of these variables in inferential statistical analyses. All Wave 1 scales demonstrated adequate to strong internal reliability with the exception of two variables: executive help seeking ($\alpha = .53$) and help-seeking threat ($\alpha = .51$). Help seeking variables are used only descriptively in reference to research question 4 and those statistics should be interpreted conservatively in light of the low reliability (see the final section of this chapter).

Post-Class Motivations and Perceptions

Descriptive statistics for post-class motivations and perceptions are presented in Table 4.1. Post-class perceptions of situational interest, self-efficacy for learning the instructed material, and perceived comprehension of the instructed material were normally distributed. The mean for situational interest indicated that, on average, students reported that they found class meetings moderately interesting. The mean for post-class self-efficacy indicated that, on average, students reported strong efficacy beliefs for learning the instructed material. The mean for perceived comprehension indicated that, on average, students reported moderate to high levels of understanding after class meetings. Reliability statistics were not calculated for post-class variables because each is a single-item scale. While an individual score on any post-class scale could vary from class meeting to class meeting (e.g., having more or less self-efficacy for learning the content present during a given class meeting), scores across class meetings were relatively stable (see ICCs reported in chapter 3).

Motivation at Wave 2

Descriptive statistics for Wave 2 motivational variables are presented in Table 4.1. Mean scores for motivation variables at Wave 2 were similar to mean scores at Wave 1. Wave 2 variables were normally distributed with the exception of Intention to remain in STEM (StaySTEM), which was negatively skewed and leptokurtic. As with this variable in Wave 1, the Wave 2 StaySTEM variable was not used in analyses with inferential statistics. Wave 2 variables had acceptable-to-strong internal reliability.

Normative Change in Motivation Constructs from Wave 1 to Wave 2

Frequencies for students who increased versus decreased on motivation variables are presented in Table 4.2. For each motivational variable, more students increased than decreased in their endorsement of motivational constructs. Recall that this categorical variable (increased vs. decreased) was created using residualized gain scores and refers to a students' growth or decline relative to the whole sample.

Table 4.1

Descriptive Statistics for Motivation Variables

Variable	Min.	Max.	*M*	*SD*	Skew.	Kurt.	α
Wave 1 Variables (*n* = 102)							
Interest (W1)	1.13	5.00	3.04	1.00	-0.163	-0.874	0.93
MAP (W1)	1.67	5.00	4.45	0.63	-1.460	2.994	0.70
MAV (W1)	1.00	5.00	3.62	1.17	-0.620	-0.607	0.72
PERF (W1)	1.00	5.00	4.19	0.93	-1.409	1.812	0.90
Self-Efficacy (W1)	2.20	5.00	4.20	0.72	-0.686	-0.286	0.85
TEMG (W1)	1.60	5.00	3.80	0.65	-0.428	0.563	0.82
StaySTEM (W1)	1.00	5.00	4.64	0.75	-2.890	9.274	0.91
HSExec (W1)	1.00	4.00	1.69	0.67	0.880	0.335	0.51
HSInstr (W1)	1.60	5.00	4.15	0.86	-1.263	1.067	0.84
HSThreat (W1)	1.00	4.00	1.41	0.61	1.819	3.486	0.67
HSAvoid (W1)	1.00	3.00	1.46	0.59	1.144	0.209	0.51
HSFormal (W1)	1.67	5.00	3.53	0.93	-0.029	-0.811	0.71

Note. Cronbach's α were not calculated for post-class variables due to the different number of question logs completed by participants (see section on ICCs in Methods section). Skew = skewness; Kurt. = Kurtosis; W1 = Wave 1; MAP = Mastery approach goal; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; StaySTEM = Intention to remain in a STEM major; HSExec = Executive help seeking; HSInstr = Instrumental help seeking; HSThreat = Help seeking threat; HSAvoid = Help seeking avoidance; HSFormal = Preference for help seeking from a formal target;

(Table 4.1 continued)

Variable	Min.	Max.	*M*	*SD*	Skew.	Kurt.	α
Post-Class Variables[a] (*n* = 103)							
Sit. Interest (PC)	1.67	10.00	6.54	1.78	-0.435	-0.148	—
Self-Efficacy (PC)	2.89	10.00	8.06	1.52	-0.877	0.604	—
PercComp (PC)	1.89	10.00	7.25	1.62	-0.764	0.596	—
Wave 2 Variables (*n* = 95)							
Interest (W2)	1.00	4.88	3.14	0.99	-0.377	-0.685	0.93
MAP (W2)	2.00	5.00	4.37	0.67	-1.086	0.839	0.75
MAV (W2)	1.00	5.00	3.59	1.14	-0.430	-0.797	0.72
PERF (W2)	1.00	5.00	4.14	0.95	-1.151	0.806	0.94
Self-Efficacy (W2)	2.60	5.00	4.24	0.73	-0.684	-0.667	0.91
TEMG (W2)	2.00	5.00	4.22	0.64	-0.944	0.802	0.86
StaySTEM (W2)	1.00	5.00	4.46	0.86	-2.395	5.685	0.91

Note. [a]Mean scores across question logs. Cronbach's α were not calculated for post-class variables due to the different number of question logs completed by participants (see section on ICCs in Methods section). Skew = skewness; Kurt. = Kurtosis; PC = Post-class; W2 = Wave 2; MAP = Mastery approach goal; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; StaySTEM = Intention to remain in a STEM major. PercComp = Perceived comprehension of the instructed material; Sit = Situational.

Table 4.2

Frequencies of Students Who Increased Versus Decreased on Motivational Variables from Wave 1 to Wave 2

Variable	*N*	Increased	Decreased
Interest	94	55.3%	44.7%
MAP	92	58.7%	41.3%
MAV	91	54.9%	45.1%
PERF	94	60.6%	39.4%
Self-Efficacy	93	61.3%	38.7%
TEMG	94	54.3%	45.7%
StaySTEM	94	64.9%	35.1%

Note. MAP = Mastery approach goal; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; StaySTEM = Intention to remain in a STEM major.

Exam Scores

Descriptives for exam scores are presented in Table 4.3. Exam scores before (Exam 1) and after (Exam 2) the use of the question logs were normally distributed.

Table 4.3

Descriptive Statistics for Scores on Exams 1 and 2 (in Percent)

Variable	*N*	Min.	Max.	*M*	*SD*	Skewness	Kurtosis
Exam 1	104	20	96	70.12	14.35	-0.838	0.943
Exam 2	99	16	100	70.91	18.14	-0.639	-0.069

Characteristics of Student-Generated Questions

Descriptive statistics for question characteristic variables are presented in Table 4.4.

Quantity of Conceptual and Procedural Questions. Students generated a mean of 9.5 questions during the 8-lecture unit and asked, on average, a statistically significantly greater number of conceptual questions ($M = 5.45$, $SD = 4.02$) compared to procedural questions ($M = 4.02$, $SD = 2.53$), $t(102) = 3.312$, $p = .001$. There was a large amount of variability in the quantity of questions, particularly for the number of conceptual questions. Approximately one-quarter (27%) of all questions were resolved for students between the time the question was recorded and the end of the class meeting, according to student self-reports.

Table 4.4

Descriptive Statistics for Question Characteristics

Variable	*N*	Min.	Max.	*M*	*SD*	Skewness	Kurtosis
No. Conceptual	103	0	21	5.45	4.02	1.070	1.480
No. Procedural	103	0	12	4.02	2.53	0.405	-0.198
% Unanswered	103	0	1	0.72	0.25	-0.960	0.654
% Closed	103	0	1	0.32	0.24	0.873	0.489
ValueSelf	103	1.67	5	3.89	0.64	-0.734	0.527
ValueOthers	103	1.67	5	3.43	0.69	-0.194	<.001

Self-Appraisals of Questions. Students tended to perceive value in their questions in terms of the usefulness of their questions for their own learning as well as the learning of their peers. Mean appraisal of *Value for Self* was 3.89 (*SD* = 0.64) indicating that students perceived the answers to their questions as being, on average, *somewhat* to *very useful* for their own learning. Mean appraisal of *Value for Others* was 3.43 (*SD* = 0.69), indicating that students perceived the answers to their questions as being, on average, *somewhat useful* to *useful* for the understanding of their classmates. Students tended to perceive their questions as being more valuable to their own learning than to the learning of their classmates.

Question syntax. Descriptives statistics and frequencies for question characteristics are presented in Table 4.5 and Figure 4.1, respectively. Approximately two-thirds of questions were phrased with open syntax (e.g., “What is the difference between a limiting reactant and a limiting reagent?”) and approximately one-third of questions were phrased with closed syntax (e.g., “Is there a difference between a reagent and a reactant?” See Table 3.6 for more examples). Only a small percentage of questions (4%) were phrased as assertions. The percentages of open and closed syntax were normally distributed while the percentage of assertions had high kurtosis. In order to simplify analyses, avoid dramatically different-sized groups, and avoid problems associated with non-normally distributed variables, questions that have an open syntax and questions phrased as assertions were collapsed into a single category, which I will refer to as "open", since both types of phrasing solicit an open response to the posed question.

Table 4.5

Descriptive Statistics for Open-Closed-Assertion Question Codes

Variable	N	Min.	Max.	*M*	*SD*	Skew.	Kurt.
% Closed	103	0	1.0	0.32	0.24	0.873	0.489
% Open	103	0	1.0	0.64	0.24	-0.536	-0.001
% Assertions	103	0	0.5	0.04	0.09	2.896	9.012
% CombinedOpen	103	0	1.0	0.68	0.24	-0.873	0.489

Note. Skew = Skewness; Kurt = Kurtosis. % CombinedOpen = %Open + % Assertions.

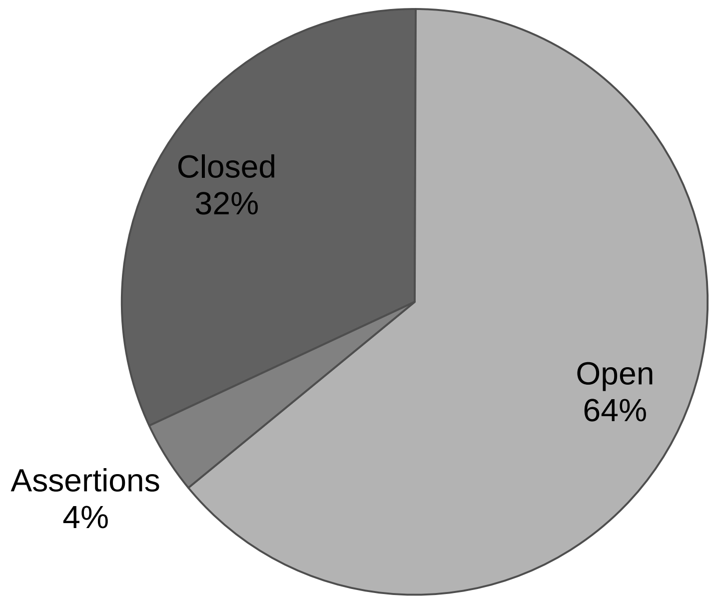

Figure 4.1. Frequency of Open-Closed-Assertion Codes

Question types. Descriptive statistics and frequencies for question characteristics are presented in Table 4.6 and Figure 4.2, respectively. The most frequently-occurring question types were definitions (12%), general procedural (11%), conditional procedural (9%), rationale procedural (9%), enablement procedural, (8%), justifications (7%), relations/comparisons (7%), causal relations (6%) and concept completion (6%). Procedural questions were relatively evenly divided among the five types of procedural question: General procedural (25.3% of all procedural questions), Rationale procedural (22.0%), Conditional procedural (20.0%), Enablement procedural (19.1%), and Strategy evaluation procedural (13.5%). Some question types were very infrequent (e.g., one instance each of quantification and exemplification questions; see Table 4.6 for descriptives on all coding categories). Across question types, frequencies tended to be very low, and for most question types, the mode was zero. As is common for many count variables, distributions for individual question types followed a Poisson distribution, with most individuals having zero occurrences of a given type of question and with the frequencies of individuals decreasing as the number of occurrences increase. When question types are grouped into two categories—procedural (43%) and conceptual (57%) questions—distributions are bell-shaped, as indicated by the descriptive statistics for *All Procedural* and *All Conceptual* at the bottom of Table 4.6.

Table 4.6

Descriptive Statistics for Question Types Using a Modified GPH Coding Scheme

Question Type	Sum	*M*	Median	Mode	Max.
Definitions	118	1.15	0	0	14
Relational/Comparative	70	0.68	0	0	4
Justification	68	0.66	0	0	5
Causal	62	0.60	0	0	4
Concept Completion	60	0.58	0	0	4
Concept Extension	52	0.51	0	0	4
Feature specification	39	0.38	0	0	3
Purpose/Value	37	0.36	0	0	4
General	29	0.28	0	0	4
Enablement	24	0.23	0	0	2
Quantification	1	0.01	0	0	1
Examples	1	0.01	0	0	1
Procedural (General)	105	1.02	1	0	5
Procedural (Rationale)	91	0.88	1	0	3
Procedural (Conditional)	83	0.81	0	0	4
Procedural (Enablement)	79	0.77	0	0	5
Procedural (Strategy Eval.)	56	0.54	0	0	3
Aggregated Types					
All Conceptual	561	5.45	4	2	21
All Procedural	414	4.02	4	5	12

Note. Eval = Evaluation.

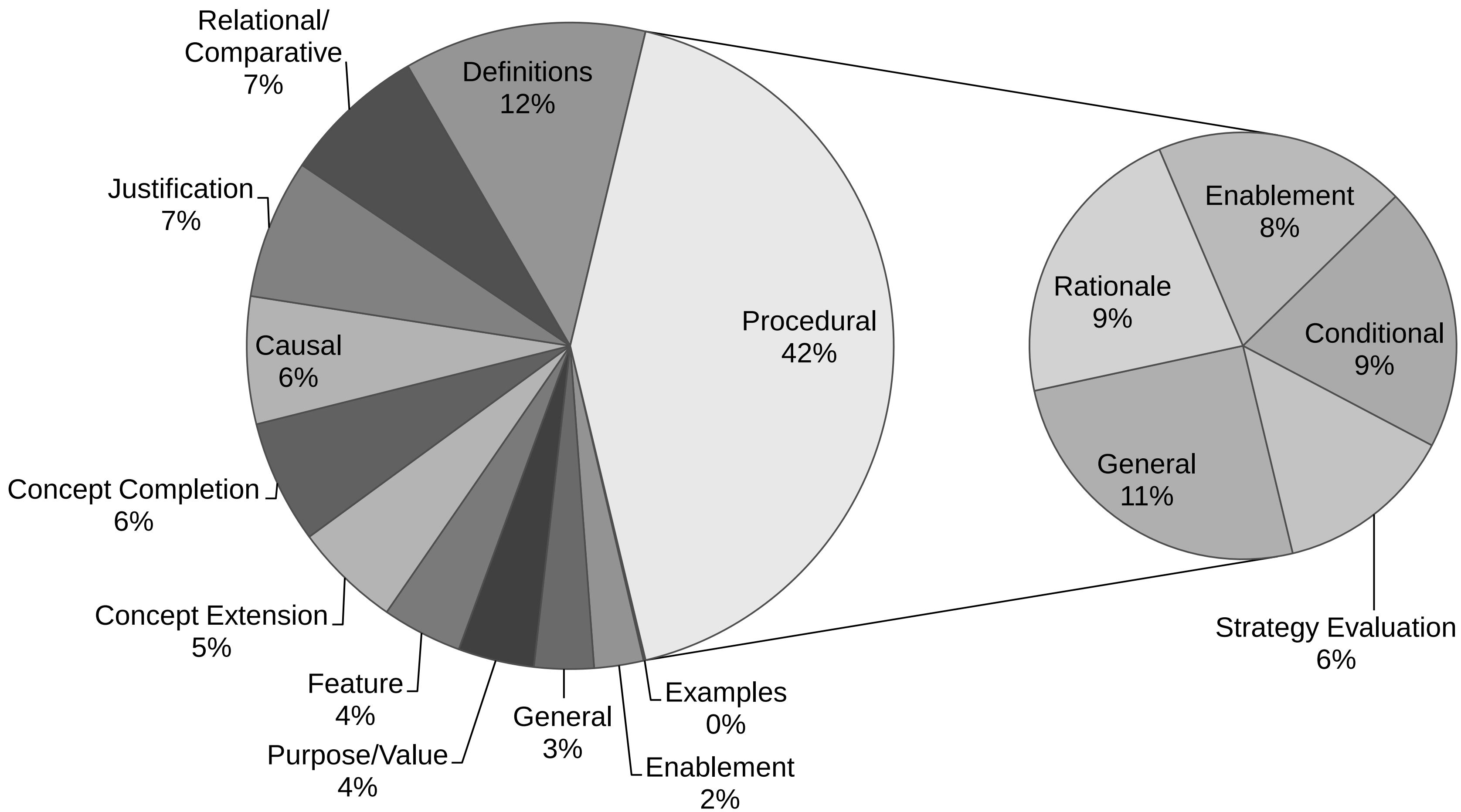

Figure 4.2. Frequencies of Question Types Using a Modified GPH Coding Scheme

Correlations

Bivariate Pearson correlations among all variables are presented in Table 4.7. In the section below, I highlight correlations among question characteristics and with motivation variables.

Relations Among Question Characteristics

Significant correlations among question characteristics

The *number of conceptual questions* had a significant weak negative correlation with *percentage of closed questions* ($r[100] = -.23$). Students who generated more conceptual questions tended to generate a smaller percentage of questions with closed syntax compared to open syntax. One reason why conceptual questions negatively associated with closed questions is that in this context conceptual questions were less likely to use closed syntax, an explanation that is supported by the chi square test presented below.

Value for Self had a significant strong positive correlation *Value for Others* ($r[100] = .68$). Thus, approximately 47% of the variance in one self-appraisal is accounted for by variance in the other self-appraisal. That over half of the variance in *Value for Self* is not accounted for in *Value for Others* suggests that the two self-appraisals are related but different constructs. Further evidence for the distinction between the two appraisals is found in the results from regression, path and commonality analyses and partial correlations described in the following section. Due to relatively strong correlation between the two appraisals, multiple methods were used to ensure that problems of multicollinearity were not present (see Screening for Assumptions in Chapter 3).

Table 4.7

Zero-Order Bivariate Pearson Correlations among Variables

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
1.	No. Conceptual	—											
2.	No. Procedural	.17	—										
3.	% Unanswered	-.03	.19	—									
4.	% Closed	**-.23**	-.05	.00	—								
5.	ValueSelf	-.08	.10	.12	-.11	—							
6.	ValueOthers	.06	.11	.13	-.01	**.68**	—						
7.	Interest (W1)	-.17	**-.22**	-.06	.03	.05	.07	—					
8.	Self-Efficacy (W1)	-.05	**-.33**	-.07	.10	-.03	.04	**.39**	—				
9.	MAP (W1)	.04	-.08	.13	.02	**.23**	**.29**	**.44**	**.45**	—			
10.	MAV (W1)	.07	-.09	**.21**	.07	**.24**	**.31**	**.24**	.16	**.32**	—		
11.	PERF (W1)	-.03	-.09	.15	.08	**.23**	**.31**	.11	.15	**.41**	**.36**	—	
12.	TEMG (W1)	.07	-.07	.04	.10	.03	.01	.17	**.34**	**.29**	.16	.18	—

Note. Correlations significant at $p < .05$ denoted with bold font. W1 = Wave 1; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals.

(Table 4.7 continued)

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
13.	StaySTEM (W1)	-.02	.07	-.15	.16	-.11	.03	.00	.09	.04	.02	.14	.15
14.	HSExecutive (W1)	.01	.14	-.12	-.01	-.17	-.18	-.11	-.16	**-.25**	-.09	-.11	-.05
15.	HSInstrument (W1)	.04	.00	.11	.11	**.28**	**.29**	.11	.199	**.36**	**.34**	**.22**	**.29**
16.	HSThreat (W1)	-.05	.14	.10	**-.21**	-.02	-.04	.03	**-.22**	-.08	-.12	.11	-.11
17.	HSFormal (W1)	.02	-.03	.05	-.16	-.16	-.14	.03	**.24**	.12	.02	-.10	.12
18.	Perc. Comp. (PC)	-.09	-.18	**-.30**	.24	-.15	-.10	.09	**.23**	-.02	-.04	-.10	.05
19.	Interest (PC)	-.05	.06	-.05	.02	.12	.01	**.33**	.03	.08	.15	.11	.21
20.	Self-Efficacy (PC)	.02	**-.30**	**-.20**	**.23**	.05	.04	.10	**.34**	.12	.02	-.05	.09
21.	Interest (W2)	-.11	-.10	-.06	.12	.14	.16	**.80**	**.27**	**.44**	**.27**	.17	.17
22.	Self-Efficacy (W2)	.05	-.17	-.01	.13	.07	.13	.18	**.51**	**.35**	.18	**.28**	**.21**
23.	MAP (W2)	.05	.09	**.23**	.04	**.24**	**.30**	**.21**	**.28**	**.64**	**.26**	**.29**	**.27**
24.	MAV (W2)	.06	.09	.10	.08	.14	**.21**	**.27**	**.22**	**.45**	**.43**	**.31**	.20

Note. Correlations significant at $p < .05$ denoted with bold font. W1 = Wave 1; PC = Post-class perceptions; W2 = Wave 2; StaySTEM = Intention to stay in STEM; HS = Help seeking; Formal = Formal help seeking target; Perc. Comp. = Perceived Comprehension; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation.

