UNIVERSITY OF CALIFORNIA, SAN DIEGO

The influence of prior knowledge, peer review, age, and gender in online philosophy discussions

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Education

in

Teaching and Learning

by

Lucas Stebbins Cuddy

Committee in Charge:

James Levin, Chair
Christopher Halter
Scott Klemmer

2016

The Dissertation of Lucas Stebbins Cuddy is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2016

Table of Contents

List of Tables

List of Figures

Acknowledgements

Thanks to my adviser, Jim Levin. Without his support and guidance on finding a topic, and focusing on that topic, I would have been lost. Thanks also to my other committee members, Scott Klemmer and Chris Halter. Scott's expertise on peer review was invaluable, as was Chris' expertise on statistics. I am also indebted to Alison Wishard Guerra for her additional help with statistical analysis, as well as Amanda Datnow for her reading of an early draft.

Vita

EDUCATION

University of California, San Diego
Doctor of Education in Teaching and Learning, June 2016
Dissertation: The Influence of prior Knowledge, Peer Review, Age and Gender in Online Philosophy Discussions

San Diego State University
Master of Arts in Philosophy, May 2008
Thesis: Video Games as Art: One Interpretation Using the Legend of Zelda as a Case Study

Chico State University
Bachelor of Arts in Philosophy, May 2005

PROFESSIONAL CREDENTIALS/CERTIFICATIONS

San Diego and Imperial Counties Community College Association (SDICCA) Intern Academic Senator for Philosophy/Humanities at Southwestern College, Chula Vista, CA

PROFESSIONAL EDUCATOR EXPERIENCE

2014 – Present: Assistant Professor of Philosophy
Southwestern College, Chula Vista, CA

2007 – 2014: Adjunct Professor of Philosophy
Southwestern College, Chula Vista, CA

2008 – 2014: Adjunct Professor of Philosophy
MiraCosta College, Oceanside, CA

2011 – 2014: Teaching Assistant with Culture, Art, and Technology (CAT) Program
University of California, San Diego

2007 – 2008: Teaching Assistant with the Philosophy Department
San Diego State University

2004 – 2005: Teaching Assistant with the Philosophy Department
Chico State University

PUBLICATIONS

BioShock and Philosophy, Editor, Wiley-Blackwell (June, 2015).

Halo and Philosophy, Editor, Open Court (June, 2011).

“A review of The Art of Videogames by Grant Tavinor” in Journal of Aesthetics and Art Criticism, Vol. 69 # 4 (Fall, 2011).

“A review of Philosophy Through Film by Mary Litch” in American Society for Aesthetics Newsletter, 31(Spring, 2011).

"Is My Kid Sitting Too Close to His Robot?" ed. Sharon Kaye, The Onion and Philosophy, Open Court (December, 2010).

“A review of Thinking Critically by John Chaffee” at Textreviews.com, Cengage (Spring, 2010).

World of Warcraft and Philosophy, Editor (w/ John Nordlinger), Open Court (November, 2009).

The Legend of Zelda and Philosophy, Editor, Open Court (November, 2008).

“The Avatar and the Ego: Reconciling Ourselves with Video Games” eds. Steiff and Tamplin, Battlestar Galactica and Philosophy, Open Court (June, 2008).

“A review of Chinese Philosophy in an Era of Globalization by Robin R. Wang” in Dialogue Volume 50 #1 (October, 2007).

“Could Beatrix Kiddo reach enlightenment? Traces of Buddhist Philosophy in Kill Bill” co-authored with Michael Bruce, eds. Greene and Mohammed, Quentin Tarantino and Philosophy, Open Court (October, 2007).

Abstract of the Dissertation

The influence of prior knowledge, peer review, age, and gender in online philosophy discussions

by

Lucas Stebbins Cuddy

Doctor of Education in Teaching and Learning

University of California, San Diego, 2016

James Levin, Chair

Using a primarily experimental design, this study investigated whether discussion boards in online community college philosophy classes can be designed in the Blackboard course management system to lead to higher order thinking. Discussions were designed using one of two teaching techniques: the activation of prior knowledge or the use of peer review. While these techniques are often identified as components of constructivist pedagogy, the study is grounded theoretically in epistemological holism. The literature calls for deeper research on prior knowledge/peer review and higher order thinking, and the relationship between them. Four eight week classes were studied, using two specific discussions from each class (one on Rene Descartes' certainty, and the other on Aristotle's eudemonia), the experimental and control groups switching from one discussion to the next. In the experimental groups, the prior knowledge technique was

implemented before the discussions in the form of multiple choice questions and open ended questions; peer review was implemented via the Blackboard peer review interface, drawing from category coding from the literature. Control groups were given standard discussion prompts. Higher order thinking was assessed from multiple choice quiz question responses as well as instructor scoring (found to be reliable) of student discussion board prompts. All quantitative data was analyzed using quantitative analysis software. My primary hypothesis, that there would be more evidence of higher order thinking in experimental groups, was not substantiated by the data. However, considering that the overall level of higher order thinking of students in the study was high, this lack of substantiation may have something to do with the ceiling effect. Also, there were statistically significant correlations between demographic variables (specifically gender and age) and higher order thinking. These correlations in some cases are supported by existing literature.

Chapter 1: Introduction

A report on online learning (Allen & Seaman, 2013) notes that 70% of chief academic officers see online learning as critical for their institution's long-term strategy. The report also notes that the belief that is difficult to retain students online has grown to 44.6% (Allen & Seaman, 2013). Considering this data, it stands to reason that adding to our growing understanding of online learning is worthwhile, and may show us new ways to increase student retention online.

While many areas of online learning have been studied, there are some areas for which more research would be beneficial. Meyer (2003) notes that much work remains to be done to enhance our understanding of online discussion boards, particularly in relation to higher order thinking. Song and McNary (2011) note that little is known about the relationship between social interaction online and learning. Rovai (2004) specifically calls for further research on online learning regarding the relationship between constructivist pedagogy and higher order thinking skills. While the focus of my study is not constructivism per say, and later I argue for grounding my study in epistemological holism (Sher, 2015), I do hone in on specific techniques that are often identified as components of constructivist pedagogy: prior knowledge and peer review.

Prior knowledge as a pedagogical technique suggests constructing an association between new knowledge and prior/existing knowledge (Pritchard, Lee, & Bau, 2008), referred to in this dissertation as *the activation of prior knowledge*. In their review of 183 publications discussing prior knowledge, Dochy, Segers, and Buehl (1999) identify six common ways prior knowledge has been operationalized. The authors found that not all of these ways are externally valid. However, I am interested in two that *are* externally

valid: open questions and multiple choice questions. Peer review as a pedagogical technique refers to the commenting of students on each other's written work (Cathey, 2007; Kulkarni, Socher, Bernstein, & Klemmer, 2014), though I will draw a more complex version of peer review from this literature. I am interested in the use of these techniques in community college philosophy discussion board assignments to support the learning of higher order thinking. Higher order thinking will be measured by student understanding of specific philosophical arguments via multiple choice questions and discussion responses, drawing from Lewis and Smith (1993) and Meyer (2003).

This study has an experimental design. It investigated whether the activation of prior knowledge or, potentially, the use of peer review (the independent variables) were more or less likely to lead to higher order thinking in relation to particular philosophical arguments (the dependent variable): Aristotle's argument for eudemonia (happiness) and Rene Descartes' argument for certainty. Four student demographic variables were also statistically analyzed in relation to higher order thinking: age, gender, interest in philosophy, and academic goal.

I said "potentially" above regarding peer review since whether or not the