(Table 4.7 continued)

	Variable	13	14	15	16	17	18	19	20	21	22	23	24
13.	StaySTEM (W1)	—											
14.	HSExecutive (W1)	-.11	—										
15.	HSInstrumental (W1)	-.13	-.11	—									
16.	HSThreat (W1)	.08	-.01	**-.36**	—								
17.	HSFormal (W1)	-.02	-.13	-.04	.15	—							
18.	Perc. Comp. (PC)	.02	.08	.06	**-.27**	-.04	—						
19.	Sit. Interest (PC)	-.10	.07	.07	-.12	.03	**.49**	—					
20.	Self-Efficacy (PC)	-.03	-.06	.06	-.12	-.02	**.58**	**.24**	—				
21.	Interest (W2)	-.04	-.02	**.22**	-.06	-.01	.19	**.45**	.14	—			
22.	Self-Efficacy (W2)	-.08	-.02	**.41**	**-.34**	.02	**.26**	**.32**	**.43**	**.40**	—		
23.	MAP (W2)	.10	-.11	**.36**	-.10	.11	.12	.20	.17	**.40**	**.44**	—	
24.	MAV (W2)	.02	.02	.12	.07	-.10	-.06	.08	.06	**.33**	.19	**.43**	—

Note. Correlations significant at $p < .05$ denoted with bold font. W1 = Wave 1; PC = Post-class perceptions; W2 = Wave 2; StaySTEM = Intention to stay in STEM; HS = Help seeking; Formal = Formal help seeking target; Perc. Comp. = Perceived Comprehension; Sit = Situational; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation.

(Table 4.7 continued)

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
25.	PERF (W2)	-.05	.12	**.24**	.05	**.27**	**.25**	.15	.06	**.47**	**.29**	**.65**	.11
26.	TEMG (W2)	.10	.07	-.07	.06	.09	.09	.07	**.23**	**.24**	.18	**.21**	**.62**
27.	StaySTEM (W2)	.10	-.07	-.10	.10	-.07	.04	-.05	.07	.04	.13	.15	.15
28.	Interest (Chg)	.09	.09	.06	.15	.13	**.21**	.14	.09	**.22**	**.24**	**.23**	.13
29.	Self-Efficacy (Chg)	-.03	-.09	.03	.14	.09	.15	.06	.14	.12	.04	.19	-.04
30.	MAP (Chg)	.08	.15	.14	.03	.09	.06	-.04	.10	.03	.11	-.02	.11
31.	MAV (Chg)	-.01	-.01	-.02	.07	.16	**.21**	**.32**	**.30**	**.45**	.16	.15	**.24**
32.	PERF (Chg)	.06	**.35**	.18	-.12	.12	.03	.03	-.08	.18	.11	.03	.09
33.	TEMG (Chg)	.11	.07	.01	-.09	.04	.04	-.03	-.02	.00	-.02	.18	-.05
34.	StaySTEM (Chg)	.11	-.06	-.09	.07	-.04	.02	-.08	.03	.01	.11	.12	.11
35.	HSCalc	.09	-.10	-.03	.05	.10	.19	-.01	.08	.10	.01	.09	.12
36.	GPA	-.11	.12	-.04	.13	-.04	**.22**	-.02	.12	-.05	-.04	-.07	.13

Note. Correlations significant at $p < .05$ denoted with bold font. W2 = Wave 2; Chg = Growth from W1 to W2 (1 = increase; 0 = decrease); PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; STEM = Intention to stay in STEM; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; HSCalc = High school calculus.

(Table 4.7 continued)

	Variable	13	14	15	16	17	18	19	20	21	22	23	24
25.	PERF (W2)	.18	.10	.12	.07	-.17	-.09	.19	-.04	**.24**	.19	**.42**	**.45**
26.	TEMG (W2)	.05	.05	**.38**	**-.23**	.02	.17	.34	.15	**.29**	**.53**	**.38**	**.27**
27.	StaySTEM (W2)	**.28**	-.17	**.26**	-.19	.07	-.07	-.15	-.09	.05	**.20**	.03	.03
28.	Interest (Chg)	-.14	.05	**.22**	-.07	.01	.09	**.27**	.14	**.57**	**.30**	**.36**	**.32**
29.	Self-Efficacy (Chg)	-.19	.11	**.28**	**-.29**	-.09	**.22**	**.29**	**.40**	**.28**	**.78**	**.32**	.09
30.	MAP (Chg)	.11	.08	.19	-.05	.11	**.30**	**.26**	.12	.10	.20	**.63**	.17
31.	MAV (Chg)	.11	-.05	.14	-.02	-.15	.07	.07	.14	**.30**	**.23**	**.48**	**.79**
32.	PERF (Chg)	.03	**.27**	.05	.06	-.11	-.09	.09	-.10	.13	-.06	**.28**	**.30**
33.	TEMG (Chg)	-.13	**.24**	.22	-.13	-.05	.16	**.26**	.00	**.21**	**.34**	**.23**	.01
34.	StaySTEM (Chg)	.00	-.12	**.28**	**-.21**	.06	-.06	-.12	-.10	.06	**.24**	-.01	.02
35.	HSCalc	-.18	.09	.02	.07	.08	-.02	-.08	.06	.01	.01	.08	.16
36.	GPA	-.05	.08	-.04	-.07	.14	.15	.14	.19	.12	.17	.08	.21

Note. Correlations significant at $p < .05$ denoted with bold font. W2 = Wave 2; Chg = Growth from W1 to W2 (1 = increase; 0 = decrease); PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; STEM = Intention to stay in STEM; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; HSCalc = High school calculus.

(Table 4.7 continued)

	Variable	25	26	27	28	29	30	31	32	33	34	35	36
25.	PERF (W2)	—											
26.	TEMG (W2)	**.24**	—										
27.	STEM (W2)	.04	**.28**	—									
28.	Interest (Chg)	**.28**	**.38**	.09	—								
29.	Self-Efficacy (Chg)	.12	**.33**	.03	.26	—							
30.	MAP (Chg)	.11	**.31**	.01	.25	.20	—						
31.	MAV (Chg)	**.32**	**.26**	.04	.16	.14	.20	—					
32.	PERF (Chg)	**.65**	.20	-.08	**.24**	-.06	**.23**	.19	—				
33.	TEMG (Chg)	.20	**.58**	.19	**.29**	**.39**	**.28**	.02	**.22**	—			
34.	StaySTEM (Chg)	-.01	**.28**	**.96**	.14	.09	-.03	.01	-.09	**.24**	—		
35.	HSCalc	.06	.02	.14	.20	-.02	-.06	.15	.13	.04	.19	—	
36.	GPA	-.11	.19	.01	**.30**	**.22**	.08	.18	-.07	.10	.02	.19	—

Note. Correlations significant at $p < .05$ denoted with bold font. W2 = Wave 2; Chg = Growth from W1 to W2 (1 = increase; 0 = decrease); PERF = Performance goal orientation; TEMG = Teacher emphasis on mastery goals; STEM = Intention to stay in STEM; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; HSCalc = High school calculus.

(Table 4.7 continued)

	Variable	1	2	3	4	5	6	7	8	9	10	11	12
37.	Exam1	.05	.01	-.02	**.20**	-.11	-.01	-.01	.13	.04	.18	.10	.00
38.	Exam2	-.16	-.01	-.03	**.34**	-.10	.13	.05	**.24**	.06	-.03	.11	.01

Note. Correlations significant at $p < .05$ denoted with bold font.

(Table 4.7 continued)

	Variable	13	14	15	16	17	18	19	20	21	22	23	24
37.	Exam1	**.22**	-.02	-.07	-.05	.04	**.23**	.09	**.35**	.10	**.33**	.03	.17
38.	Exam2	**.24**	-.05	.12	**-.21**	.02	**.41**	.12	**.28**	.24	**.47**	**.21**	.11

Note. Correlations significant at $p < .05$ denoted with bold font.

(Table 4.7 continued)

	Variable	25	26	27	28	29	30	31	32	33	34	35	36	37
37.	Exam1	.01	.12	.08	.08	**.22**	-.08	.00	-.14	.02	.07	.08	**.25**	—
38.	Exam2	.04	**.33**	.12	.19	**.37**	.20	.14	-.06	**.28**	.11	-.03	**.32**	**.57**

Note. Correlations significant at $p < .05$ denoted with bold font.

Relations among question type and syntax

Question syntax was distributed non-randomly across all question types as indicated by a chi square test on syntax by question type (2 [open vs. closed] x 13 [question types]), χ^2 [12, N = 980] = 205.570, p < .001. Similarly, syntax was distributed non-randomly across procedural question types (2 [open vs. closed] x 5 [procedural question types]), χ^2 [4, N = 415] = 228.217, p < .001. The chi square results indicated that some question types were more strongly associated with either a closed or open syntax. For example, nearly all questions about definitions used open syntax (e.g., "What is molarity?") while nearly all strategy evaluation questions used closed syntax (e.g., "Does it matter what unit of temperature you use for gas laws?").

Summary

There were few significant correlations among question characteristics. The number of conceptual questions had a significant weak negative correlation with the percentage of closed questions. Self-appraisal were strongly and positively correlated with each other. In general, however, question characteristic variables were independent of each other.

The magnitude of bivariate correlations among question characteristics and motivation constructs—especially at Wave 1 and post-class—were used to make decisions about the models to be tested. Quantity and percentage variables tended to have significant correlations with self-efficacy and interest at Wave 1 and with post-class self-efficacy and tended not to be significantly correlated with goal orientation variables. Self-appraisals, on the other hand, tended to be significantly correlated with goal orientations and not with interest and self-efficacy at Wave 1. Given the constraints on statistical power, two separate models were tested—one examining the relations among interest, self-efficacy, quantity and percentage variables, and exam scores (Analysis A) and one examining the relations among goal orientations, self-appraisals and exam scores (Analysis B; see section below).

Relations Among Question Characteristics, Motivation Constructs and Exam Scores: Variable-Centered Analyses

I conducted three sets of variable-centered analyses examining the relations among question characteristics, motivations, and exam scores; these analyses pertain to RQ2 and RQ3. Analysis A tests a path model examining the relations among quantity and percentage question variables, interest and self-efficacy, and exam scores. The path analysis is followed up with two commonality analyses. Analysis B tests a path model examining the relations among self-appraisals, goal orientations, and exam scores. The path analysis is followed up with two commonality analyses. Analysis C consists of a set of independent samples *t* tests examining the differences in question characteristics between students who increased versus decreased on motivations from Wave 1 to Wave 2.

Analysis A: Relations among Interest, Self-Efficacy, Selected Question Characteristics and Exam Scores

Analysis A tested three sets of relations pertaining to motivation, questions, and achievement: (1) the relations between Wave 1 interest and self-efficacy and question quantity and percentage variables; (2) the relations between question characteristics and post-class self-efficacy; and (3) the relations among question characteristics, post-class self-efficacy and exam scores. Fit statistics indicate that the model tested in Analysis A has excellent fit ($\chi 2 = 13.371$, $df = 15$, $p = .574$, CFI = 1.00, RMSEA = .000; SRMR = .051). Significant paths[1] are presented in Figure 4.3 and full model results are presented in Table 4.8.

Wave 1 self-efficacy negatively predicted the number of procedural questions ($\beta = -.342$). Students who reported having higher self-efficacy for the chemistry course tended to generate fewer procedural questions. However, the model accounted for a non-significant 10.1% of variance in the number of procedural questions. Thus, the negative association between self-efficacy and the number of procedural questions was relatively minor.

Post-class self-efficacy was negatively predicted by the number of procedural questions ($\beta = -.194$) and positively predicted by the percentage of closed questions ($\beta = .225$), after accounting for the variance explained by Wave 1 self-efficacy ($\beta = .255$). The paths from percentage of unanswered questions ($\beta = -.147$, $p = .10$) and the number of conceptual questions ($\beta = .111$, $p = .22$) to post-class self-efficacy were non-significant. The error variance between the number of conceptual questions and the percentage of closed questions was significantly and negatively correlated, indicating that residual variance (i.e., what is not explained) for the former is significantly and negatively correlated with residual variance of the latter. Given the small amount of variance in quantity of conceptual questions explained by the model, the correlation of the errors likely reflects the negative correlation between the variables themselves. This corroborates the finding from the chi square test that conceptual questions tended to have open rather than closed syntax.

[1] I interpret path coefficients as significant at $p < .05$. The variable *Exam 2* was used in four analyses, thereby increasing the potential for Type I error. A more conservative Bonferroni-corrected alpha level is $p < .0125$ ($p = .05/4$).

In sum, students who generated more procedural questions or who generated a higher percentage of open questions during class meetings reported lower self-efficacy for learning the material that was presented, even after controlling for students general self-efficacy for the course. Post-class self-efficacy and percentage of closed questions each significantly and positively predicted exam scores. The model accounted for 15.4% of the variance in exam scores and 21.7% of variance in post-class self-efficacy.

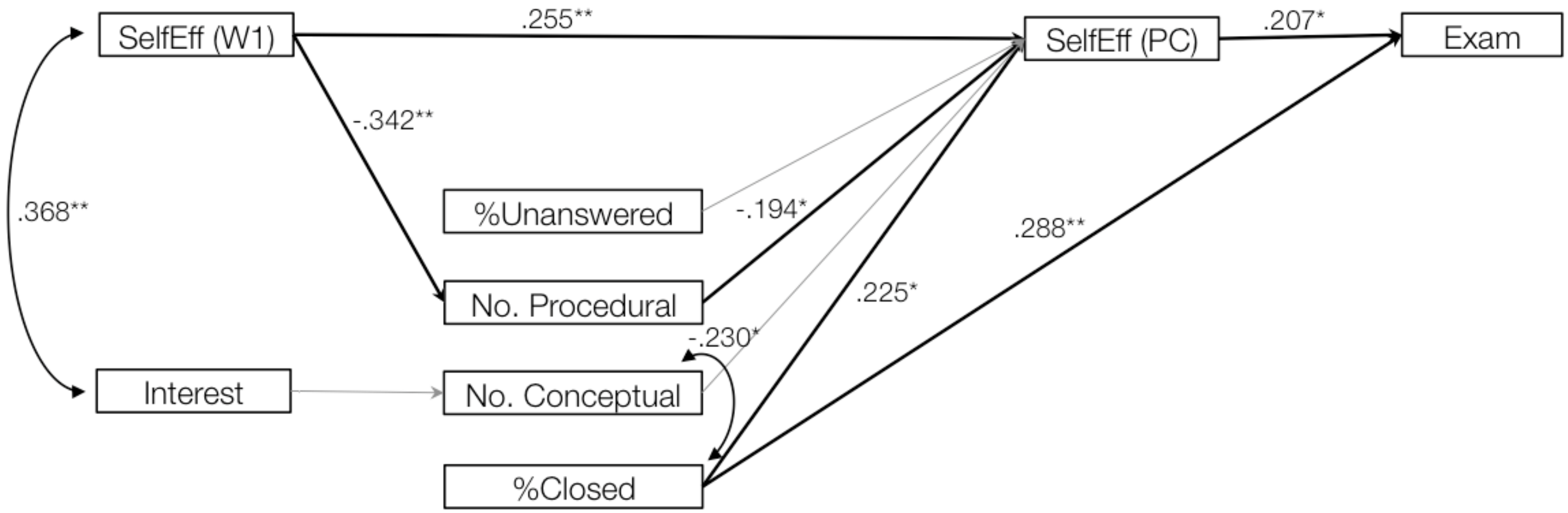

Figure 4.3. Path Model of Interest, Self-Efficacy and Question Characteristics Predicting Exam Scores.

Note. Errors for endogenous variables not shown. W1 = Wave 1; PC = Post-class; SelfEff = Self-Efficacy.
p < .05*;* ***p* < .01.

Table 4.8

Estimated Parameters for Model Tested in Analysis A

Parameters	β	*SE*	*p*
Direct Effects			
Self-Efficacy (PC) => Exam	.207	.093	.03
%Closed => Exam	.288	.091	< .01
No. Conceptual => Self-Efficacy (PC)	.111	.090	.22
No. Procedural => Self-Efficacy (PC)	-.194	.094	.04
%Unanswered => Self-Efficacy (PC)	-.147	.089	.10
%Closed => Self-Efficacy (PC)	.225	.088	.01
Self-Efficacy (W1) => Self-Efficacy (PC)	.254	.090	.01
Interest (W1) => No. Conceptual	-.175	.096	.07
Self-Efficacy (W1) => No. Procedural	-.324	.088	< .01
Indirect Effects			
SelfEff (W1) => SelfEff (PC) => Exam	.053	.030	.08
SelfEff (W1) => No. Proc => SelfEff (PC) => Exam	.013	.009	.17
No. Procedural => Self-Efficacy (PC) => Exam	-.040	.027	.13
%Closed => Self-Efficacy (PC) => Exam	.047	.028	.09
Correlated Errors			
No. Conceptual with %Closed	-.230	.093	.01

Note. PC = Post-class; W1 = Wave 1; SelfEff = Self-Efficacy; Proc = Procedural.

(Table 4.8 continued)

	β	*SE*	*p*
Covariance of Exogenous Variables			
Interest with Self-Efficacy (W1)	.368	.086	< .01
$\boldsymbol{R^2}$	R^2	*SE*	*p*
Exam	.154	.066	.02
Self-Efficacy (PC)	.217	.070	.00
No. Conceptual	.031	.034	.36
No. Procedural	.105	.057	.07

Note. PC = Post-class; W1 = Wave 1.

Commonality Analyses for Analysis A

To examine the unique and shared variance of significant predictors with endogenous variables in Analysis A, I conducted separate commonality analyses on post-class self-efficacy and exam scores. To conduct the commonality analyses, the dependent variable of interest (post-class self-efficacy or exam scores) was regressed on significant predictors from the path model. Note that the regression R^2 values were slightly smaller compared to the R^2 values calculated in the path model since the regressions included fewer predictors (i.e., only significant predictors from the path model) and because they were calculated in SPSS, which does not employ a Full Information Maximum Likelihood framework, as does *Mplus*.

The first commonality analysis examined the unique and shared variance of post-class self-efficacy with significant predictors—self-efficacy (W1), quantity of procedural questions, percentage of unanswered questions, and percentage of closed questions. Results are presented in Table 4.9. Results indicate self-efficacy at Wave 1 accounted for total of 11.8% of the variance in post-class self-efficacy, with 5.9% of that variance shared with other predictors, primarily with the quantity of procedural questions. The quantity of procedural questions accounted for a total of 8.7% of the variance in post-class self-efficacy, with 6.1% of the variance shared with other predictors, primarily Wave 1 self-efficacy. The percentage of unanswered questions accounted for a total of 4.1% of the variance in post-class self-efficacy, with 1.8% of the variance shared with other predictors. Finally, the percentage of closed questions accounted for a total of 5.4% of the variance in post-class self-efficacy, with 1.5% of the variance shared with other predictors. Thus, while the significant predictors in the model explained unique variance in post-class self-efficacy, these predictors also explained overlapping variance. For example, Wave 1 self-efficacy and the quantity of procedural questions shared 3.6% of the variance in post-class self-efficacy, indicating that Wave 1 self-efficacy and the number of procedural questions explained shared variance in post-class self-efficacy that cannot be attributed to one or other. The amount of shared variance among question characteristics was small, indicating that these variables explain unique variance in self-efficacy beliefs.

Table 4.9

Commonality Analysis of Post-Class Self-Efficacy

	1	2	3	4
Variance	Self-Eff (W1)	No. Procedural	% Unanswered	% Closed
Unique	.059	.026	.023	.039
Common	.059	.061	.018	.015
Total	.118	.087	.041	.054
% of R^2	54.4%	40.1%	18.9%	24.9%
C12	.036	.036	—	—
C13	.001	—	.001	—
C14	.010	—	—	.010
C23	—	.010	.010	—
C24	—	.001	—	.001
C34	—	—	-.001	-.001
C123	.008	.008	.008	—
C124	.005	.005	—	.005
C134	-.001	—	-.001	-.001
C234	—	.001	.001	.001
C1234	<.001	<.001	<.001	<.001

Note. Unique = Unique effect. Common = Σ common effects. Total = Unique + Common. % of R^2 = Total/R^2; W1 = Wave 1; Self-Eff = Self-efficacy.

Turning to the unique and shared variance explained in exam scores, I conducted a commonality analysis of exam scores by the two significant predictors from the model in Analysis B: post-class self-efficacy and percentage of closed questions. Results are presented in Table 4.10. Results indicated that post-class self-efficacy accounted for a total of 11.4% of the variance in exam scores, with 8.0% of the variance uniquely explained by this predictor. Percentage of closed questions accounted for a total of 7.5% of the variance with 4.1% of the variance uniquely explained by this predictor. Post-class self-efficacy and percentage of closed questions shared 3.4% of the variance in exam scores.

Table 4.10

Commonality Analysis of Exam Scores

Variance	Self-Efficacy (IC)	%Closed
Unique	.080	.041
Common	.034	.034
Total	.114	.075
%R^2	73.5%	48.4%

Note. Unique = Unique effect. Common = Σ common effects. Total = Unique + Common. % of R^2 = Total/R^2.

Summary of Analysis A. I summarize the results from the well-fitting model and the pattern of significant path coefficients as follows. For this sample and context, Wave 1 interest and self-efficacy had relatively minor effects[2] on question characteristics. Wave 1 self-efficacy negative predicted the number of procedural questions but accounted for a non-significant amount of variance. Interest and self-efficacy were not strongly related to either question syntax or the resolution of questions during class meeting.

Question characteristics predicted students' self-efficacy beliefs at the end of class meetings, after accounting for the effect of their self-efficacy beliefs for the course. The number of procedural questions was negatively associated with post-class efficacy beliefs while the percentage of closed questions was positively associated with post-class efficacy beliefs. The effects of question characteristics on post-class self-efficacy beliefs were mainly unique and explained a modest amount of variance.

With the exception of question syntax, question characteristics did not directly predict exam scores, nor were the indirect effects via post-class self-efficacy significant. Percentage of closed questions did have a direct effect on exam scores, however the indirect effect via self-efficacy beliefs was non-significant. Whatever is captured by generating closed versus open questions (see discussion in Chapter 5) was an important predictor of efficacy beliefs and achievement in this context.

[2] The term *effect* is commonly used in the description of regression results, however it should not be interpreted to imply causality.

Analysis B: Relations Among Goal Orientations, Self-Appraisals, and Exam Scores

Analysis B tested two sets of relations pertaining to motivation, questions, and achievement: (1) the relations between Wave 1 goal orientation variables and self-appraisals of questions; and (2) the relations between self-appraisals and exam scores. Fit statistics indicate that the model has excellent fit (χ^2 = 1.087, *df* = 4, *p* = .890, CFI = 1.00, RMSEA = .000; SRMR = .014). Significant paths[3] are presented in Figure 4.4 and full model results are presented in Table 4.11.

Wave 1 Mastery avoidance goal orientation significantly and positively predicted *Value for Others* (β = .201). Students who reported greater concern about not learning as much as they could tended to appraise their questions as having greater value for others' learning. Mastery approach goal orientation (β = .181, *p* = .076) and performance goal orientation (the sum of performance avoidance and performance approach scores; β = .201, *p* = .058) were positively related to *Value for Others* but were non-significant. The model accounted for a statistically significant 17.5% of the variance in *Value for Others*. In terms of *Value for Self*, the three individual goal orientations were positively related to students' appraisals of the usefulness of their questions for their own learning, however these relations were not statistically significant. Perceived teacher emphasis on mastery goals was not significantly related to either *Value for Self* or *Value for Others*.

[3] I interpret path coefficients as significant at $p < .05$. The variable *Exam 2* was used in four analyses, thereby increasing the potential for Type I error. A more conservative Bonferroni-corrected alpha level is $p < .0125$ ($p = .05/4$).

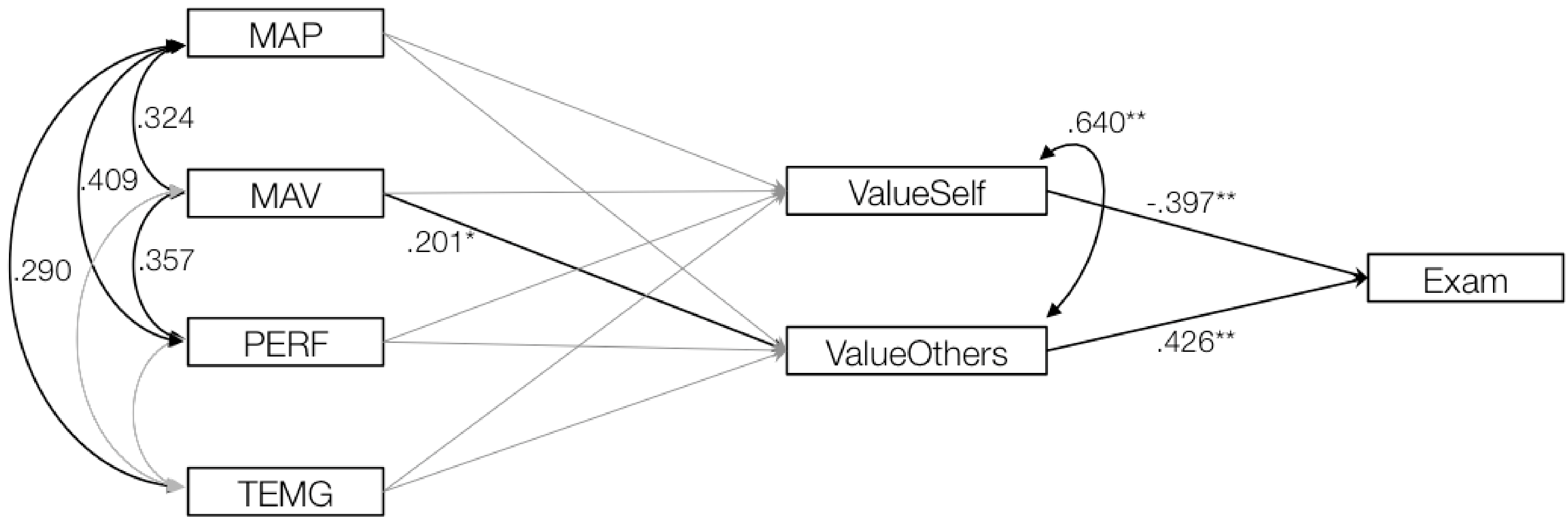

Figure 4.4. Model of Individual and Perceived Teacher Goal Orientations Predicting Self-Appraisals and Exam Scores.

Note. Solid black lines represent significant paths; Solid grey lines represent non-significant paths; Black-dotted lines represent paths that are significant at $p < .10$. MAP = Mastery Approach Goal Orientation; MAV = Mastery Avoidance Goal Orientation; PERF = Performance Goal Orientation; TEMG = Teacher emphasis on mastery goals.
*p<.01; **p<.01

Table 4.11

Estimated Parameters for Model Tested in Analysis B

Parameters	β	*SE*	*p*
Direct Effects			
ValueSelf => Exam	-.397	.131	.00
ValueOthers => Exam	.426	.132	.00
MAP => ValueSelf	.139	.107	.19
MAV => ValueSelf	.156	.102	.13
PERF => ValueSelf	.130	.106	.22
TEMG => ValueSelf	-.056	.098	.57
MAP => ValueOthers	.181	.102	.08
MAV => ValueOthers	.201	.097	.04
PERF => ValueOthers	.191	.101	.06
TEMG => ValueOthers	-.109	.094	.25
Indirect Effects			
MAP => Exam via ValueOthers	.077	.050	.12
MAV => Exam via ValueOthers	.086	.050	.08
PERF => Exam via ValueOthers	.081	.050	.11
Correlated Errors			
ValueSelf with ValueOthers	.640	.058	.00

Note. MAP = Mastery approach goal orientation; MAV = Mastery avoidance Goal Orientation; PERF = Performance avoidance goal orientation; TEMG = Teacher emphasis on mastery goals

(Table 4.11 continued)

	β	*SE*	*p*
Covariance of Exogenous Variables			
MAP with MAV	.324	.089	.00
MAP with PERF	.409	.082	.00
MAP with TEMG	.290	.091	.00
MAV with PERF	.357	.086	.00
MAV with TEMG	.155	.097	.11
PERF with TEMG	.182	.096	.06
$\boldsymbol{R^2}$	R^2	*SE*	*p*
Exam	.108	.063	.09
ValueSelf	.097	.055	.08
ValueOthers	.175	.068	.01

Note. MAP = Mastery approach goal orientation; MAV = Mastery avoidance Goal Orientation; PERF = Performance avoidance goal orientation; TEMG = Teacher emphasis on mastery goals

Value for Self and *Value for Others* each significantly predicted Exam 2 scores but in opposite directions. *Value for Self* had a medium-sized negative effect on exam scores while *Value for Others* had a medium-sized positive effect on exam scores. *Value for Self* and *Value for Others* had opposite effects on exam scores, even as they were positively correlated with each other, as indicated by the bivariate Pearson correlation (r = .68). In the model, the error terms for *Value for Self* and *Value for Others* were strongly and positively correlated, indicating that what the model fails to explain about variance in one self-appraisal is highly correlated with what the model fails to explain about the other self-appraisal. Within an individual, the negative relations of *Value for Self* on exam scores and the positive relation of *Value for Other* on exam scores act to dampen the individual effect of each variable. This is evident in the non-significant bivariate correlations between *Value for Self* and *Value for Others* with exam scores. However, once accounting for the effect of one, the effect of the other become significant and substantial. This suppression effect is examined further in the commonality analysis below. The model accounts for a statistically significant 10.8% of variance in exam scores, indicating a small to moderate effect of self-appraisals on exam scores.

Commonality Analysis for Analysis B

To examine the unique and shared variance of significant predictors with endogenous variables in the model tested in Analysis B, I conducted separate commonality analyses on *Value for Others* and exam scores. As with the commonality analyses in Analysis A, to conduct the commonality analyses on the model in Analysis B, the dependent variable of interest (*Value for Others* or exam scores) was regressed on significant predictor from the path model.

The first commonality analysis examines the unique and shared variance of *Value for Others* with significant predictors—mastery approach, mastery avoidance, and performance goal orientations. Results are presented in Table 4.12. Results indicated a large amount of variance in *Value for Others* shared by individual goal orientations. Mastery approach goal orientation accounted for a total of 9.1% of the variance in *Value for Self* while uniquely explaining only 1.9% of the variance. Mastery avoidance goal orientation accounted for a total of 9.7% of the variance in *Value for Self* while uniquely explaining only 3.2% of the variance. Performance goal orientation accounted for a total of 10.1% of the variance in *Value for Self* while uniquely explaining only 2.7% of the variance. Mastery approach, mastery avoidance, and performance goal orientations respectively accounted for 55.5%, 59.1%, and 61.6% of the total R^2. Interestingly, performance goal orientation explained the greatest amount of total variance, even though it had a smaller (and non-significant) path coefficient compared to mastery avoidance. This indicates that while performance goal orientation explained the greatest amount of total variance, most of that variance overlapped with at least one other goal orientation. A given goal orientation had two to three times more shared variance with at least one other goal orientation than it explained by itself. The amount of shared variance common to all three goal orientations was 3.7% of the variance in *Value for Others*. This shared variance reflects the moderate positive correlations among the goal orientations. Thus, the picture that emerges from the commonality analysis is the *collective* contribution of goal orientations to explaining variance in *Value for Others.*

Table 4.12

Commonality Analysis of Value for Others

	1	2	3
Variance	MAP	MAV	PERF
Unique	.019	.032	.027
Common	.072	.065	.074
Total	.091	.097	.101
% of R^2	55.5%	59.1%	61.6%
C12	.013	.013	—
C13	.022	—	.022
C23	—	.015	.015
C123	.037	.037	.037

Note. Unique = Unique effect. Common = Σ common effects. Total = Unique + Common. % of R2 = Total/R2; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance goal orientation.

Turning to the unique and shared variance of exam scores, I conducted a commonality analysis of exam scores by the two significant predictors from the model tested in Analysis B: *Value for Self* and *Value for Others*. Results are presented in Table 4.13. The results indicated a strong suppressor effect between the two appraisals, as indicated by a negative shared variance (Pedhazur, 1997). *Value for Self* uniquely explains 7.8% of the variance in exam scores and *Value for Others* uniquely explains 8.6% of the variance in exam scores. The large positive correlation between these variables causes the effect of one to suppress the effect of the other. The commonality analysis reveals that 6.9% of the variance in each appraisal is suppressed by the other appraisal, leading to small values for total variance explained when unique and common variances are combined.

Table 4.13

Commonality Analysis of Exam Scores

	ValueSelf	ValueOthers
Total	.009	.017
Common	-.069	-.069
Unique	.078	.086

Note. Unique = Unique effect. Common = Σ common effects; Total = Unique + Common.

Partial correlations. To examine the suppression effect further, I examined the partial correlations among *Value for Self* (controlling for *Value for Others*) and *Value for Others* (controlling for *Value for Self*) as well as correlations with a *ValueGap* score (*ValueGap = Value for Self* minus *Value for Others*). Table 4.14 presents the (1) zero-order Pearson correlations among *Value for Self*, *Value for Others*, and all primary variables in the study; (2) partial Pearson correlations for *Value for Self*, *Value for Others*, and all primary variables in the study; and (3) zero-order Spearman correlations among *ValueGap* and all primary variables in the study (Note: I calculated Spearman rather than Pearson correlation because *ValueGap* was non-normally distributed).

Regarding individual goal orientations, the partial correlations indicated that while *Value for Self* was correlated with individual goal orientations at the zero-order level, these correlations were not significant after controlling for *Value for Others*. That is, after partialling out the variance attributed to *Value for Others*, the variance in *Value for Self* that remained was not significantly correlated with individual goal orientations. The relations among individual goal orientations and *Value for Others* remained significant after controlling for the effect of *Value for Self.* Interestingly, *ValueGap* exhibited a significant and negative Spearman correlation with performance goals but not mastery approach or mastery avoidance goals. Student with higher endorsement of performance goals had smaller gaps between *Value for Self* and *Value for Others.*

Two additional significant correlations are noteworthy. *ValueGap* was negative correlated with GPA and Wave 2 intention to remain in a STEM major. Students with higher GPA before the course tended to have smaller differences between the two appraisals. Similarly, students who tended to have smaller differences between the two appraisals reported stronger intentions to remain in a STEM major.

Table 4.14

Zero-Order and Partial Correlations for Self-Appraisals

Variable	Zero-Order			Partial	
	Value Self[a]	Value Others[a]	Value Gap [b]	Value Self[a]	Value Others[a]
Exam 1	-.11	-.01	-.05	-.15	.10
Exam 2	-.10	.13	-.22*	-.26*	.27**
GPA	.03	.18	-.35**	-.26**	.33**
No. Conceptual	-.08	.06	-.13	-.16	.15
No. Procedural	.10	.11	.01	.03	.06
% Unanswered	.12	.13	-.04	.04	.07
% Closed	-.11	-.01	-.05	-.14	.10
Interest (W1)	.05	.07	-.01	.00	.05
MAP (W1)	.23*	.29**	-.12	.04	.19†
MAV (W1)	.24*	.31**	-.06	.04	.21*
PERF (W1)	.22*	.30**	-.27**	.02	.22*
Self-Efficacy (W1)	-.03	.04	-.06	-.08	.09
StaySTEM (W1)	-.11	.03	-.10	-.17	.14
TEMG (W1)	.03	.01	-.01	.04	-.02
HSExec (W1)	-.17	-.18	-.03	-.07	-.09

Note. [a]Pearson's r; [b]Spearman's rho; W1 = Wave 1; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance avoidance goal orientation; StaySTEM = Intention to remain in STEM major; TEMG = Teacher emphasis on mastery goals; HSExec = Executive help seeking. †$p < .10$; *$p < .05$; **$p < .01$

(Table 4.14 continued)

Variable	Zero-Order			Partial	
	Value Self	Value Others	Value Gap	Value Self	Value Others
HSInstr (W1)	.28**	.29**	-.08	.13	.13
HSThreat (W1)	-.02	-.04	.06	.01	-.04
HSFormal (W1)	-.16	-.14	-.10	-.09	-.04
Interest (PC)	.12	.01	.18	.15	-.10
Perc. Comp (PC)	-.15	-.10	.01	-.11	.00
Self-Efficacy (PC)	.05	.04	.04	.03	.01
Interest (W2)	.14	.16	-.04	.04	.10
MAP (W2)	.24*	.30**	-.16†	.06	.18†
MAV (W2)	.14	.21*	-.17	-.01	.16
PERF (W2)	.27**	.25*	-.06	.14	.09
Self-Efficacy (W2)	.07	.13	-.07	-.03	.12
StaySTEM (W2)	-.07	.04	-.23*	-.13	.12
TEMG (W2)	.09	.09	-.07	.04	.04
Interest (Chg)	.13	.21*	-.13†	-.02	.17
Self-Efficacy (Chg)	.09	.15	-.13	-.02	.13
MAP (Chg)	.09	.06	-.02	.06	.01

Note. [a]Pearson's *r*; [b]Spearman's rho; PC = Post-class; W2 = Wave 2; Chg = Normative change from Wave 1 to Wave 2; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance avoidance goal orientation; StaySTEM = Intention to remain in STEM major; TEMG = Teacher emphasis on mastery goals; HSInstr = Instrumental help seeking; HSThreat = Help seeking threat; HSFormal = Preference for a formal help seeking target. †$p < .10$; *$p < .05$; **$p < .01$

(Table 4.14 continued)

Variable	Zero-Order			Partial	
	Value Self	Value Others	Value Gap	Value Self	Value Others
MAV (Chg)	.16	.21*	-.12	.02	.14
PERF (Chg)	.12	.03	.09	.13	-.07
TEMG (Chg)	.04	.04	-.04	.02	.02

Note. [a]Pearson's *r*; [b]Spearman's rho; PC = Post-class; W2 = Wave 2; Chg = Normative change from Wave 1 to Wave 2; MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance avoidance goal orientation; StaySTEM = Intention to remain in STEM major; TEMG = Teacher emphasis on mastery goals; HSInstr = Instrumental help seeking; HSThreat = Help seeking threat; HSFormal = Preference for a formal help seeking target. †p < .10; *p < .05; **p < .01

Further Analyses on the Relations Among Self-Appraisals, Post-Class Perceptions, and Achievement

I examined to the extent to which post-class self-efficacy and perceived comprehension scores each accurately predicted exam scores. The reason for this analysis was to explore why the *Value Gap* was not related to post-class self-efficacy scores, given that the *Value Gap* may indicate perceptions that questions were basic or rudimentary, as I discuss in Chapter 5. One reason for the non-significant relations may be that post-class perceptions (i.e., post-class self-efficacy and perceived comprehension) were not accurately calibrated with exam scores. To explore this possibility, I regressed Exam 2 scores on post-class self-efficacy and saved the standardized residuals. Residualized scores represented the extent to which students' efficacy beliefs failed to predict exam scores, with negative residuals representing overly confident self-efficacy beliefs (relative to actual performance) and positive residuals representing under-confident self-efficacy beliefs. I then examined the Spearman correlations between *Value Gap* and residualized scores. The correlations, presented in Table 4.15, indicated that high *Value Gap* scores were associated with overly confident post-class efficacy belief and low *Value Gap* scores were associated with under-confident efficacy beliefs.

I repeated the analyses described above for perceived comprehension: regressing perceived comprehension on Exam 2 scores, saving the standardized residuals, and examining the Spearman correlations with *Value Gap*, presented in Table 4.15. As with self-efficacy beliefs, high *Value Gap* scores were associated with overly confident evaluations of comprehension and low *Value Gap* scores were associated with under-confident evaluations.

Table 4.15

Spearman Correlations Among Post-Class Residualized Scores and the Value Gap

		1	2	3
1	Value Gap	—		
2	Self-Efficacy Residuals	-.252*	—	
3	Perceived Comprehension Residuals	-.289**	.956**	—

Note. *Value Gap* = (*Value for Self* minus *Value for Others*). Residualized gain scores are standardized residuals from regressions in which Exam 2 scores were regressed on each post-class perception. Negative and positive residuals represent over-estimations and under-estimations, respectively, relative to performance on Exam 2.
*p < .05; **p < .01.

Alternative Analysis with Winsorized Scores

As I screened the variables used in Analysis B for potential outliers, outlier diagnostic statistics indicated the presence of several individual cases (8-10 students) with extreme scores. Examination of these cases revealed that each case had an abnormally high or low score on a single variable, such as a very low test score (e.g., 16 out of 100) or endorsement of a motivational construct far below the mean for the sample (a mastery avoidance score of 1.00). To examine the effect of these potential outliers, I tested the model using Winsorized scores, in which scores were forced to fall within 2 *SD* of the original mean for a given variable.

As with the analyses presented above, results with the Winsorized scores indicated a good fitting model (χ^2 = 1.864, *df* = 4, *p* = .761, CFI = 1.00, RMSEA = .000; SRMR = .020). Compared to the non-trimmed scores, results from the model with Winsorized scores estimated slightly weaker but significant paths from *Value for Self* and *Value for Others* to Exam 2 scores—β = -.341 vs. -.397 and .380 vs. .426, respectively—and the model explained a non-significant 8.3% (vs. a non-significant 10.8%) of the variance in exam scores. The two models demonstrated some differences with regard to the relations among individual goal orientations and *Value for Others*. Recall that in the model with non-trimmed scores, *Value for Others* was significantly predicted by mastery avoidance scores (β = .201, *p* = .04) while mastery approach goals (β = .201, *p* = .08) and performance goal (β = .191, *p* = .06) were significant at *p* < .10. The opposite pattern occurred with the Winsorized scores: *Value for Others* was significantly predicted by mastery approach goals (β = .209, *p* = .04) and performance goals (β = .215, *p* = .03) while mastery avoidance (β = .178, *p* = .07) was significant at *p* < .10. As the commonality analyses above revealed, the majority of the variance in *Value for Others* was shared by individual goal orientations. While the model with the Winsorized scores produced slightly different results regarding the unique variances explained by individual goal orientations, the overall pattern of results was similar.

Summary of Analysis B. Based on the good fit of the model tested in Analysis B and the pattern of significant path coefficients, the results can be summarized as follows. For this sample and context, goal orientations are positively related with how student appraise their questions, particularly the extent to which they view their question as having value for others. Because the individual goal orientations were moderately and positively correlated with each other, the goal orientations tended to collectively explain variance in *Value for Others*, with the greatest unique effect attributable to variance in mastery avoidance goal orientation.

The model revealed a large suppression effect of one appraisal on the other. *Value for Self* negatively predicted exam scores. Assigning more *individual* importance to knowledge deficits was associated with lower exam scores. *Value for Others* positively predicted exam scores. Assigning more importance for others to knowledge deficits was associated with higher exam scores.

Multiple Regression of Exam 2 Scores on Question Characteristics

To understand the unique and collective relationship between question characteristics and exam scores, I regressed Exam 2 scores on the number of conceptual and procedural questions, the percentage of unanswered questions, the percentage of closed questions, and *Value for Self* and *Value for Others.* Results are presented in Table

4.16. The regression accounted for significant[4] 20.7% of the variance in exam scores. Exam scores were significantly and positively predicted by the percentage of closed questions (β = .201) and *Value for Others* (β = .415) and significantly and negatively predicted by *Value for Self* (β = -.383). The number of conceptual and procedural questions and the percentage of unanswered questions were not significantly related to exam scores after accounting for other question characteristics.

Table 4.16

Multiple Regression of Exam 2 Scores on Question Characteristics

Variable	β	*p*
No. Conceptual	-.144	.150
No. Procedural	.053	.591
% Unanswered	-.061	.526
% Closed	.270	.007
ValueSelf	-.383	.007
ValueOthers	.415	.003

Note. $R^2 = .207$; $p = .001$.

[4] I interpret path coefficients as significant at $p < .05$. The variable *Exam 2* was used in four analyses, thereby increasing the potential for Type I error. A more conservative Bonferroni-corrected alpha level is $p < .0125$ ($p = .05/4$).

Analysis C: Question Characteristics and Normative Changes in Motivation Constructs

The results from a series of independent samples *t* tests examining mean differences in question characteristics between students who increased versus decreased on motivational constructs from Wave 1 to Wave 2 are presented in Table 4.17.

Results indicated that growth in three motivational constructs—mastery avoidance goal orientation, performance goal orientation, and interest—were associated with statistically significant[5] higher means for at least one question characteristic. Growth in mastery avoidance goal orientation and growth in interest were each associated with higher *Value for Other* appraisals. Thus, students who become more mastery avoidant from Wave 1 to Wave 2 tended on average to appraise their questions as more valuable for others compared to students who decreased in mastery avoidance. Similarly, students who increased in interest from Wave 1 to Wave 2 reported higher *Value for Others*. The effect sizes of the mean differences were small to medium sized (d = .42 and d = .43, respectively).

Students who grew in performance goal orientation generated on a greater mean number of procedural questions compared to students who reported a decrease in performance goal orientation. The effect size for this mean difference was large (d = .77).

[5] For this set of *t* tests I interpret results as statistically significant at $p < .05$. Given that 6 *t* tests are repeated with each variable, a more conservative Bonferroni-corrected alpha level is $p < .008$ (p = .05/6).

Table 4.17

T-tests on Mean Differences in Question Characteristics between Students Who Increased Versus Decreased on Motivational Constructs

	M (SD)									
Variable	Decr.		Incr.		n_{Decr}	n_{Incr}	*df*	*t*	*p*	*d*
MAP										
No. Conceptual	5.03	(3.37)	5.67	(4.47)	38	54	90.00	-0.75	.46	0.16
No. Procedural	3.74	(2.53)	4.48	(2.53)	38	54	90.00	-1.39	.17	0.29
% Unanswered	0.67	(0.26)	0.74	(0.23)	38	54	90.00	-1.37	.17	0.29
% Closed	0.33	(0.27)	0.34	(0.22)	38	54	90.00	-0.27	.79	0.06
ValueSelf	3.83	(0.67)	3.94	(0.64)	38	54	90.00	-0.81	.42	0.17
ValueOthers	3.37	(0.74)	3.46	(0.65)	38	54	90.00	-0.59	.56	0.12
MAV										
No. Conceptual	5.34	(4.20)	5.26	(3.97)	41	50	89.00	0.10	.93	-0.02
No. Procedural	4.20	(2.29)	4.16	(2.82)	41	50	89.00	0.06	.95	-0.01
% Unanswered	0.72	(0.21)	0.71	(0.27)	41	50	89.00	0.23	.82	-0.05

Note. MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation.

(Table 4.17 continued)

	M (SD)									
Variable	Decr.		Incr.		n_{Decr}	n_{Incr}	*df*	*t*	*p*	*d*
% Closed	0.33	(0.20)	0.36	(0.28)	41	50	89.00	-0.62	.54	0.13
ValueSelf	3.76	(0.76)	3.97	(0.56)	41	50	89.00	-1.51	.14	0.32
ValueOthers	3.23	(0.74)	3.52	(0.60)	41	50	89.00	-2.01	.05*	0.42
PERF										
No. Conceptual	5.03	(4.33)	5.51	(3.87)	37	57	92.00	-0.56	.58	0.12
No. Procedural	3.08	(2.03)	4.88	(2.63)	37	57	92.00	-3.53	< .01	0.77
% Unanswered	0.66	(0.28)	0.75	(0.21)	37	57	92.00	-1.79	.08	0.37
% Closed	0.37	(0.29)	0.32	(0.20)	37	57	58.35	1.03	.31	-0.23
ValueSelf	3.79	(0.74)	3.95	(0.58)	37	57	64.38	-1.08	.29	0.23
ValueOthers	3.39	(0.71)	3.43	(0.67)	37	57	92.00	-0.29	.77	0.06
Personal Interest										
No. Conceptual	4.93	(3.58)	5.63	(4.38)	42	52	92.00	-0.84	.40	0.18
No. Procedural	3.90	(2.43)	4.38	(2.66)	42	52	92.00	-0.90	.37	0.19

Note. PERF = Performance avoidance goal orientation.

(Table 4.17 continued)

Variable	*M (SD)* Decr.		Incr.		n_{Decr}	n_{Incr}	*df*	*t*	*p*	*d*
% Unanswered	0.70	(0.25)	0.73	(0.24)	42	52	92.00	-0.59	.55	0.12
% Closed	0.30	(0.23)	0.37	(0.25)	42	52	92.00	-1.44	.15	0.30
ValueSelf	3.79	(0.75)	3.96	(0.55)	42	52	72.82	-1.21	.23	0.26
ValueOthers	3.25	(0.71)	3.54	(0.64)	42	52	92.00	-2.10	.04*	0.43
Self-Efficacy										
No. Conceptual	5.50	(3.72)	5.26	(4.27)	36	57	91.00	0.27	.79	-0.06
No. Procedural	4.44	(2.78)	3.98	(2.43)	36	57	91.00	0.84	.40	-0.18
% Unanswered	0.71	(0.23)	0.72	(0.25)	36	57	91.00	-0.29	.77	0.06
% Closed	0.29	(0.21)	0.36	(0.25)	36	57	91.00	-1.33	.19	0.29
ValueSelf	3.81	(0.66)	3.93	(0.65)	36	57	91.00	-0.84	.40	0.18
ValueOthers	3.27	(0.67)	3.49	(0.69)	36	57	91.00	-1.48	.14	0.32
TEMG										
No. Conceptual	4.84	(3.29)	5.73	(4.57)	43	51	92.00	-1.06	.29	0.23

Note. = Teacher emphasis on mastery goals.

(Table 4.17 continued)

	M (SD)									
Variable	Decr.		Incr.		n_{Decr}	n_{Incr}	*df*	*t*	*p*	*d*
No. Procedural	3.98	(2.85)	4.33	(2.29)	43	51	92.00	-0.67	.50	0.14
% Unanswered	0.71	(0.26)	0.72	(0.24)	43	51	92.00	-0.07	.95	0.01
% Closed	0.36	(0.29)	0.32	(0.19)	43	51	70.03	0.83	.41	-0.18
ValueSelf	3.86	(0.64)	3.91	(0.67)	43	51	92.00	-0.40	.69	0.08
ValueOthers	3.38	(0.72)	3.44	(0.65)	43	51	92.00	-0.40	.69	0.08

Relations Among Question Characteristics, Motivation, and Exam Scores: Profile-Centered Analyses

To explore the possibility that group statistics obscured subsets of participants with similar patterns of scores, I conducted three latent class analyses on Wave 1 motivation variables only (LCA1); question characteristics only (LCA2); and Wave 1 motivation variables with question characteristics (LCA3). Results from LCA1 and LCA2 indicated insufficient evidence of more than one class; the results are described briefly below. Results from LCA3 indicated sufficient evidence of a two-class solution.

Latent Class Analysis of Wave 1 Motivation Constructs

I conducted a latent class analysis on the mean scores for the following Wave 1 motivation variables: interest, self-efficacy, mastery approach goal orientation, mastery avoidance goal orientation, performance goal orientation and teacher emphasis on mastery goals. Fit indices are presented in Table 4.18. Collectively, model indices provided insufficient evidence of more than one latent class. Although the AIC, BIC, and ABIC were lower for a 2- vs. 1-class solution, the Entropy was low (<.80) and the Lo Mendel Rubin test for a significantly improved fit for the 2-class solution was non-significant (p = .276). Furthermore, the two classes had similar relations among variables, with one class having slightly higher means across variables compared to the other class. Absent compelling evidence that homogenous latent sub-groups existed in the sample, person-centered analyses for Wave 1 motivation variables alone were not pursued.

Table 4.18

Fit Statistics for Latent Class Analysis of Wave 1 Motivation Variables

No. classes	*AIC*	*BIC*	*ABIC*	Entropy	*LMR*
1	1519.827	1551.327	1513.423	—	—
2	1455.282	1505.156	1445.142	0.765	n.s.

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; ABIC = Adjusted Bayesian Information Criterion; LMR = Lo Mendel Rubin Likelihood Ratio Test for 1 vs. 2 classes.

Latent Class Analysis of Question Characteristics

I conducted a latent class analysis on the mean scores for the following question characteristics: number of conceptual questions, number of procedural questions, percentage of unanswered questions, percentage of closed questions, value for self, and value for others. Fit indices are presented in Table 4.19. Collectively, model indices provided insufficient evidence of more than one latent class. Although the AIC, BIC, and ABIC were lower for a 2- vs. 1-class solution, the Entropy was low (<.80) and the Lo Mendel Rubin test for a significantly improved fit for the 2-class solution was non-significant (p = .074). Furthermore, the two classes had similar relations among variables, with one class having slightly higher means across variables compared to the other class. Absent compelling evidence that homogenous latent sub-groups existed in the sample, person-centered analyses for question characteristics alone were not pursued.

Table 4.19

Fit Statistics for Latent Class Analysis of Question Characteristics

No. classes	*AIC*	*BIC*	*ABIC*	Entropy	*LMR*
1	1503.152	1534.769	1496.863	—	—
2	1465.949	1516.009	1455.991	0.741	n.s.

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; ABIC = Adjusted Bayesian Information Criterion; LMR = Lo Mendel Rubin Likelihood Ratio Test for 1 vs. 2 classes.

Latent Class Analysis of Wave 1 Motivation Constructs and Question Characteristics

Determining the Number of Latent Classes

I conducted a latent class analysis on the mean scores for the following Wave 1 motivation variables and question characteristics: interest; self-efficacy; mastery approach, mastery avoidance, performance goal orientations; teacher emphasis on mastery goals; number of conceptual questions; number of procedural questions; percentage of unanswered questions; percentage of closed questions; value for self; and value for others. Fit indices and class probabilities for a 2-class solution are presented in Table 4.20 and Table 4.21, respectively. Together, model indices provided evidence for a 2-class solution. AIC, BIC, and ABIC were lower for a 2- vs. 1-class solution, entropy was adequate (>.80) and the Lo Mendel Rubin test of a significantly improved fit for the 2-class solution was significant (p = .049). While the AIC, BIC, and ABIC were lower for a 3- vs. 2-class solution and the entropy was slightly higher, two pieces of evidence suggested the 2-class solutions was preferable. First, the MLR was non-significant, indicating the improvement in fit was not large enough to warrant the less parsimonious solution. Second, the 3-class solution included one class with a small number of participants (*n* = 11). Therefore, a 2-class solution was adopted. Descriptive statistics for motivational and question characteristic variables by latent class are presented and illustrated in Figure 4.5.

Table 4.20

Fit Indices for Competing Models in a LCA of Wave 1 Motivation Variables and Question Characteristics

No. classes	*AIC*	*BIC*	*ABIC*	Entropy	*LMR*	Smallest class
1	3022.98	3086.213	3010.401	1	—	—
2	**2924.907**	**3022.392**	**2905.515**	**.880**	**.049**	**25%**
3	2905.024	3036.761	2878.82	.879	.526	13%

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; ABIC = Adjusted Bayesian Information Criterion; LMR = Lo Mendel Rubin Likelihood Ratio Test for 2- vs. 1- and 3- vs. 2-classes.

Table 4.21

Class probabilities

Class		Class 1	Class 2
	1	.970	.030
	2	.033	.967

	Int (W1)	SelfEff (W1)	MAP (W1)	MAV (W1)	PERF (W1)	TEMG (W1)	No. Conc.	No. Proc.	% Unans	% Closed	Value Self	Value Other
Class 1 (25%)	2.56	3.90	3.77	2.63	3.26	3.54	5.13	4.16	0.31	0.65	3.39	2.81
Class 2 (75%)	3.23	4.31	4.70	3.99	4.54	3.90	5.56	3.97	0.33	0.75	4.08	3.65

Figure 4.5. Mean Scores for Wave 1 Motivation and Question Characteristics for Two Latent Classes

Note. W1 = Wave 1; Int = Interest; SelfEff = Self-efficacy; MAP = Mastery approach goal orientation; MAV = Mastery Avoid Goal Orientation; PERF = Performance Goal Orientation; TEMG = Teacher emphasis on mastery goals; Conc. = Conceptual questions; Proc = Procedural questions; % Unans. = Questions that remained unanswered.

Description of Latent Classes

Results from a latent class analysis on Wave 1 motivation variables and question characteristics identified two latent classes, with class 1 constituting 25% (n = 26) and class 2 constituting 75% (n = 77) of the sample. An examination of the pattern of mean scores between the two classes suggested few differences in question characteristics. Students in both latent classes generated a similar number of procedural and conceptual questions and similar percentages of closed questions. The two classes differed more substantially in the percentage of questions that remained unanswered by the end of class meetings, with the smaller latent class having fewer of their question unanswered (65% of questions for this class) compared to the larger class (75% of questions). The greatest differences in question characteristics between the two classes were the self-appraisals. Compared to the larger class, the smaller class assigned lower *Value for Self* and *Value for Others*. In both classes, *Value for Self* was higher than *Value for Others*.

Regarding Wave 1 motivation variables, the smaller class had lower scores on all motivation variables compared to the larger class. The smaller class had mean scores for motivation constructs that indicated *somewhat agreement* with the items while the larger group had mean scores that indicated *agreement* to *strong agreement* with the scale items. Overall, the results suggested that students in the two latent classes differed in the *extent* to which they endorsed motivational constructs (i.e., class 1 = med; class 2 = high) but had similar *patterns* of motivation in terms of the relative relations among motivational constructs. For example, both classes showed similar patterns of individual goal orientations in which mean scores for mastery approach and performance were similar to each other and mean scores for mastery avoidance were slightly lower.

The results suggested that students who had medium scores on motivational variable had similar question characteristics as students who had high scores on motivational variables. Students who had higher scores for motivational variables also had higher self-appraisals compared to students with lower scores on motivational variables.

t *test on Exam 2 scores by Latent Class Membership*

An independent samples *t*-test of Exam 2 scores by class membership indicated a non-significant difference in mean exam scores, $t(97) = -1.054$, $p = .295$. Thus, the two profiles of Wave 1 motivational variables and questions characteristics did not significantly differ in their Exam 2 performance.

In summary, the latent class analyses suggested that while there was some evidence that two latent classes existed in the sample with regard to their Wave 1 motivations and question characteristics, classes were characterized by medium versus high endorsement of motivation constructs and medium versus high appraisals. However, classes appeared to have similar relations among motivation variables and question characteristics and mean Exam 2 scores for the two classes were not significantly different.

Comparison between Question Generation and Question Asking

Descriptive statistics for motivation constructs and question characteristics for askers and non-askers are presented in Table 4.22. With regard to Wave 1 motivation variables, askers and non-askers had similar mean scores for many constructs, namely interest, mastery avoidance and performance goal orientation, perceived teacher emphasis on mastery goals, and intention to remain in STEM. Askers and non-askers showed more substantial differences in mean scores for mastery approach goal orientation and self-efficacy, with askers reporting lower levels on both variables compared to askers.

Askers and non-askers demonstrated similar mean scores for help seeking variables. The six students who asked questions did not differ dramatically from non-askers in their endorsement of instrumental or executive help seeking, their perceived threat of help seeking, or their preference for formal vs. informal targets.

Table 4.22

Descriptive Statistics Comparing Askers Versus Non-Askers

	Askers			Non-askers		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*
Wave 1 Motivation						
HSExec	6	1.89	1.22	95	1.68	0.63
HSInstr	6	3.93	1.14	96	4.16	0.84
HSThreat	6	1.50	0.55	96	1.40	0.62
HSFormal	6	3.44	0.91	96	3.53	0.93
Interest	6	2.67	0.98	96	3.07	1.00
MAP	6	3.94	0.33	96	4.48	0.63
MAV	6	3.83	1.09	96	3.61	1.18
PERF	6	4.19	0.92	94	4.17	0.94
Self-Efficacy	6	3.47	0.86	96	4.24	0.69
StaySTEM	6	4.61	0.80	96	4.50	0.75
TEMG	6	3.33	1.09	96	3.83	0.61
Question characteristics						
No. Conceptual	6	6.67	2.88	97	5.37	4.08
No. Procedural	6	3.67	2.50	97	4.04	2.54

Note. MAP = Mastery approach goal orientation; MAV = Mastery avoidance goal orientation; PERF = Performance avoidance goal orientation; TEMG = Teacher emphasis on mastery goals; StaySTEM = Intention to remain in a STEM major; HSExec = Executive help seeking; HSInstr = Instrumental help seeking; HSThreat = Help seeking threat; HSFormal = Preference for help seeking from a formal target; % Closed = Percentage of questions with closed syntax; ValueSelf = Value for Self; ValueOth = Value for Others.

(Table 4.22 continued)

	Askers			Non-askers		
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*
% Unanswered	6	0.55	0.21	97	0.73	0.24
% Closed	6	0.18	0.16	97	0.33	0.25
ValueSelf	6	4.13	0.53	97	3.88	0.65
ValueOthers	6	3.39	0.68	97	3.43	0.70

With regard to question characteristics, askers and non-askers differed in mean scores for several question characteristics. Compared to non-askers, askers generated more conceptual questions, fewer unresolved questions, fewer questions with closed syntax, and assigned higher *Value for Self* to their questions.

In sum, askers and non-askers demonstrated some differences in their motivations and question characteristics. These differences were moderate and, given the small number of askers, cannot be evaluated as to whether the differences are statistically significant. Furthermore, any apparent differences may be due to sampling error.

CHAPTER 5

DISCUSSION

In this study I examined the questions that undergraduate students generated during chemistry lectures and the relations among students' questions, motivations, and achievement. Question-generation has been proposed as a critical mechanism in learning and cognitive development (Bruner, 1960; Chouinard, 2007; Papert, 1980; Piaget, 1952; Sternberg, 1987; Wittrock, 1992) and an important indicator of student interest, metacognition, and understanding (Baram-Tsabari & Yarden, 2007; Chin & Osborne, 2006; Palincsar & Brown, 1984; Renninger, 2000). Yet, relatively little is known about students' questions during common instructional contexts, such as listening to a lecture, or relations among the act of generating questions, motivation, and achievement. To address these gaps in the literature, I had four primary aims for the study. First, I sought to examine the characteristics of student-generated questions and their inter-relations, including the number and nature of students' questions and students' appraisals of their questions during chemistry lectures. The second aim of the study was to examine the relations among student-generated questions and students' motivational self-beliefs and perceptions of the classroom environment. The third aim of the study was to examine the relations between student-generated questions and achievement on a course exam. A final aim of the study was to compare behaviors and motivations related to *generating* questions with those related to *asking* questions during class meetings.

Accordingly, I investigated the following research questions. (RQ1) What are the relations among the characteristics and self-appraisals of student-generated questions during chemistry lectures? (RQ2) What are the relations among motivational self-beliefs, perceived teacher emphasis on mastery goals, and the quantity, qualities, and self-appraisals of student-generated questions? Research Question 3 (RQ3) investigated the relations among exam scores, the quantity and qualities of student-generated questions, and motivational perceptions; RQ3 was posed in two parts: (RQ3a) Can exam scores be predicted by the quantity and qualities of student-generated questions? and (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and exam scores? (RQ4) Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, help seeking beliefs, achievement, and the quantity, qualities, and self-appraisals of the questions they generate?

In the following sections, I address each of the research questions in light of the context of the learning task, domain and course. It is important to emphasize that the methods used in this study—having students record questions in logs during class meetings—constituted a potentially important influence on questioning behavior itself. Asking students to be cognizant of their questions, providing time for students to identify and articulate questions, and reinforcing the value of students' questions by integrating them into later instruction undoubted influenced questioning behavior. Therefore, I do not presume to have examined student-generated questioning in a naturalistic setting, strictly defined. Although the methods employed in this study intervened on the object of study—students' questions—the methods allowed questions to be recorded and observed with relative privacy in the context of an authentic classroom and therefore afforded a novel glimpse at students' questions that typically would not be publically shared.

Research Question 1: What are the relations among the characteristics and self-appraisals of student-generated questions?

Question characteristics were, in general, independent of each other. The number of conceptual and procedural questions, the extent to which questions were perceived to be resolved during class meetings, and students' appraisals of their questions captured unique and independent aspects of student-generated questions in this sample and context. There were two exceptions to the general trend of question characteristics being independent of each others: 1) the number of conceptual questions was negatively related to the percentage of closed questions and 2) *Value for Self* and *Value for Others* exhibited a strong positive correlation. In the following section, I discuss what each question characteristic and its relations may indicate about student cognition and knowledge in the context of the study.

The Number of Conceptual and Procedural Questions

Across 8 lectures, students generated on average 1-2 questions per 50-minute class meeting. This count excluded questions that addressed exclusively logistical concerns (e.g., when review sessions were scheduled). Not surprisingly, the type of questions that students generated reflected the nature of the content that was taught. Students generated a large number of both conceptual and procedural questions, reflecting the relatively balanced attention paid to both conceptual and procedural knowledge during class meetings.

What did the number of student-generated questions indicate about student cognition and knowledge deficits in this context? At the most basic level, the number questions represented the number of knowledge deficits of which a student was aware and willing and able to record. Students recorded questions within the time constraints of the class meetings (a couple of minutes each class period) and in the context of competing goals (e.g., recording a question vs. taking notes or whispering with a friend seated nearby) and therefore students may not have recorded all of their questions. Nevertheless, the number of conceptual and procedural questions provided a glimpse, albeit a rudimentary one, into students' metacognition and knowledge during lectures.

Questions can indicate metacognitive monitoring of understanding (Palincsar & Brown, 1984), and therefore in some contexts, the number of questions indicate the extent to which students monitor their understanding. In this study, however, students were prompted to monitor for gaps in their understanding in several ways, including verbal reminders, allocated class time, and visual cues, such as the question log on their desk and the behaviors of classmates completing the logs. Therefore, it is difficult to draw conclusions about the relations among the number of questions and metacognitive monitoring in the context of this study.

Question logs may have prompted students to generate self-explanations. The act of generating explanations while learning has been associated with increased comprehension and problem solving (Chi et al., 1989), as have other strategies associated with generative learning (Wittrock, 1974, 1992), such as summarizing, elaborating, clarifying, analogizing, and questioning (Brown & Palincsar, 1984; King, 1994; Wittrock & Alesandrini, 1990). While self-explanation may facilitate multiple cognitive processes that support learning, one key mechanism is that explanations help learners uncover gaps in knowledge (VanLehn & Jones, 1993). Findings from intervention studies suggest that the self-explanation effect can be triggered by prompting students to self-explain while learning (e.g., Chi et al., 1994). Thus, prompting students to identify questions during lectures may be akin to prompting students to engage in self-explanation or use other generative learning strategies. In addition to metacognitive monitoring, student-generated questions can reveal information about other aspects of cognition, such as whether the student has engaged in shallow or deep processing (Graesser & Person, 1994; Scardamalia & Bereiter, 1992). Graesser and colleagues (Graesser, Ozuru, & Sullins, 2010; Graesser & Person, 1994) have demonstrated that some question types are positively correlated with higher-order cognition as defined by Bloom's taxonomy (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Krathwohl, 2002).

Following Graesser and colleagues' categorization of question types (Graesser, McMahen, & Johnson, 1994; Graesser, Person & Huber, 1992), the majority of questions that students generated in this sample and context suggested higher-order thinking. Several frequently-occurring types of procedural and conceptual questions suggested deep rather than shallow processing, such as questions that sought information about: rules or conditions for applying procedures; the purpose, value, or rationale behind procedural or conceptual knowledge; how one could know when knowledge was relevant (Procedural enablement, Enablement); causal explanation or justifications; or the relations among concepts.

It is important to emphasize that not all question types that are associated with higher-order processing necessarily reflect complex or sophisticated inquiries or even deep processing (Graesser, Ozuru, & Sullins, 2009). Furthermore, it is important to keep in mind that I used the term *conceptual questions* to refer to inquiries that addressed non-procedural concepts in any form, including factual information. Thus, conceptual questions did not necessarily address deep conceptual understanding, as the term could be interpreted. Nevertheless, in general, many frequent question types appeared to indicate attempts by students to integrate, evaluate and apply knowledge. For this sample and context, such attempts constituted a majority of students' questions.

The tendency for students in this sample and context to generate higher-order questions is inconsistent with findings by Graesser and Person (1994), who found that undergraduate students tended to ask mostly lower-order questions while working with a tutor learning introduction to psychology material. The difference between these findings likely reflects the influence of the nature of the learning task and academic domain on question generation. In this study, students generated questions while learning complicated, objective material with a strong emphasis on procedural problem solving that required learning and applying procedural and conceptual knowledge. Given the context, it is not surprising that students asked many questions that suggested higher-order processing when the instructional context emphasized the application of concepts and procedures. The prominence of questions that suggested higher-order thinking highlighted the importance of understanding and interpreting questioning behavior in the context in which it occurred (Yang, 2006).

Resolving Questions

Most students' questions were not resolved during class meetings. Students reported that approximately 25% of their questions had been resolved in the period of time between when they generated them in class and the end of the class meeting. What does variation between students in the percentage of unresolved questions indicate about students' cognitions and knowledge deficits in this context?

The act of resolving a question during lectures (without asking it) may have indicated adaptive cognitive and metacognitive processes, such as when an individual searched for and successfully integrated information or prior knowledge to resolve a knowledge deficit. For example, a student may have resolved her question during the lecture by thinking about how the question related to the material she read the night before, searching for an unknown term in her textbook, or considering how new information presented in the lecture related to her knowledge gap. Therefore, resolving questions may have indicated effortful and adaptive cognitive processes and therefore be associated with perceptions of self-efficacy and comprehension (a point I discuss in further detail below).

In addition to indicating how students process information, the act of resolving a question may have indicated *the nature of the deficit itself*, such as how substantial or central a deficit was to understanding the instructed material. Deeper or more substantial knowledge deficits may have been more likely to have remained unresolved during a class meeting compared to surface or minor gaps. At the same time, question that were central to the lecture content—regardless of whether they represented large or small deficits—may have been more likely to have been covered later in the lecture and thus resolved during the class meeting compared to question that were peripheral to the instructed material.

Closed vs. Open Questions

Approximately one-third of students' questions were phrased using closed syntax. Closed questions articulated a knowledge deficit in such a way that it could be resolved by a single-word answer (e.g., yes or no; this or that). What does the variation in the percentage of closed (vs. open) questions indicate about students' metacognition and knowledge deficits?

Questions with closed syntax may have indicated effortful cognitive processes, such as attempts to refine a knowledge deficit, construct plausible answers to one's question, or test a tentative understanding. Questions that articulated knowledge deficits such that they could be answered with a single word may have reflected the effortful process of refining what it was *exactly* that a student did not understand. In doing so, a student may have articulated the precise nature of the knowledge deficit, exposing the crux of the lack of understanding. Thus, students who generated a greater percentage of closed questions may have exerted more effort in precisely articulating what was not understood. In addition, questions that articulated knowledge deficits with closed syntax may have resulted from students testing or confirming a tentative or emerging understanding. Testing a tentative or emerging understanding may have indicated students' tendency to attempt to construct a plausible answer to their questions, which could be confirmed or disconfirmed using a closed question.

In addition to indicating cognitive processes, closed questions may have indicated the extent of the knowledge deficit itself. Closed questions may have represented smaller, less substantial deficits while questions with open syntax may have represented more extensive gaps in knowledge. That is, closed questions may have indicated a tentative understanding or a degree of uncertainty while open questions may have been more likely to represent missing information or complete gaps in knowledge. Thus, a greater percentage of closed questions may have indicated smaller or less substantial knowledge deficits.

Consider, for example, two questions involving the simple definition of a dilute solution. One student generated an open question, "What is a dilute solution?" A second student generated a closed question, "Does a dilute solution have more or less solute?" While both questions address essentially the same content knowledge, the closed questions may have indicated more effortful processing or a smaller knowledge deficit than is indicated by the open question. The student who generated the closed question may have worked to refine the exact nature of the lack of understanding—the relative amount of solute (to solvent, presumably) in dilute solutions. In addition, the closed question may arguably constitute a slightly smaller knowledge deficit than the open question. By contrast, the student who generated the open question may have failed to articulate the precise nature of her deficit (i.e., the part of the definition that was unclear) or may have had a greater deficit (e.g., not understanding what either the terms *dilute* or *solution* means).

The percentage of closed questions had a significant small and negative relation with the number of conceptual questions. Therefore, conceptual questions were more often articulated using open rather than closed syntax. In this context, frequent conceptual question types, such as definitions, justifications, causal explanations, or concept completion questions tended to be phrased using open syntax. In the context of an introductory chemistry course, the negative relationship between the number of conceptual questions and percentage of closed syntax may have indicated that conceptual knowledge deficits were more likely to require extended responses—such as a causal explanation, justification, or definition, and less likely to require a clarification or confirmation. In addition, the inverse relation may have reflected cultural-linguistic norms for making certain types of inquiries. For example, beginning a question about causality with *Why* and questions about the meaning of terms with *What*.

Self-Appraisals

Students tended to generate question that they appraised as being important for their individual learning and the learning of others. Students consistently perceived greater *Value for Self* in their questions than *Value for Others,* and the two appraisals were strongly and positively correlated (points that are discussed in greater detail below). Students' appraisals of their questions were not significantly correlated with other question characteristics. Thus, students' appraisals were independent of whether they generated many or few conceptual or procedural questions, whether they answered many or few of their questions, and whether they articulated more or less of their questions using closed syntax. Note that students' appraisals were independent of other question characteristics *when these variables were aggregated across time and questions*, but this may have ignored dependent relationships at the level of specific questions (see the Limitations section below for further discussion of this point). What do students' appraisals tell us about students' metacognition and knowledge in this context?

Value for Self. Individuals evaluate their questions in light of their learning goals (Flammer, 1981). Presumably, high *Value for Self* appraisals indicated students' tendency to identify knowledge deficits that they viewed as important relative to their learning goals. Low appraisals, on the other hand, presumably indicated the tendency for students to identify knowledge deficits that they perceived as being peripheral or unimportant to learning goals.

Students' mean *Value for Self* appraisal across questions was 3.9 (out of 5), indicating that student tended to generate questions that they viewed as important and relevant knowledge deficits in light of their learning goals. The relatively high mean for *Value for Self* may have been the product of internal editing on the part of students. For example, students may have recorded only those questions that they deemed most valuable for their own learning and dismissed questions they perceived as unimportant or irrelevant. Importantly, the high mean score for *Value for Self* suggested that prompting students to generate questions during class and providing time and incentive to do so did not result in questions that students perceived to have little value for their learning.

Value for Others. Appraising a question's *Value for Others* presumably involved evaluating (1) the extent that peers were likely to have a similar knowledge deficit (include deficits that others are unaware of) and (2) the extent that resolving the knowledge deficit would be useful in light of peers' goals. Presumably, high mean *Value for Others* scores indicated the tendency to identify knowledge deficits that a student thought would be shared by others and which was relevant to their goals. Low mean *Value for Others* scores presumably indicated the tendency to identify knowledge deficits that were perceived to be unique to the individual (i.e., *not* shared by others, such as basic knowledge presumed to be known by others) or knowledge deficits that were perceived to be irrelevant to others (e.g., a curiosity question perceived to be unrelated to the course assessments).

Value for Self and *Value for Others* were related in two ways: 1) the variables were significantly and positively correlated with each other and 2) the mean *Value for Self* score was significantly and consistently higher than the mean *Value for Others* score. The two appraisals were correlated at .68, so approximately half of the variance in one appraisal was explained by the other. Appraisals of *Value for Self* were significantly and consistently higher than appraisals of *Value for Others*. Ninety-five percent of students had mean scores for *Value for Self* that were equal to or greater than mean scores for *Value for Others.* That is, for nearly all students questions first needed to be perceived to be valuable for one's individual learning, if they were to have value for classmates' learning.

On the surface, the strong positive correlation between the two appraisals may suggest that *Value for Self* and *Value for Others* were tapping a single construct, such as a question's general value or importance. As I discuss in later sections, however, these two appraisals predicted outcomes differently and were predicted by different motivational variables, suggesting they are unique appraisals.

What is indicated about a learner when *Values for Self* and *Values for Others* are aligned or misaligned? Consider three scenarios in which students appraised *Value for Self* and *Value for Others* as 1) both low; 2) both high; or 3) when *Value for Self* was high and *Value for Others* was low. For students in the first scenario, their appraisals suggest that they saw little value in their questions for their own learning and assumed their questions would not likely hold much value for others as well. Such students may have generated questions that they perceived to be minor knowledge deficits or addressed knowledge that was peripheral to the instructed material or course assessments.

For students in the second scenario—both high appraisals—their appraisals suggest that they perceived their questions to constitute important knowledge deficits that would be useful for both their individual learning and their peers' learning. Such students may have generated questions that they perceived to be important, relevant, or substantial deficits and believed that the answers to their questions would help others. By contrast, for students in the third scenario—high *Value for Self* and low *Value for Others*—their appraisals suggested that they generated questions that they perceived to be important, substantial deficits for themselves yet they reported that they thought the answers to their questions would hold relatively little value for others. Students who appraised their questions as having high value for their own learning and low value for peers' learning may have done so because they believed their questions were basic or rudimentary or because they believed their knowledge deficit was not important for how students would be assessed in the course (e.g., questions that were interesting to an individual but deemed irrelevant for the course exam). I further discuss what *Value for Self* and *Value for Others* may mean in later sections in light of their relations with motivation and achievement.

In summary, question characteristics were largely independent of each other. Frequently-occurring question types included several types that have been associated with higher-order thinking. The characteristics of students' questions have the potential to indicate aspects of student cognition that produced the questions as well as indicate information about students' knowledge and knowledge deficits.

Research Question 2: What are the relations among motivational self-beliefs and perceived teacher emphasis on mastery goals and the quantity, qualities, and self-appraisals of student-generated questions?

Overall, the results indicated that some motivational self-beliefs had statistically significant relations with question characteristics. Self-efficacy, personal interest, and individual goal orientations predicted different question characteristics. Several question characteristics, in turn, predicted in-class self-efficacy perceptions.

Self-Efficacy and Question Characteristics

Results from Analysis A indicated that more self-efficacious students tended to generate fewer procedural questions. However, the model accounted for a modest and non-significant amount of variance in the number of procedural questions. Self-efficacy beliefs are perceptions about one's competence for a specific task. To the extent that students' competency beliefs were calibrated with actual competence, it is logical that more competent student had fewer questions regarding the execution of procedures, a set of essential skill for being successful in the course. Much of the required and optional homework focused on developing procedural knowledge. Efficacy beliefs predict student initiation, effort, and persistence at academic tasks (Bandura, 1997; Schunk & Zimmerman, 2006). Therefore, highly self-efficacious students may have invested time and energy solving chemistry problems outside of class and consequently had fewer questions while watching the instructor demonstrate how to solve problems.

Turning to post-class self-efficacy, two question characteristics predicted students' self-efficacy beliefs at the end of class meetings, after accounting for the effect of initial self-efficacy beliefs. The number of procedural questions had a small negative effect on end-of-class efficacy beliefs. In addition, the percentage of closed questions had a small positive effect on end-of-class efficacy beliefs. Taken together, the results suggest that for this sample and context the qualities of the questions students generated while listening to lectures were indicative of students' end-of-class efficacy beliefs for learning the instructed material. Students who generated more procedural questions or a greater percentage of closed questions tended to report lower self-efficacy, beyond what would be predicted by initial efficacy beliefs. Why might this be?

Self-efficacy beliefs for academic tasks are shaped by previous mastery experiences, vicarious experiences, social persuasion, and emotional and physiological states (Bandura, 1977, 1986; Usher & Pajares, 2008). As students listened to lectures and monitored, recorded, and appraised their knowledge deficits, their questions may have been linked to mastery experiences, vicarious experiences, and emotional and physiological states during class meetings and thus directly shaped their post-class self-efficacy beliefs.

Regarding mastery experiences, the questions students generated during class lectures may have been a part of students' monitoring of the extent to which they had successfully understood the instructed material. More questions, especially procedural questions, may have indicated to students that they had more extensive knowledge deficits. In addition, the more students were able to sharply identify the exact nature of their knowledge deficit or the more knowledge deficits themselves were narrow (as indicated by closed syntax), the more they may have felt they had understood or were capable of understanding the material.

Vicarious experiences observing classmates monitor, record, and appraise their questions may have shaped post-class self-efficacy beliefs. Students sat just inches away from other students in the large lecture hall and could have easily viewed the number and possibly the contents of classmates' questions, including whether they had ticked off whether questions had been resolved. Information regarding the number and nature of peers' questions may have influenced students' self-efficacy beliefs at the end of class meetings. For example, a student's confidence for learning the material may have been reduced when considering her many procedural questions after observing that peers seated nearby had few.

Finally, the emotional and physiological states that students experienced during class in response to questions may have influenced post-class efficacy beliefs. The act of generating a question has been associated with cognitive states such a confusion, perplexity, and anxiety (Dillon, 1988; Festinger, 1962; van der Meij, 1994) and emotions, such as frustration and interest (D'Mello & Graesser, 2011; Silvia, 2008). Students who generated more procedural questions or more questions that suggested larger knowledge gaps might have experienced feelings of anxiety, confusion, or frustration, which lead to lowered self-efficacy perceptions at the end of class meetings.

It bears mentioning that the end-of-class self-efficacy beliefs may have influenced the number and nature of student-generated questions rather than having been the result of them. While post-class self-efficacy beliefs were reported at the end of class meetings, these beliefs may not have changed during lectures. For example, students who prepared for class by practicing problem sets and reading relevant parts of the textbook *before coming to class* may have arrived to class feeling self-efficacious and consequently generated fewer procedural questions or were more likely to generate closed questions.

Goal Orientation and Question Characteristics

Individual Goal Orientations and Self-Appraisals

Each individual goal orientation examined in this study—mastery approach, mastery avoidance, and performance—was significantly and positively related to students' appraisals of their questions' value for furthering their individual learning as well as the learning of others. The more students endorsed mastery approach, mastery avoidance, or performance goals, the greater the value they perceived in their questions. The positive relationships between individual goal orientations and self-appraisals suggested that for this sample and context perceiving value in one's questions was associated with both the motive to *develop* competence as well as the motive to *demonstrate* competence.

In this study individual goal orientations were moderately and positively correlated with each other. Students who endorsed performance goals also tended to endorse mastery approach and mastery avoidance goals. Performance approach and performance avoidance goals were strongly correlated with each and factor analyses suggested a single factor. Despite theoretical distinctions between performance approach and performance avoidance goals (Elliot & Church, 1997; Elliot & McGregor, 1999; Elliot & Harackiewicz, 1996), researchers commonly find that performance approach and avoidance scales demonstrate strong positive correlations (e.g., Elliot & Murayama, 2008; Pugh, Linnenbrink-Garcia, Koskey, Stewart, & Manzey, 2010; Ross, Shannon, Salisbury-Glennon, & Guarino, 2002; Senko & Harackiewicz, 2005).

Murayama, Elliot, and Yamagata (2011) suggested that perceived competence may be a factor that moderates the correlation between performance approach and avoidance. The authors hypothesized that for individuals with low perceived competence, concerns for demonstrating competence and avoiding demonstrating incompetence commingle whereas for individuals with higher perceived competence concerns with demonstrating incompetence (i.e., performance avoidance) are reduced and are therefore more distinct from performance approach goals. In this study students were enrolled in CHEM1031 in the spring semester, an atypical course sequence for STEM majors who typically take the course in the fall. According to the instructor of the course, students often take the course during the spring may having failed or withdrawn from the course previously or needed to take the course off-sequence due to missing pre-requisite math courses, factors which may indicate that students in this sample were likely to have relatively low perceived competence.

Mastery approach, mastery avoidance, and performance goal orientations exhibited weak to moderate positive correlations (r =.32 to .41), indicating that students tended to endorse multiple achievement goals. Positive correlations among mastery approach, mastery avoidance and performance goals are typical in the goal orientation literature (for a meta-analysis, see Hulleman et al., 2010). It is understandable why students may have endorsed multiple achievement goals in a competitive gateway course in their major. For many students, the desire to *develop* competence in solving objective chemistry problems was tied to their desire to *demonstrate* that competence in assessments that measured achievement by normative standards.

Multiple goal orientations may combine to exert additive, interactive, or other effects beyond the main effects of individual goal orientations (Barron & Harackiewicz, 2000; Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002). Results from the LCA suggested that multiple goals were not interactive for this sample and context. The two latent classes exhibited different levels of endorsement for the three individual goal orientations (medium vs. high) and different levels of *Value for Self* and *Value for Others* (medium vs. high) but the two classes did not differ in the relations among goal orientations (similar scores for performance and mastery approach and relatively lower scores for mastery avoidance) or the relations among self-appraisals (*Value for Self* were higher than *Value for Others*). The finding that different levels of endorsement of goal orientations between Class 1 and Class 2 were associated with a similar pattern of self-appraisals suggested that endorsing multiple goals did not have an interactive effect on questioning.

In Analysis B, individual goal orientations accounted for a significant amount of variance in *Value for Others* but not *Value for Self*. Results from the path analysis and partial correlations suggested that the zero-order positive correlation among *Value for Self* and individual goal orientations occurred via the effect of individual goal orientations on *Value for Others.* Individual goal orientations did not account for a significant amount of variance in *Value for Self* after accounting for shared variance (in the partial correlations) or shared error variance (in the path model) with *Value for Others*.

Questions can have different meanings in the minds of learners. For example, a question can be perceived as a weakness to be exploited by others, a direction for self-improvement, a falling short of complete understanding, an opportunity to demonstrate mental acuity or interest to others, or a path to being more academically competitive.

Accordingly, the meaning of a question's value for others may be influenced by a students' goal orientation. Students who held performance goals may have viewed the answer to their question—their knowledge *deficit*—as a potential asset to others as students competed for grades. In CHEM1031, final grades were established using a curve and therefore a student's achievement in the course was directly influenced by the performance of others. By contrast, students who adopted mastery approach or mastery avoidance goals may have perceived value in their questions for different reasons. For example, mastery-oriented students may have been more likely to perceive the remediation of their knowledge deficits as useful to others who, like themselves, may be striving for growth and self-improvement or who were seeking to avoid falling short of such goals.

In this study, the learning task was to comprehend a fast-paced lecture in which the instructor disseminated a large amount of information quickly and moved rapidly from one worked example to the next. In this context, students who adopted a mastery *avoidance* goal orientation, which explained the most unique variance in *Value for Others*, may have been particularly apt to perceive value for others in their questions because they perceived their question as evidence that they had not mastered the material. The unique relationship between *Value for Others* and mastery avoidance in this context was corroborated by the results from the *t* tests comparing questions characteristics between students who increased or decreased on motivational constructs. Students who increased in mastery avoidance goal orientation from Wave 1 to Wave 2 assigned significantly higher *Value for Others* to their questions compared to students who decreased in the endorsement of mastery avoidance goals.

The slightly different pattern of results from the analysis with the Winsorized scores, however, suggested that the strength of the relations between mastery avoidance and *Value for Others* may have been driven by a small number of students with low scores on both variables. Furthermore, *Value Gap* exhibited a significant negative Spearman correlation with performance goals but not with mastery approach or mastery avoidance goals. Spearman correlations are based on rank-ordered relationships and less susceptible to the influence of outliers (Abdullah, 1990). Accordingly, in this sample and context performance goals appeared to have a more consistent negative relationship with the value gap compared to mastery approach and mastery avoidance goals, which demonstrated greater variability between students.

Consistent with the overlapping nature of individual goal orientations in this sample, mastery approach, mastery avoidance, and performance goals explained shared variance in *Value for Others*. While mastery avoidance accounted for the greatest amount of unique variance, individual goal orientations explained more shared variance in *Value for Others* (6.5% to 7.4%) than unique variance (1.9% to 3.2%). Students who adopted multiple goals may have perceived value for others in their questions for a variety of different reasons and these reasons may have varied question-to-question and context-to-context.

Relations Among Individual Goal Orientations, Teacher Emphasis on Mastery Goals and Other Question Characteristics

In addition to the relations with self-appraisals, individual goal orientations and teacher emphasis on mastery goals exhibited significant relations with other question characteristics and across time, which I describe briefly below.

Mastery avoidance goals exhibited a significant weak positive correlation with the percentage of unanswered questions. The more students endorsed mastery avoidance goals, the greater the percentage of questions that remained unresolved by the end of class meetings. As I discussed previously, the percentage of unanswered questions may have indicated the degree to which students used strategies to resolve knowledge deficits or may have indicated the extent to which knowledge deficits were central or substantial. Studies that have examined relations among goal orientations and how students process information (e.g., deep vs. shallow processing) or use strategies while learning have not typically included mastery avoidance scales. Yet prior research has found that mastery approach goals were associated with deep processing, effective strategy use, and self-regulation (e.g., Ablard & Lipschultz, 1998; Ames & Archer, 1988; Elliot, McGregor, & Gable, 1999; Karabenick, 2004). In light of prior research, one explanation for the positive correlation among mastery avoidance and the percentage of unanswered questions is that the more students reported striving to avoid falling short of mastering the material, the more they tended to identify substantial gaps in their understanding, ones that would not likely be easily resolved during the remaining lecture.

Growth in performance goals from Wave 1 to Wave 2 were positively associated with the number of procedural questions. Students who increased in their endorsement of performance goals relative to the sample generated a greater number of procedural questions, with a large effect size (d = .77). One possible explanation is that in the context a fast-paced lecture in a challenging, competitive course, generating questions might have exacerbated some students' concerns with normative performance, especially when their questions targeted key knowledge and skills.

Students' perceptions of the instructor's emphasis on mastery goals at Wave 1 were unrelated to all question characteristics. Interestingly, students' perceptions of the instructor's emphasis on mastery goals increased significantly during the time period in which students were using the question logs. By contrast, all other motivational self-beliefs including individual goal orientations did not exhibit mean changes between Wave 1 and Wave 2. Students' perceptions of the environment as being supportive of mastery goals may have been influenced by the use of question logs. The instructor provided time for students to think about and record their questions, verbally reinforced the importance of students' questions as feedback for planning future lectures, and devoted time to responding to questions from previous lectures. It is not surprising, then, that students may have changed their perception of the goal structure of the course in light of the introduction of instructional practices that supported mastery learning.

Relations with Personal and Situational Interest

Personal Interest

Results from Analysis A indicated that personal interest was not significantly related to question characteristics. On the surface, the absence of significant relations may seem counterintuitive, since one might expect that students with greater interest in a topic may be more likely to exhibit curiosity about the instructed material or to think more deeply or expansively during lectures and therefore generate more questions or perceive more value in their questions. Yet, the weak and non-significant relations among personal interest and question characteristics are understandable in light of the nature of the material and the learning task. Students who reported greater personal interest in chemistry may also have brought more prior knowledge to bear on their learning during lectures and therefore had few conceptual knowledge gaps.

Furthermore, the learning environment provided few opportunities for students to engage in deep or expansive processing of the presented information or to pursue ideas of interest. The instructed material consisted largely of objective conceptual and procedural knowledge. In the context of a fast-paced lecture covering content that students often find challenging, students may have had little time to generate curiosity questions. Anecdotally, I observed very few questions during the coding process that clearly suggested curiosity and wonderment. Overwhelmingly, students' questions tended to indicate practical concerns with specific gaps or deficits—the nuts and bolts of comprehension—rather than more expansive pondering about the topics, a finding that is consistent with prior research (Chin, Brown, & Bruce, 2002; Pedrosa de Jesus et al., 1997; Texiera-Dias et al., 2003). In fact, the lack of evidence of curiosity questions influenced my decision to abandon some coding schemes that focused on interest-driven questions (Pedrosa et al., 1997, 2003; Scardamalia & Bereiter, 1992). In addition, prior research has found that the association between personal interest and achievement is the strongest when achievement reflected deep, conceptual understanding since personal interest has been more associated with deep processing than with surface processing (Wigfield & Cambria, 2010). Achievement in the context of the present study emphasized efficiency while solving well-defined chemistry word problems rather than deep conceptual knowledge.

While personal interest at Wave 1 was not significantly related to self-appraisals, students who increased in personal interest from Wave 1 to Wave 2 tended to perceive greater value for others in their questions. As students develop personal interest they are theorized to accrue both knowledge and value for knowing in a domain (Hidi & Renninger, 2006; Renninger, 2000). One explanation is that as students gained knowledge—and valued knowledge and knowing in the domain—they may have been more likely to perceive that their questions addressed knowledge deficits that were likely to be important and not simply rudimentary, and thus believed others may benefit from the answer.

Situational Interest

Situational interest was not significantly related to any question characteristics. Thus, students' questioning behaviors appeared to be independent of whether they found class meetings stimulating or boring. Again, at the surface, this pattern of non-significant relations with situational interest may seem surprising. One might expect that students who were very interested during lectures would generate more or qualitatively different questions or perceive more value in their questions compared with students who found class meetings dull or tedious. Alternatively, one can imagine that students reported high interest for lectures that they perceived to be comprehensible and during which knowledge deficits were few, were deemed relatively unimportant, or which were resolved during class meetings. The significant positive correlation between perceived comprehension and situational interest supports this explanation. It may be that situational interest was related to student-generated questioning in different and possibly opposite ways for different students and within students across time, and therefore, for the sample as whole, there was no significant relationship.

Person-Centered Analyses

The latent class analysis (LCA3) found evidence of two latent classes that were homogeneous with regard to their Wave 1 motivations and question characteristics. Mean scores for the two classes indicated that approximately one-quarter of the students reported lower levels of personal interest, self-efficacy, and individual goal orientations compared to the rest of the sample. Students who had lower levels on these motivational constructs also had lower levels of appraisals, but others question characteristics were relatively similar between the two classes. The two classes had similar Exam 2 scores.

Thus, the person-centered analysis did not suggest the presence of groups within the sample who differed dramatically in how their patterns of motivational variables related to patterns of questioning.

Summary

Results suggested that several motivational self-beliefs were related to students' questions during lectures. Personal interest and self-efficacy were associated with fewer procedural and conceptual questions, respectively. In the context of this study, efficacy beliefs or personal interest may have been associated with students' out-of-class preparation for lectures, which led students to generate fewer questions during class meetings. Individual goal orientations were significantly and positively related to the extent to which students saw value for others in their questions and negatively related to the gap between the two appraisals. Individual goal orientations tended to explain shared variance in *Value for Others.* Results suggested it was primarily the extent to which students reported a desire to achieve in the course that was associated with this appraisal, rather than the purpose behind the achievement striving. For students in this sample, individual goal orientations were themselves moderately correlated, indicating that many students adopted multiple goals. An important finding of the study was that question characteristics explained variance in post-class self-efficacy beliefs after accounting for initial general self-efficacy beliefs for the course. Self-generated questions may have provided students with information about their level of mastery of the material, which might have influenced their situated efficacy beliefs.

Research Question 3: Relations Among Question Characteristics and Exam Scores

I asked two research questions about the relations among question characteristics and exam scores. (RQ3a) Can exam scores be predicted by the quantity and qualities of student-generated questions? and (RQ3b) Do self-perceptions at the end of class meetings mediate the relations among question characteristics and Exam 2 scores?

Regarding RQ3a, results from the multiple regression and two path models indicated that question characteristics can predict a medium-to-large amount of variance in exam scores. The *percentage of closed questions*, *Value for Self*, and *Value for Others* accounted for a significant 18.6% of the variance in exam scores. The percentage of closed questions and *Value for Others* had a positive relationship with exams while *Value for Self* had a negative relationship. I discuss the meaning of each predictor in turn.

Closed Questions and Exam Scores

Why would the percentage of closed questions positively predict exam performance? As I described previously, the percentage of closed questions may have indicated cognitive processes as well as the nature of the knowledge deficits. Regarding cognitive processes, the percentage of closed questions may have indicated the tendency to refine knowledge deficits to arrive at the precise nature of the gap or to construct plausible answers, which could be confirmed, refuted, or clarified. Regarding the nature of the deficit itself, closed question may have indicated smaller gaps in knowledge compared to open questions. These cognitive processes and types of knowledge deficits were both adaptive in this environment.

It is important to emphasize that while closed questions appeared to be adaptive in this context such questions may be maladaptive in other contexts. For example, posing closed questions may prematurely narrow possible solutions to ill-defined problems or may misguide or confuse a learner by pursuing a question that contains false presumptions (van der Meij, 1994). In other settings, open questions may be more adaptive. For example, open questions may be more likely to reveal unforeseen answers or support expansive or creative thinking and therefore be adaptive in settings that value these aspects of learning.

Self-Appraisals and Exam Scores

Turning to self-appraisals, *Value for Self* and *Value for Others* accounted for a small-to-medium amount variance (9.5% in Analysis B) in exam scores, with *Value for Self* negatively predicting exam scores and *Value for Others* positive predicting exam scores. To understand the relations of these variables with exam scores it is important to keep three relationships in mind. First, *Value for Self* and *Value for Others* were positively correlated with each other (r = .68). Second, *Value for Self* and *Value for Others* exerted opposite and approximately equal effects on Exam 2 scores ($\beta_{\text{Value for Self}}$ = -.40; $\beta_{\text{Value for Others}}$ = .43). Third, for nearly all individuals (only 3 exceptions in 103 participants) the mean *Value for Self* was equal to or higher than the mean *Value for Others.* With regard to the first and second point, *within an individual* one appraisal tended to dampen the effects of the other, such that when *Value for Self* and *Value for Others* were the same, the effects of the two appraisals canceled each other and were essentially unrelated to exam scores. When *Value for Other* scores were lower than *Value for Self*, the dampening effect of *Value for Others* on *Value for Self* was reduced and the net result was a negative association with exam scores. When *Value for Self* was high and *Value for Others* was low, the former exerted a strong negative effect on exam scores. Thus, the greater the value gap between *Value for Self* and *Value for Others*, the lower the exam score.

Value for Self was negatively related to exam scores after accounting for *Value for Others.* One reason why *Value for Self* may have had a negative relation with exam scores is that this appraisal may have indicated the extent and relevance of one's knowledge gap in light of the exam. Students with higher *Value for Self* scores may have perceived their knowledge deficits as more substantial while students with lower *Value for Self* scores perceived their knowledge deficits as less substantial. That *Value for Self* negatively explained exam scores suggested that students tended to accurately appraise the extent to which their knowledge deficits were important in light of how they would be assessed.

Why did *Value for Others* have a positive relationship with exam scores after accounting for *Value for Self*? Stated differently, why was the gap between *Value for Self* and *Value for Others* negatively related to achievement? As I described previously, the gap between *Value for Others* and *Value for Self* may have reflected the extent to which knowledge deficits were perceived to be shared and relevant. A large gap between *Value for Self* and *Value for Others* may have reflected the extent to which students perceived their questions as representing basic or rudimentary gaps in knowledge, gaps that few others would have. This explanation was supported by the significant negative relationship between prior GPA and the value gap. Lower-achieving students in the study tended to report greater gaps between the two appraisals.

The value gap appeared to capture the extent to which students believed their questions were shared and relevant to others, with greater value gaps associated with lower performance. These relations may shed light on why lower-achieving students tend to *ask* fewer questions than high-achieving students (Good, 1981; Good et al., 1987). Lower-achieving students may be less likely to ask their questions to others because they perceive that their questions are rudimentary and may therefore expose embarrassing knowledge deficits (Karabenick & Knapp, 1988; Karabenick & Sharma, 1991; Ryan & Pintrich, 1997) or take up class time to resolve gaps in understanding that are irrelevant to others (Bergey et al, 2012). In colloquial language, the value gap may reflect the extent to which students think theirs is a “stupid question.”

One might assume that students who perceived their questions as less sophisticated would feel less efficacious and perceive lower comprehension, yet the value gap was unrelated to post-class self-efficacy or perceived comprehension. Results suggested that students with different value gaps differed in the extent to which their efficacy beliefs appeared to be accurately calibrated with performance on the exam. Relative to how students would actually perform on Exam 2, students with large value gap scores tended to report overly confident self-efficacy beliefs while students with small or no value gap tended to report under-confident self-efficacy beliefs. Thus, an interesting picture emerges with regard to students’ perceptions of their questions relative to others’ questions and relative to their efficacy for learning. If we assume that the value gap represented perceptions of basic versus sophisticated questions (large vs. small value gaps, respectively), then students appeared to accurately evaluate the extent to which their questions indicated a basic or sophisticated gap in knowledge. At the same time, students who perceived their questions as having less value for others than for themselves reported overly confident efficacy beliefs. Therefore, students may have made accurate appraisals regarding the level of sophistication of their questions but had efficacy beliefs that were somewhat inaccurate regarding their actual performance on the exam. The relations among students’ appraisals, their goal orientations, efficacy beliefs, and achievement appear to be complex, and constitute an aspect of the study that merits further analysis and future research.

Achievement, Closed Questions, and Self-Efficacy

In Analysis A, the percentage of closed questions and post-class self-efficacy beliefs significantly and positively predicted exam scores. As described above, closed questions may indicate effortful cognitive processes as well as smaller knowledge gap, both of which appear to have been adaptive in this context. Students who worked to articulate the precise nature of their knowledge deficit may have been more likely to resolve their questions outside of class. More generally, the percentage of closed questions may have reflected more general cognitive skills and habits, such as effortful processing and active (versus passive) learning.

Self-efficacy beliefs at Wave 1 and at the end-of-class meetings were each positively associated with exam scores, a finding consistent with a wide-range of studies (Lau & Roeser, 2002; Multon, Brown, & Lent, 1991; Schunk & Zimmerman, 2006; cf. Lawson, Banks, & Logvin, 2007). In Analysis A, Wave 1 self-efficacy had only a modest effect on post-class self-efficacy ($\beta = .255$). After accounting for the various ways in which Wave 1 self-efficacy was associated with exam scores (e.g., via post-class self-efficacy and interest), the remaining variance shared between Wave 1 self-efficacy and exam scores was small (as indicated by the good model fit). This is consistent with prior research that has found efficacy beliefs are most predictive of achievement when the individuals are asked about specific (versus general) tasks and when the tasks are aligned with outcome measures (Bong, 2004; Pajares, 1996). Compared with the Wave 1 measure of self-efficacy, post-class efficacy beliefs were more specific (i.e., successfully mastering the instructed material vs. success in the course) and more closely aligned with the outcome (i.e., Exam 2). As discussed above, post-class self-efficacy was predicted by several questions characteristics after controlling for initial self-efficacy for the course. The indirect effects between question characteristics and Exam 2 scores via post-class self-efficacy, however, were not significant. Thus, the number of procedural or conceptual questions and the percentage of unanswered questions did not directly or indirectly predict exam scores.

Summary

Several characteristics of students' questions were predictive of exam scores. Students who generated a greater percentage of closed questions performed better on the course exam. The percentage of closed questions may have indicated cognitions that refined knowledge deficits or tested plausible answers or they may have indicated less substantial knowledge deficits, both of which appeared to be adaptive in the context of the course and the assessment. Students who perceived their questions as having similar value for their own learning compared with the value for others' learning performed better of the course exam. The inverse relation between the value gap and GPA and Exam 2 scores indicated that students who perform more poorly academically tended to assume their questions would have less value for others compared to the value for their own learning. It may be that higher performing students generated more sophisticated questions or questions that were more aligned with assessments compared with lower performing students. Understanding how and why students appraise a question's *Value for Others* and its relation with achievement and motivations merits further research.

Research Question 4: Question Generation and Question Asking

Research Question 4 asked: Do students who ask questions during class differ from students who do not in terms of their motivational self-beliefs, perceived teacher emphasis on mastery goals, help seeking beliefs, achievement, and the quantity, qualities, and self-appraisals of generated questions? Students asked few questions during class meetings, consistent with prior research on question asking in university settings (Pedrosa de Jesus et al., 1997; Person & West, 1991; Texiera-Dias et al., 2003; West & Person, 1993). While students rarely asked their questions, this sample of students generated many, ranging from 97 to 158 questions each class period. The number of asked questions constituted less than 1% of the questions students recorded in question logs. Therefore, the study revealed that in this context very few of the questions students generate were asked as a part of class-wide discourse during class meetings.

Asking a question in public can be a vulnerable act, since questions reveal what one does not know along with other information about one's beliefs, intentions and interests (Dillon, 1988). Therefore, seeking help by asking a question can present a threat to feelings of competence (Karabenick & Sharma, 1991; Ryan & Pintrich, 1997; van der Meij, 1990). In addition, there were several reasons why students in CHEM1031 may not have asked their questions. Approximately 120 students attended each lecture; speaking in front of such a large group likely intimidated many students. Furthermore, normative behavior in large-enrollment lectures courses is for students to listen quietly while the instructor lectures. Moreover, students had access to other channels of communication with instructors where logistical and psychological barriers to question asking may have been minimized. For example, the instructor regularly held supplementary study sessions and students attended a weekly recitation class with conditions more hospitable to asking questions (e.g., student–teacher ratios of approximately 25 to 1).

The question logs themselves constituted a means for students to communicate their questions to the lecture instructor and recitation instructors. It is possible that questions that were recorded might have been asked if this communication channel were not available. Rates of question asking during the 8 lectures when students used the question logs, however, were similar to rates of question asking before and after the use of the logs. Therefore, it is unlikely that students withheld questions during class meetings because of the question logs. To the contrary, some of the questions that were asked aloud were asked during the pause in instruction when students were allotted time to record their questions.

Implications

Implications for Theory and Research

The theoretical framework that I presented in this study has implications for interpreting and designing studies of student questioning. Findings from this study highlighted the importance of examining students' questions and their relation to motivation and achievement *in light of* the context of the learning task and the instructional setting in which the questioning took place. Echoing Yang's (2006) critique of research that decontextualizes students' questions, this study demonstrated the strong connection between students' questions, the learning task, and the learning context. For example, in the context of this study, open questions and procedural questions were negatively related to achievement, reflecting the nature of the instructed material and course assessments, both of which strongly emphasized mastering procedural skills and knowledge required to solve chemistry word problems.

Furthermore, how students' individual motivations related to the number and nature of their questions reflected the learning task and context. Consider, for example, the absence of significant relations among questions and mastery goal structure and personal interest in this study. Students were tasked with comprehending the content of fast-paced lectures and had few opportunities to pursue personal interests or engage in deep learning. In this context, questions appeared to serve the function of identifying aspects of the lecture that were not completely understood rather than areas of interest, curiosity, or personal relevance that went beyond the material as it was presented. A very different pattern of relations among these variables and questions would be expected in contexts with greater opportunities to examine topics of interest and with learning tasks that are more open, flexible, or subjective. Thus, researchers interested in studying student questioning in authentic learning environments should be mindful of the inter-related influences of individuals' perceptions and prior knowledge, the learning task, and the social context, and might employ methods that examine the integrated and interactive nature of these factors.

Findings from this study also indicated that students' evaluations of their questions may be an important indicator of how students perceive their level of understanding relative to peers. In this study students distinguished between a question's value for their individual learning and for peers' learning, and the relation among these appraisals was significantly associated with achievement. How and why students appraise their questions and how self-appraisals of questions relate to information-seeking and help-seeking behaviors—and ultimately learning—merits further investigation.

The results revealed a complex relationship between student questioning, learning, and achievement. This complexity stems in part from the fact that questions can indicate adaptive cognitive *processes* (e.g., comprehension monitoring, self-explanation, clarification) as well as knowledge *deficits*. Identifying a deficit in understanding does not ensure that the deficit will be resolved and therefore learning is a *potential*—but not a *necessary*—outcome of generating a question. Even when self-questioning results in resolving knowledge deficits, questions may indicate different levels of understanding and therefore be associated with either high or low achievement. Understanding the cognitive, motivational, and social factors that mediate and moderate successful information-seeking behavior and the conditions under which question generation lead to individual learning as well as normative achievement are important areas for future research.

Question logs are a promising method for studying questioning in authentic classroom contexts. In this study, the use of the logs appeared to be minimally disruptive to instruction, even in a context with a high teacher–student ratio and with teacher-directed, content-heavy instruction. Furthermore, the use of the question logs appeared to make changes to the classroom environment that resulted in students' perceiving the teacher as more supportive of mastery goals. In addition, question logs afforded the opportunity to collect data on a small number of situated perceptions at the end of class meetings. Sampling students' motivational beliefs across time may reveal how experiences during learning relate to changes in self-beliefs. For example, in this study, post-class self-efficacy was only modestly predicted by initial self-efficacy beliefs, suggesting that self-efficacy beliefs for learning specific aspects of the instructed material differed substantially from initial self-efficacy beliefs for succeeding in the course in general.

Implications for Teaching

In this study, characteristics of students' questions were predictive of perceived levels of understanding, efficacy beliefs, and exam performance. Thus, questions have the potential to provide students and teachers with important information about levels of understanding, as Chin and Osborne (2006) argued in their extensive review of student questioning in science. Questions represent unique sources of information regarding student learning because questions articulate what students perceive as relevant gaps in their understanding. When students' questions are fed back to instructors in a timely and private manner, questions can be an important source of information for formative assessment. Teachers might explore ways to become more aware of students' questions and to support information seeking behaviors so that questions are likely to result in productive answers.

Limitations

The results of this study should be interpreted in light of several limitations. One limitation of the present research was that I was unable to code for some aspects of question quality. The GPH taxonomy identifies the type of cognition elicited by a question and its logical answer but does not capture content-specific or context-specific aspects of question quality. For example, the GPH taxonomy treats all causal questions as a single question type, when causal questions might range from very basic to very sophisticated or inquire about a central idea in the lecture or a very peripheral one. A content expert's evaluation of the sophistication or centrality of questions could illuminate important variation in question quality not captured in the present study.

Another limitation is that the analyses in this study aggregated question characteristics across time and question types. Analyzing students' questions at level of the 8-lecture unit ignored potentially important aspects of or relationships among question characteristics. By aggregating question characteristics across questions, I ignored relationships at the level of the individual questions. For example, students may have appraised closed or procedural questions differently than open or conceptual ones. Similarly, by aggregating question characteristics across time, I ignored changes within the individual across time or sequencing effects. For example, *when* students generated different types of questions may have been related to motivation and achievement in different ways than for the unit as a whole.

An additional limitation is that the self-appraisal items were previously untested and students may have interpreted them in different ways. Drawing conclusions from the patterns of self-appraisals—individually and in relation to each other—was complicated by the various ways these items could be interpreted by students. For example, students may have calculated a question's *Value for Others* in different ways, such as considering a specific peer (e.g., a study partner) or a group within the class (e.g., low- or high-performing students) or by estimating an average value for all students. Understanding how students interpret these items may lead to the development of better measures of students' appraisals and a clearer understanding of what these items indicate, including in different contexts.

The sample size was relatively small for some of the path models tested in the study. Accordingly, I may have lacked statistical power to detect small effects (Type II error). In addition, the large number of variables and analyses included in this study increase the possibility that some significant findings were spurious (Type I error).

Finally, it is important to keep in mind that the study examined question generation in a single instructional setting with a particular instructor while learning a particular subject matter. Patterns and relationships found in the data may reflect idiosyncratic aspects of the sample, learning task, or environment. For example, taking CHEM1031 in the spring semester represented a non-traditional course sequence and therefore the sample many not be representative even of chemistry students at Temple University. To facilitate the evaluation of the extent to which findings from the present study may generalize to other populations, tasks and settings, I have presented my findings within a theoretical framework that describes student-generated questioning occurring at the intersection of the nature of the learning task, the social context, and individual characteristics.

Summary

This study examined the questions that students generated during chemistry lectures, students' appraisals of their questions, and the relations among various question characteristics, motivations, and achievement on a course exam. Students' questions—recorded in logs during class lectures—provided a unique window on student cognition, knowledge, and motivation during learning not often observable to educators and researchers. As is typical of many classroom settings, very few questions were asked aloud during class meetings, yet students' generated nearly a thousand questions that, on average, they believed were important for their individual learning and the learning of their peers.

The characteristics of students' questions significantly predicted and were predicted by various motivational variables. Individual goal orientations—individually and in combination—explained variance in appraisals, particularly the extent to which students thought their questions held the value for peers' learning that it held for their individual learning. This value gap, in turn, explained variance in exam scores and was associated with growth in mastery avoidance and personal interest. Results suggested that low-achieving students or students who declined in personal interest or mastery avoidance over time tended to perceived their knowledge deficits as unique and possibly rudimentary, and unlikely to be of value for peers' learning.

Importantly, students' questions were associated with normative changes in motivational beliefs. For example, growth in performance goal orientations over time was associated with more procedural questions and more unresolved questions. A particularly important finding of the study was that characteristics of students' questions explained variance in students' efficacy beliefs at the end of class meetings, beyond what was predicted by initial efficacy beliefs. Post-class self-efficacy along with the percentage of closed questions directly predicted performance on a course exam.

Overall, the findings from this study suggested that students' questions—their quantity, qualities, and students' appraisals of them—constituted important indicators of student understanding, motivation and performance in an undergraduate chemistry course. Prior research on question generation had primarily examined relations with cognitive variables and investigated these relations in one-on-one or laboratory settings. The pattern of significant relations in the present study among students' questions, motivational beliefs, and achievement suggests the promise of research that examines and facilitates students' questioning in authentic learning environments.

REFERENCES CITED

Abdullah, M. B. (1990). On a robust correlation coefficient. *Journal of the Royal Statistical Society, 39*(4), 455–460. doi:10.2307/2349088

Ainley, M., & Hidi, S. (2002). Dynamic measures for studying interest and learning. In P. R. Pintrich & M. L. Maehr (Eds.), Advances in motivation and achievement: New directions in measures and methods (Vol. 12, pp. 43–76). Amsterdam: JAI, Elsevier Science.

Ainley, M., & Patrick, L. (2006). Measuring self-regulated learning processes through tracking patterns of student interaction with achievement activities. Educational Psychology Review, 18(3), 267–286. doi:10.1007/s10648-006-9018-z

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317–332. doi:10.1007/BF02294359

Albergaria-Almeida, P. (2011). Critical thinking, questioning and creativity as components of intelligence. *Procedia - Social and Behavioral Sciences*, *30*, 357–362. doi:10.1016/j.sbspro.2011.10.070

Alexander, P.A. (1997). Mapping the multidimensional nature of domain learning: The interplay of cognitive, motivational, and strategic forces. In M.L. Maehr & P.R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10), (pp. 213-250). Greenwich, CT: JAI.

Alkharusi, H., & Aldhafri, S. (2010). Gender differences in the factor structure of the 2x2 achievement goal framework. *College Student Journal*, *44*(3), 795–804.

Allwood, C. M., & Eliasson, M. (1988). Question asking when learning a text-editing system. *International Journal of Man-Machine Studies*, *29*(1), 63–79.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261-271.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260-267.

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research*, *64*(2), 287–309.

Anderman, E. M., & Wolters, C. (2006). Goals, values and affects: Influences on student motivation. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (2nd ed.). New York: Simon & Schuster.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. Oxford, England: Grune & Stratton.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523-535.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, *29*(3), 344–370. doi:10.1016/j.cedpsych.2003.09.002

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior*, *11*(4), 365–374.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman.

Baram-Tsabari, A., & Yarden, A. (2007). Interest in biology: A developmental shift characterized using self-generated questions. *The American Biology Teacher*, *69*(9), 532–540.

Baram-Tsabari, A., & Yarden, A. (2009). Identifying meta-clusters of students' interest in science and their change with age. *Journal of Research in Science Teaching*, *46*(9), 999–1022.

Beck, I. L., McKeown, M. G., Hamilton, R. L., & Kucan, L. (1997). *Questioning the author: An approach for enhancing student engagement with text.* Newark, DE: International Reading Association.

Bergey, B. W., & Cromley, J. G. (2012). *How and when graduate students ask questions in a statistics course: Patterns and profiles.* Presented at the annual meeting of the American Education Research Association, Vancover, BC.

Bergey, B. W., Horvat, E. M., & Cromley, J. G. (2012). *The costs of questioning: Why graduate students withhold their questions from the instructor in an introductory statistics classroom.* Presented at the annual meeting of the American Education Research Association, Vancover, BC.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York, NY: McGraw-Hill Book Company.

Bernacki, M. L., Byrnes, J. P., & Cromley, J. G. (2012). The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology*, *37*(2), 148–161. doi:10.1016/j.cedpsych.2011.12.001

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, *17*(5), 564–577. doi:10.1016/j.learninstruc.2007.09.007

Besterfield-Sacre, M., Atman, C. J., & Shuman, L. J. (1997). Characteristics of freshman engineering students: Models for determining student attrition in engineering. *Journal of Engineering Education*, *86*, 139–150.

Bickel, R. (2007). *Multilevel Analysis for Applied Research: It's Just Regression!* New York, NY: Guilford Press.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain.* New York: David McKay.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, *54*(2), 199–231. doi:10.1111/j.1464-0597.2005.00205.x

Bong, M. (2004). Academic motivation in self-efficacy, task value, achievement goal orientations, and attributional beliefs. *The Journal of Educational Research*, *97*(6), 287–298.

Bong, M., & Hocevar, D. (2002). Measuring self-efficacy: Multitrait-multimethod comparison of scaling procedures. *Applied Measurement in Education*, *15*(2), 143–171. doi:10.1207/S15324818AME1502_02

Brown, J. S., Collins, A., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-493). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bruner, J. S. (1960). *The process of education*. Harvard University Press.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1986). *A study of thinking*. New Brunswick, NJ: Transaction Publishers.

Bugg, J. M., & McDaniel, M. A. (2012). Selective benefits of question self-generation and answering for remembering expository text. *Journal of Educational Psychology*, *104*(4), 922-931. doi:10.1037/a0028661

Buseri, J. C. (1988). Questions in Nigerian science classes. *Questioning Exchange*, *2*(3), 275–280.

Butler, R. (2007). Teachers' achievement goal orientations and associations with teachers' help seeking: Examination of a novel approach to teacher motivation. *Journal of Educational Psychology*, *99*(2), 241-252.

Butler, R., & Neuman, O. (1995). Effects of task and ego achievement goals on help-seeking behaviors and attitudes. *Journal of Educational Psychology*, *87*(2), 261–271. doi:10.1037/0022-0663.87.2.261

Cabin, R. J., & Mitchell, R. J. (2000). To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America*, *81*(3), 246–248.

Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research*, *61*(2), 157–178.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, *13*(2), 145–182.

Chi, M. T., De Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science*, *18*(3), 439–477.

Chin, C., Brown, D. E., & Bruce, B. C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education*, *24*(5), 521–549. doi:10.1080/09500690110095249

Chin, C., & Chia, L. G. (2004). Problem-based learning: Using students' questions to drive knowledge construction. *Science Education*, *88*(5), 707–727. doi:10.1002/sce.10144

Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, *44*(1), 1–39.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*(1), 1–49. doi:10.3102/00346543063001001

Chouinard, M. M. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, *72*(1), 1–108. doi:10.1111/j.1540-5834.2007.00412.x

Collins, A., Brown, J. S. and Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 453-494) Hillsdale, NJ: Lawrence Erlbaum Associates.

Conley, A. M. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology*, *104*(1), 32-47.

Cook, R. D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association*, *74*(365), 169–174. doi:10.1080/01621459.1979.10481634

Cook, R. D., & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, *70*(1), 1–10.

Costa, J., Caldeira, H., Gallastegui, J. R., & Otero, J. (2000). An analysis of question asking on scientic texts explaining natural phenomena. *Journal of Research on Science Teaching*, *37*(6), 602–614.

Craig, S. D., Sullins, J., Witherspoon, A., & Gholson, B. (2006). The deep-level-reasoning-question effect: The role of dialogue and deep-level-reasoning questions during vicarious learning. *Cognition and Instruction*, *24*(4), 565–591.

Cromley, J. G., Horvat, E. M., Tanaka, J., Michel, J., Perez, A. C., Tancredi-Brice, E., Raja, L., Vorndran, P., Wills, M., & Wills, T. W. (2011). *Cognition and motivation in science technology engineering and math retention.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *102*(3), 687-700.

Davey, A., & Savla, J. (2010). *Statistical power analysis with missing data: A structural equation modeling approach*. Philadelphia, PA: Routledge.

Dewey, J. (1910). Natural resources in the training of thought. In J. Dewey (Ed.), *How We Think* (pp. 29–44). Lexington, MA: D C Heath.

Dillon, J. T. (1982). The multidisciplinary study of questioning. *Journal of Educational Psychology*, *74*(2), 147–165.

Dillon, J. T. (1988). *Questioning and teaching: A manual of practice.* New York, NY: Teachers College Press.

Dillon, J. T. (1990). *The practice of questioning*. New York: Routledge.

Dillon, J. T. (1998). Theory and practice of student questioning. In S. A. Karabenick (Ed.), *Strategic help seeking: Implications for learning and teaching* (pp. 171-193). Mahwah, NJ: Lawrence Erlbaum Associates.

Diseth, A. A., Danielsen, A. G., & Samdal, O. (2012). A path analysis of basic need support, self-efficacy, achievement goals, life satisfaction and academic achievement level among secondary school students. *Educational Psychology*, *32*(3), 335–354.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. doi:10.1016/j.learninstruc.2011.08.003

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040-1048.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *3*(21), 215-225.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum. *The Journal of the Learning Sciences*, *8*(3&4), 391–450.

Elliot, A. J., & McGregor, H. A. (2001). A 2x2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501–519.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613-628.

Fenker, R. M. (1975). The organization of conceptual materials: A methodology for measuring ideal and actual cognitive structures. *Instructional Science*, *4*(1), 33–57.

Festinger, L. (1962). Cognitive dissonance. *Scientific American, 207*(4), 93-107.

Field, A. (2005). *Discovering statistics using SPSS, 2nd edition.* London England: Sage.

Fishbein, H. D., Eckart, T., Lauver, E., Van Leeuwen, R., & Langmeyer, D. (1990). Learners' questions and comprehension in a tutoring setting. *Journal of Educational Psychology, 82*(1), 163-170.

Flammer, A. (1981). Towards a theory of question asking. *Psychological Research*, *43*(4), 407–420.

Friedel, J. M., Cortina, K. S., Turner, J. C., & Midgley, C. (2007). Achievement goals, efficacy beliefs and coping strategies in mathematics: The roles of perceived parent and teacher goal emphases. *Contemporary Educational Psychology*, *32*(3), 434–458. doi:10.1016/j.cedpsych.2006.10.009

Glogger, I., Holzäpfel, L., Schwonke, R., Nückles, M., & Renkl, A. (2009). Activation of learning strategies in writing learning journals: The specificity of prompts matters. *Zeitschrift für Pädagogische Psychologie/ German Journal of Educational Psychology*, *23*(2), 95–104. doi:10.1024/1010-0652.23.2.95

Glogger, I., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, *104*(2), 452–468. doi:10.1037/a0026683

Good, T. L., Slavings, R. L., Harel, K. H., & Emerson, H. (1987). Student passivity: A study of question asking in K-12 classrooms. *Sociology of Education*, 181–199.

Graesser, A. C. (1985). *The psychology of questions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Graesser, A. C., Baggett, W., & Williams, K. (1996). Question-driven explanatory reasoning. *Applied Cognitive Psychology*, *10*, 17–31.

Graesser, A.C., Lang, K., & Horgan, D. (1988). A taxonomy for question generation. *Questioning Exchange*, *2*(1), 3–15.

Graesser, A. C., Lu, S., Olde, B. A., Cooper-Pye, E., & Whitten, S. (2005). Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition*, *33*(7), 1235–1247. doi:10.3758/BF03193225

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, *85*(1), 136-151.

Graesser, A. C., McMahen, C. L., & Johnson, B. K. (1994). Question asking and answering. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 517–538). San Diego, CA: Academic Press.

Graesser, A. C., Ozuru, Y., & Sullins, J. (2009). What is a good question? In M. G. McKeown & L. Kucan (Eds.), *Threads of coherence in research on the development of reading ability* (pp. 112-141). New York: Guilford.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, *31*(1), 104–137.

Graesser, A. C., Person, N., & Huber, J. (1992). Mechanisms that generate questions. In T. W. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and Information Systems* (pp. 167–187). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Guthrie, J. T., & Scafiddi, N. T. (2004). Reading comprehension for information text: Theoretical meanings, developmental patterns, and benchmarks for instruction. In J. T. Guthrie, A. Wigfield, & K. C. Perencevich (Eds.), *Motivating Reading Comprehension: Concept-oriented Reading Instruction.* (pp. 225–248). Mahwah, NJ US: Lawrence Erlbaum Associates.

Hackett, G., Betz, N. E., Casas, J. M., & Rocha-Singh, I. A. (1992). Gender, ethnicity, and social cognitive factors predicting the academic achievement of students in engineering. *Journal of Counseling Psychology*, *39*(4), 527-538.

Hancock, G. R., & Mueller, R. O. (Eds.). (2010). *The reviewer's guide to quantitative methods in the social sciences*. New York: Routledge.

Hidi, S. (1995). A re-examination of the role of attention in learning from text. *Educational Psychology Review*, *7*, 323–350.

Hidi, S., & Baird, W. (1986). Interestingness—A neglected variable in discourse processing. *Cognitive Science*, *10*(2), 179–194.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127.

Hidi, S., Renninger, K. A., & Krapp, A. (2004). Interest, a motivational variable that combines affective and cognitive functioning. In D. Y. Dai & R. J. Sternberg (2004). *Motivation, emotion, and cognition: Integrative perspectives on intellectual functioning and development* (pp. 89–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800-802. doi:10.1093/biomet/75.4.800

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1-55. doi:10.1080/10705519909540118

Huber, F., Beckmann, S. C., & Herrmann, A. (2004). Means–end analysis: Does the affective state influence information processing style? *Psychology and Marketing*, *21*(9), 715–737.

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422. doi:10.1037/a0018947

James, W. (1890). *The principles of psychology* (Vol. 1). New York, NY: Henry Holt and Co.

Kaplan, A., Gheen, M., & Midgley, C. (2002). Classroom goal structure and student disruptive behaviour. *British Journal of Educational Psychology*, *72*(2), 191–212. doi:10.1348/000709902158847

Kaplan, A., Lichtinger, E., & Margulis, M. (2011). The situated dynamics of purposes of engagement and self-regulation strategies: A mixed-methods case study of writing. *Teachers College Record*, *113*(2), 284–324.

Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review*, *19*(2), 141–184.

Kaplan, A., & Midgley, C. (1999). The relationship between perceptions of the classroom goal structure and early adolescents' affect in school: The mediating role of coping strategies. *Learning and Individual Differences*, *11*(2), 187–212. doi:10.1016/S1041-6080(00)80005-9

Karabenick, S. A. (1994). Relation of perceived teacher support of student questioning to students' beliefs about teacher attribution for questioning and perceived classroom learning environment. *Learning and Individual Differences*, *5*(2), 187–204.

Karabenick, S. A. (1996). Social influences on metacognition: Effects of colearner questioning on comprehension monitoring. *Journal of Educational Psychology*, *88*(4), 689-703.

Karabenick, S. A. (1998). *Strategic help seeking: Implications for learning and teaching*. Mahwah, NJ: Lawrence Erlbaum Associates.

Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary Educational Psychology*, *28*(1), 37–58. doi:10.1016/S0361-476X(02)00012-7

Karabenick, S. A. (2004). Perceived achievement goal structure and college student help seeking. *Journal of Educational Psychology*, *96*(3), 569–581. doi:10.1037/0022-0663.96.3.569

Karabenick, S. A. & Dembo, M. (2011). Understanding and facilitating self-regulated help seeking. *New directions for teaching and learning*, *126*, 33-43.

Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, *83*(2), 221-230.

Karabenick, S. A., & Newman, R. S. (2006). *Help seeking in academic settings: Goals, groups, and contexts*. Mahwah, NJ: Lawrence Erlbaum Associates.

Karabenick, S. A., & Sharma, R. (1994). Perceived teacher support of student questioning in the college classroom: Its relation to student characteristics and role in the classroom questioning process. *Journal of Educational Psychology*, *86*(1), 90–103. doi:10.1037/0022-0663.86.1.90

Keith, T. Z. (2006). *Multiple regression and beyond*. Boston, MA: Pearson.

King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal*, *29*(2), 303–323.

King, A. (1994). Autonomy and question asking: The role of personal control in guided student-generated questioning. *Learning and Individual Differences*, *6*(2), 163–185.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Koskey, K. L. K., Karabenick, S. A., Woolley, M. E., Bonney, C. R., & Dever, B. V. (2010). Cognitive validity of students' self-reports of classroom mastery goal structure: What students are thinking and why it matters. *Contemporary Educational Psychology*, *35*(4), 254–263.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, *41*(4), 212–218. doi:10.1207/s15430421tip4104_2

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment*, *8*(2), 139–162.

Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. Journal of Research in Science Teaching, 44(5), 706–724. doi:10.1002/tea.20172

Lehnert, W. G. (1978). *The process of question answering: A computer simulation of cognition*. Hillsdale, NJ: Laurence Erlbaum Associates.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood, NJ: Ablex Publishing.

Lent, R. W., Brown, S. D., & Larkin, K. C. (1987). Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. *Journal of Counseling Psychology*, *34*(3), 293–298. doi:10.1037/0022-0167.34.3.293

Lent, R. W., Brown, S. D., Schmidt, J., Brenner, B., Lyons, H., & Treistman, D. (2003). Relation of contextual supports and barriers to choice behavior in engineering majors: Test of alternative social cognitive models. *Journal of Counseling Psychology*, *50*(4), 458-465.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., & Harackiewicz, J. M. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement*, *70*(4), 647–671. doi:10.1177/0013164409355699

Linnenbrink-Garcia, L., Middleton, M. J., Ciani, K. D., Easter, M. A., O'Keefe, P. A., & Zusho, A. (2012). The strength of the relation between performance-approach and performance-avoidance goal orientations: Theoretical, methodological, and instructional implications. *Educational Psychologist*, *47*(4), 281–301. doi:10.1080/00461520.2012.722515

Linnenbrink-Garcia, L., Tyson, D. F., & Patall, E. A. (2008). When are achievement goal orientations beneficial for academic achievement? A closer look at main effects and moderating factors. *Revue Internationale de Pschologie Sociale*, *21*(1/2), 19–70.

Lipstein, R. & Renninger, K. A. (2006). Putting things into words: The development of 12-15-year-old students' interest for writing. In Boscolo, P. & Hidi, S. (Eds.), *Motivation and writing: Research and School Practice* (pp. 113- 140). New York: Kluwer Academic/Plenum.

Lo, Y., Mendell, N. R., & Rubin, D. B., (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767-778.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, *116*(1), 75-98.

Maehr, M. L. (1984). Meaning and motivation: Toward a theory of personal investment. In C. Ames & R. Ames (Eds.), *Research on motivation in education* (Vol. 1), (pp. 115– 144). New York: Academic Press.

Magidson, J., & Vermunt, J. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research, 20*(1), 36–43.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among US students. *Science Education*, *95*(5), 877–907.

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*, 487–503.

Meece, J. L., & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of Educational Psychology*, *85*(4), 582-590.

Michou, A., Mouratidis, A., Lens, W., & Vansteenkiste, M. (2013). Personal and contextual antecedents of achievement goals: Their direct and indirect relations to students' learning strategies. *Learning and Individual Differences*, *23*, 187–194. doi:10.1016/j.lindif.2012.09.005

Middleton, M. J., Kaplan, A., & Midgley, C. (2004). The change in middle school students' achievement goals in mathematics over time. *Social Psychology of Education*, *7*(3), 289–311. doi:10.1023/B:SPOE.0000037484.86850.fa

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor: University of Michigan.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.

Miller, R. G. 1981. *Simultaneous statistical inference* (2nd Ed). New York, NY: Springer-Verlag.

Mills, C. M., Danovitch, J. H., Grant, M. G., & Elashi, F. B. (2012). Little pitchers use their big ears: Preschoolers solve problems by listening to others ask questions. *Child Development*, *83*(2), 568-580. doi:10.1111/j.1467-8624.2011.01725.x

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, *85*(3), 424-436.

Miyake, N., & Norman, D. A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior, 18*, 357-364.

Moos, D. C., & Azevedo, R. (2006). The role of goal structure in undergraduates' use of self-regulatory processes in two hypermedia learning tasks. *Journal of Educational Multimedia and Hypermedia*, *15*(1), 49–86.

Moran, M. D. (2003). Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, *100*(2), 403–405.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30-38.

Murayama, K., & Elliot, A. J. (2009). The joint influence of personal achievement goals and classroom goal structures on achievement-relevant outcomes. *Journal of Educational Psychology*, *101*(2), 432–447. doi:10.1037/a0014221

Murayama, K., Elliot, A. J., & Yamagata, S. (2011). Separation of performance-approach and performance-avoidance achievement goals: A broader analysis. *Journal of Educational Psychology*, *103*(1), 238–256. doi:10.1037/a0021948

Murdock, T. B., Miller, A., & Kohlhardt, J. (2004). Effects of classroom context variables on high school students' judgments of the acceptability and likelihood of cheating. *Journal of Educational Psychology*, *96*(4), 765-777.

Muthén, L.K. & Muthén, B.O. (1998-2012). *Mplus User's Guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Myers, J. L., & Well, A. D. (2002). *Research design and statistical analysis* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, *15*(6), 1044–1045.

National Research Council. (2002). *Improving undergraduate instruction in science, technology, engineering, and mathematics: Report of a workshop*. Washington, DC: National Academies Press.

Nelson-Le Gall, S. (1981). Help-seeking: An understudied problem-solving skill in children. *Developmental Review, 1*(3), 224–246. doi:10.1016/0273-2297(81)90019-8

Nelson Le-Gall, S. (1985). Help-seeking behavior in learning. *Review of Research in Education*, *12*, 55–90.

Newman, R. S. (1990). Children's help-seeking in the classroom: The role of motivational factors and attitudes. *Journal of Educational Psychology*, *82*(1), 71-80.

Newman, R. S. (1991). Goals and self-regulated learning: What motivates children to seek academic help? In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (pp. 151-183). Greenwich, CT: JAI.

Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning* (pp. 283-301). Hillsdale, NJ: Lawrence Erlbaum Associates.

Newman, R. S. (1998). Adaptive help seeking: A role of social interaction in self-regulated learning. In. S. A. Karabenick (Ed.), *Strategic help seeking: Implications for learning and teaching* (pp. 13-37). Mahwah, NJ: Lawrence Erlbaum Associates.

Newman, R. (2000). Social influences on the development of children's adaptive help seeking: The role of parents, teachers, and peers. *Developmental Review*, *20*(3), 350–404. doi:10.1006/drev.1999.0502

Newman, R. S., & Goldin, L. (1990). Children's reluctance to seek help with schoolwork. *Journal of Educational Psychology*, *82*(1), 92-100.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*(3), 328-346.

Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MA: Harvard University Press.

Nolen, S. B. (2003). Learning environment, motivation, and achievement in high school science. *Journal of Research in Science Teaching*, *40*(4), 347–368. doi:10.1002/tea.10080

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction*, *19*(3), 259–271. doi:10.1016/j.learninstruc.2008.05.002

Nylund, K.L., Asparouhov, T. & Muthén, B.O. (2004). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo Simulation study. *Structural Equation Modeling, 14*, 535-569.

Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse processes*, *35*(2), 135–198.

Olney, A. M., Graesser, A. C., & Person, N. K. (2012). Question generation from concept maps. *Dialogue and Discourse*, *3*(2), 75–99.

Otero, J. C., & Campanario, J. M. (1990). Comprehension evaluation and regulation in learning from science texts. *Journal of Research in Science Teaching*, *27*(5), 447–460. doi:10.1002/tea.3660270505

Otero, J., & Graesser, A. C. (2001). PREG: Elements of a model of question asking. *Cognition and instruction*, *19*(2), 143–175.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction*, *1*(2), 117–175.

Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings. *Review of Educational Research*, *66*(4), 543–578. doi:10.3102/00346543066004543

Papert, S. (1980). *Mindstorms: Children, computers and powerful ideas*. New York: Basic Books.

Pearson, J. C., & West, R. (1991). An initial investigation of the effects of gender on student questions in the classroom: Developing a descriptive base. *Communication Education*, *40*(1), 22–32. doi:10.1080/03634529109378823

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. Fort Worth, TX: Harcourt Brace.

Pedrosa de Jesus, H. (2003). Questions of chemistry. *International Journal of Science Education*, *25*(8), 1015–1034.

Perez, A. C. (2012). *Valuing STEM majors: The role of occupational-academic ego identity status and task values in STEM persistence* (Unpublished doctoral dissertation). Temple University, Philadelphia, PA.

Perez, T., Cromley, J. G., & Kaplan, A. (2013). The Role of Identity Development, Values, and Costs in College STEM Retention. *Journal of Educational Psychology*. (Volume and pages not yet assigned). doi:10.1037/a0034027

Piaget, J. (1952). *The origins of intelligence in children.* New York, NY: W.W. Norton & Co.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boedaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation: Theory, research and applications.* San Diego, California: Academic Press.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33-40.

Pomerantz, J. (2005). A linguistic analysis of question taxonomies. *Journal of the American Society for Information Science and Technology*, *56*(7), 715–728. doi:10.1002/asi.20162

Rebok, G. W., & Balcerak, L. J. (1989). Memory self-efficacy and performance differences in young and old adults: The effect of mnemonic training. *Developmental Psychology*, *25*(5), 714.

Renninger, K. A. (1990). Children’s play interests, representation, and activity. In R. Fivush & K. Hudson (Eds.), *Knowing and remembering in young children* (pp. 127–165). New York: Cambridge University Press.

Renninger, K. A. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 373–404). San Diego, CA: Academic Press.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168–184. doi:10.1080/00461520.2011.587723

Renninger, K., Hidi, S. E., & Krapp, A. E. (1992). *The role of interest in learning and development.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution*, *43*(1), 223–225. doi:10.2307/2409177

Robinson, W. P., & Rackstraw, S. J. (1972). *A question of answers*. Routledge & Kegan Paul.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, *66*(2), 181–221. doi:10.3102/00346543066002181

Ryan, A. M., Gheen, M. H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social–emotional role, and the classroom goal structure. *Journal of Educational Psychology*, *90*(3), 528–535. doi:10.1037/0022-0663.90.3.528

Ryan, A. M., Hicks, L., & Midgley, C. (1997). Social goals, academic goals, and avoiding seeking help in the classroom. *The Journal of Early Adolescence*, *17*(2), 152–171.

Ryan, A. M., & Pintrich, P. R. (1997). "Should I ask for help?" The role of motivation and attitudes in adolescents' help seeking in math class. *Journal of educational psychology*, *89*(2), 329-341.

Ryan, A. M., Pintrich, P. R., & Midgley, C. (2001). Avoiding seeking help in the classroom: Who and why? *Educational Psychology Review*, *13*(2), 93–113.

Ryan, A. M., & Shin, H. (2011). Help-seeking tendencies during early adolescence: An examination of motivational correlates and consequences for achievement. *Learning and Instruction*, *21*(2), 247–256. doi:10.1016/j.learninstruc.2010.07.003

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge based questioning by children. *Cognition and Instruction*, *9*(3), 177–199.

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3-4), 299–323. doi:10.1080/00461520.1991.9653136

Schiefele, U. (2009). Situational and individual interest. In K.Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197–222). New York, NY: Routledge.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development*, (pp. 183-212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, *31*(1), 64–96.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading &Writing Quarterly*, *19*(2), 159–172.

Schunk, D. H., & Zimmerman, B. J. (2006). Competence and control beliefs: Distinguishing the means and ends. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (Vol. 2), (pp. 349–367). Mahwah, NJ, US: Lawrence Erlbaum Associates.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464. doi:10.1214/aos/1176344136

Sclove, L. S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, *52*, 333–343. doi: 10.1007/BF02294360

Shernoff, D. J., Csikszentmihalyi, M., Shneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, *18*(2), 158.

Skaalvik, S., & Skaalvik, E. M. (2005). Self-concept, motivational orientation, and help-seeking behavior in mathematics: A study of adults returning to high school. *Social Psychology of Education*, *8*(3), 285–302. doi:10.1007/s11218-005-3276-3

Sternberg, R. J. (1987). Questioning and intelligence. *Questioning Exchange*, *1*(1), 1–14.

Stokal, R. R. (1974). Classification. *Science*, *185*, 115-123.

Susskind, E. (1979). Encouraging teachers to encourage children's curiosity: A pivotal competence. *Journal of Clinical Child Psychology*, *8*(2), 101–106.

Taboada, A., & Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research*, *38*(1), 1–35. doi:10.1207/s15548430jlr3801_1

Tan, S. C., & Seah, L. H. (2011). Exploring relationship between students' questioning behaviors and inquiry tasks in an online forum through analysis of ideational function of questions. *Computers & Education*, *57*(2), 1675–1685. doi:10.1016/j.compedu.2011.03.007

Teixeira-Dias, J., Pedrosa de Jesus, H., Neri de Souza, F., & Watts, M. (2005). Teaching for quality learning in chemistry. *International Journal of Science Education*, *27*(9), 1123–1137. doi:10.1080/09500690500102813

Thorndike, E. L. (1935). *The psychology of wants, interests and attitudes*. Oxford, England: Appleton-Century.

Tuominen-Soini, H., Salmela-Aro, K., & Niemivirta, M. (2011). Stability and change in achievement goal orientations: A person-centered approach. *Contemporary Educational Psychology, 36*(2), 82–100.

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*(2), 438-456.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796. doi:10.3102/0034654308321456

van der Meij, H. (1987). Assumptions of information-seeking questions. *Questioning Exchange, 1*, 111-117.

van der Meij, H. (1988). Constraints on question asking in classrooms. *Journal of Educational Psychology, 80*(3), 401-405.

van der Meij, H. (1990). Question asking: To know that you do not know is not enough. *Journal of Educational Psychology*, *82*(3), 505-512.

van der Meij, H. (1994). Student questioning: A componential analysis. *Learning and Individual Differences*, *6*(2), 137–161.

van Zee, E.H. (2000). Analysis of a student-generated inquiry discussion. *International Journal of Science Education. 22*, 115-142.

VanLehn, K., & Jones, R. M. (1993). What mediates the self-explanation effect? Knowledge gaps schemas or analogies? In M. Polson (Ed.), *Proceedings on the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 1034-1039). Hillside, NJ: Erlbaum.

Watt, H. M. G. (2006). Helen M. G. Watt (2006) The role of motivation in gendered educational and occupational trajectories related to maths, *Educational Research and Evaluation: An International Journal on Theory and Practice, 12*(4), 305-322.

Watts, M., Gould, G., & Alsop, S. (1997). Questions of understanding: Categorising pupils' questions in science. *School Science Review, 79*(286), 57–63.

Wertsch, J. V. (1991). Beyond Vygotsky: Bakhtin's contribution. In J. V. Wertsch, *Voices of the Mind: A socio-cultural approach to mediated action* (pp. 18-45). Cambridge, MA: Harvard University Press.

West, R., & Pearson, J. C. (1994). Antecedent and consequent conditions of student questioning: An analysis of classroom discourse across the university. *Communication Education, 43*(4), 299–311.

White, M. D. (1998). Questions in reference interviews. *Journal of Documentation, 54*(4), 443–465. doi:10.1108/EUM0000000007177

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1–35. doi:10.1016/j.dr.2009.12.001

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. doi:10.1006/ceps.1999.1015

Wilen, W. W., & Clegg Jr., A. A. (1986). Effective questions and questioning: A research review. *Theory & Research in Social Education, 14*(2), 153–161.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation.* (pp. 531–566). San Diego, CA: Academic Press.

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist, 11*(2), 87–95. doi:10.1080/00461527409529129

Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist, 27*(4), 531–541.

Wittrock, M. C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. American Educational Research Journal, 27(3), 489–502.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, *96*(2), 236-250.

Wong, B. Y. L. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, *55*(2), 227–268.

Wood, D., & Wood, H. (1988). Questioning versus student initiative. In J. T. Dillon (Ed.), *Questioning and discussion: A multidisciplinary study* (pp. 280-305). Westport, CT: Ablex Publishing.

Xu, J., & Corno, L. (1998). Case studies of families doing third-grade homework. *Teachers College Record, 100*(2), 402-436.

Yang, M. (2006). A critical review of research on questioning in education: Limitations of its positivistic basis. *Asia Pacific Education Review*, *7*(2), 195–204.

Zientek, L. R., & Thompson, B. (2009). Matrix summaries improve research reports: secondary analyses using published literature. *Educational Researcher, 38*(5), 343–352.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*(3), 329-339.

Zimmerman, B.J. (2000). Attainment of self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. doi:10.3102/0002831207312909

Zimmerman, B. J. & Labuhn, A. S. (2012). Self-regulation of learning: Process approaches to personal development. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook (Vol. 1): Theories, constructs, and critical issues* (pp. 399-425). Washington, DC: American Psychological Association.

Zoller, U. (1987). The fostering of question-asking capability: A meaningful aspect of problem-solving in chemistry. *Journal of Chemical Education*, *64*(6), 510-512.

APPENDIX A

WAVE 1 QUESTIONNAIRE

Question 1
Please enter the 4-digit research study identification number you were given.

Question 2
My chemistry lecture professor wants us to understand our work, not just memorize it.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 3
If I were having difficulty learning the material in my chemistry lecture course I would find someone who could help me.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 4
I am striving to avoid performing worse than others.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 5
I am unlikely to switch to (or select) a non-science major before the end of this semester.

1. Not true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 6
Thinking like a chemist is an important part of who I am.

1. Not true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 7
I would seek help from someone if I could not make sense of something in my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 8

I seek help in my chemistry lecture course when it means I have to work less.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 9
I like chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 10
My aim is to completely master the material presented.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 11
At the present time, how likely are you to remain in (or select) a science major?

1. Not at all likely
2. A little likely
3. Somewhat likely
4. Likely
5. Very likely

Question 12
My aim is to avoid learning less than I possibly could.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 13
My goal is to learn as much as possible.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 14
I would rather do worse on an assignment I couldn't finish than ask for help in my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 15
At the present time, how likely are you to switch to (or select) a major that is not in a science, engineering, technology, or math field (e.g., history, English, art, education, urban studies, Spanish, Psychology, etc.)

1. Not at all likely
2. A little likely
3. Somewhat likely
4. Likely
5. Very likely

Question 16
Even if the work were too hard to do on my own, I wouldn't ask for help with my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 17
I am striving to understand the content of the lectures as thoroughly as possible.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 18
Even if the work in my chemistry lecture course is hard, I can learn it.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 19
If I didn't understand something in my chemistry lecture course I would guess rather than ask someone for assistance.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 20
I would not want anyone to find out that I needed help in my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 21
I would seek help from someone about material in my chemistry lecture course if I couldn't figure it out on my own.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 22
Chemistry is exciting to me.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 23
Getting help would be one of the first things I would do if I were having trouble in my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 24
I enjoy doing chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 25
I am striving to avoid an incomplete understanding.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 26
I would prefer asking another student for help in my chemistry lecture course rather than the instructor.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 27
My goal is to avoid learning less than it is possible to learn.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 28
My goal is to avoid performing poorly compared to others.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 29
If I were to seek help in my chemistry lecture course I would ask the teacher rather than another student.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 30
Getting help in my chemistry lecture course would be an admission that I am just not smart enough to do the work on my own.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 31
I would feel like a failure if I needed help in my chemistry lecture course.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 32
My chemistry lecture professor gives us time to really explore and understand new ideas.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 33
I am likely to switch to (or select) a non-science major at the end of this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 34
I will likely switch to (or select) a major in the social sciences, humanities or arts after this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 35
I enjoy the subject of chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 36
I can do even the hardest work in my chemistry lecture course if I try.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 37
I'm certain I can figure out how to do the most difficult coursework in my chemistry lecture courses.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 38
My aim is to avoid doing worse than other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 39
My chemistry lecture professor recognizes us for trying hard.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 40
My chemistry lecture professor thinks mistakes are okay as long as we are learning.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 41
I am striving to do well compared to other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 42
The purpose of asking somebody for help in my chemistry lecture course would be to succeed without having to work as hard.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 43
It is important to me to be a person who understands chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 44
My aim is to perform well relative to other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 45
I am likely to switch to (or select) a major in the social sciences, humanities or arts before the semester ends.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 46
Chemistry is practical for me to know.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 47
Getting help in my chemistry lecture course would be a way of avoiding doing some of the work.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 48
I'm certain I can master the skills taught in my chemistry lecture course this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 49
I can do almost all the work in my chemistry lecture courses if I don't give up.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 50
My chemistry lecture professor really wants us to enjoy learning new things.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 51
Chemistry helps me in my daily life outside of school.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 52
In my chemistry lecture course, the teacher would be better to get help from than would a student.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 53
My goal is to perform better than the other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 54
If I were having trouble understanding the material in my chemistry lecture course I would ask someone who could help me understand the general ideas.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

APPENDIX B

DEMOGRAPHIC QUESTIONNAIRE

Question 1
Please enter the 4-digit research identification number you were given

Question 2
What is your age?

Question 3
What is your year in school?

- a. Freshman
- b. Sophomore
- c. Junior
- d. Senior
- e. Post-baccalaureate

Question 4
What is your sex?

- Male
- Female

Question 5
What is your race/ethnicity? (Please select all that apply)

- Asian
- Black/African American
- Caucasian/White
- Hispanic/Latino(a)
- Middle Eastern
- Indian subcontinent (India, Pakistan, Bangladesh, Nepal, Bhutan, Sri Lanka, Maldives)

- Indian/Native American peoples
- Other

Question 6
What is the level of education of your mother (or female guardian)?

- Did not graduate from high school
- Graduated from high school
- Some college/Community college
- Bachelor's degree
- Graduate degree (master's, professional, or doctoral)

Question 7
What is the level of education of your father (or male guardian)?

- Did not graduate high school
- Graduated from high school
- Some college/Community college
- Bachelor's degree
- Graduate degree (master's, professional or doctoral)

Question 8
Have you enrolled in CHEM1031 in a previous semester at Temple University?

If no, select "never previously enrolled"?

If yes, select the grade you received?

- ◯ Never previously enrolled
- ◯ A
- ◯ A-
- ◯ B+
- ◯ B
- ◯ B-
- ◯ C+
- ◯ C
- ◯ C-
- ◯ D+
- ◯ D
- ◯ D-
- ◯ F
- ◯ I dropped the course
- ◯ I withdrew from the course

Question 9
Which math courses did you take in high school? (check all that apply)

- ☐ Algebra I
- ☐ Algebra II
- ☐ Geometry
- ☐ Trigonometry and/or Pre-calculus
- ☐ Calculus

Question 10
What is your cumulative GPA for Temple University as of the end of Fall 2012? (Please provide one decimal point if possible; If you have transferred to Temple this semester, please provide your cumulative GPA from your most recent educational institution)

Question 11

What was your score for the Math section of the SAT? (Ranges from 200-800) If you do not recall, please state so.

Question 12

What was your score for the Reading section of the SAT? (Ranges from 200-800) If you do not recall, please state so.

Question 13

What is the highest level of education you aspire to complete (check one):

- Some college
- Bachelor's degree
- Master's degree (MS, MA)
- Professional degree (MD, DPharm, DDS, JD, etc.)
- Doctorate (Ph.D., Ed.D., Psy.D., etc.)

APPENDIX C

WAVE 2 QUESTIONNAIRE

Question 1
Presently, what is your major?

Question 2
Has your major changed over the course of the semester?

- No
- Yes

Question 3
If you changed majors, what was your previous major? (If no change, please enter "NA")

Question 4
Please enter your 9-digit Temple University identification number

Question 5
My chemistry lecture professor wants us to understand our work, not just memorize it.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 6
I am striving to avoid performing worse than others.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 7

I am unlikely to switch to (or select) a non-science major before the end of this semester.

1. Not true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 8

Thinking like a chemist is an important part of who I am.

1. Not true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 9

I like chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 10

My aim is to completely master the material presented.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 11
At the present time, how likely are you to remain in (or select) a science major?

1. Not at all likely
2. A little likely
3. Somewhat likely
4. Likely
5. Very likely

Question 12
My aim is to avoid learning less than I possibly could.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 13
My goal is to learn as much as possible.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 14
At the present time, how likely are you to switch to (or select) a major that is ***NOT*** in a science, engineering, technology, or math field (e.g., history, English, art, education, urban studies, Spanish, Psychology, etc.)

1. Not at all likely
2. A little likely
3. Somewhat likely
4. Likely
5. Very likely

Question 15
I am striving to understand the content of the lectures as thoroughly as possible.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 16
Even if the work in my chemistry lecture course is hard, I can learn it.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 17
Chemistry is exciting to me.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 18
I enjoy doing chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 19
I am striving to avoid an incomplete understanding of chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 20
My goal is to avoid learning less than it is possible to learn.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 21
My goal is to avoid performing poorly compared to others.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 22
My chemistry lecture professor gives us time to really explore and understand new ideas.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 23
I am likely to switch to (or select) a ***non***-science major at the end of this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 24
I will likely switch to (or select) a major in the social sciences, humanities or arts after this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 25
I enjoy the subject of chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 26
I can do even the hardest work in my chemistry lecture course if I try.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 27
I'm certain I can figure out how to do the most difficult coursework in my chemistry lecture courses.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 28
My aim is to avoid doing worse than other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 29
My chemistry lecture professor recognizes us for trying hard.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 30
My chemistry lecture professor thinks mistakes are okay as long as we are learning.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 31
I am striving to do well compared to other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 32
It is important to me to be a person who understands chemistry.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 33
My aim is to perform well relative to other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 34
I am likely to switch to (or select) a major in the social sciences, humanities or arts before the semester ends.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 35
Chemistry is practical for me to know.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 36
I'm certain I can master the skills taught in my chemistry lecture course this semester.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 37
I can do almost all the work in my chemistry lecture courses if I don't give up.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 38
My chemistry lecture professor really wants us to enjoy learning new things.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 39
Chemistry helps me in my daily life outside of school.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

Question 40
My goal is to perform better than the other students.

1. Not at all true
2. A little true
3. Somewhat true
4. Mostly true
5. Very true

APPENDIX D

QUESTION LOG

First Name:

TUID:

During this class period, please do the following:

1. Write down the questions that occur to you about the material being taught.
2. Note the time of each question.
3. Evaluate your question on two dimensions and note whether you asked the question aloud or not
4. Complete the end of class meeting survey questions

	My questions	The answer would be:	Very little (1)	A small amount (2)	Somewhat (3)	Quite a bit (4)	A great deal (5)
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		

Please use reverse side for additional questions

At the end of the class meeting, please complete the following brief survey.

Today, I found today's class interesting	Not true of me	1	2	3	4	5	6	7	8	9	10	Very true of me
Today, I really understood the content of the lecture	Not true of me	1	2	3	4	5	6	7	8	9	10	Very true of me
Today, I am certain that I can master the material from today's lesson	Not true of me	1	2	3	4	5	6	7	8	9	10	Very true of me

Don't Forget! Put an "X" next to any question to which you <u>now</u> know the answer

Please use opposite side first

Record additional questions below

	My questions	The answer would be:	Very little (1)	A small amount (2)	Somewhat (3)	Quite a bit (4)	A great deal (5)
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		
Time		Useful to my learning	1	2	3	4	5
Put "X" here if answered		Useful to others' learning	1	2	3	4	5
		Did you ask it?	No		Yes		

Thank you for your participation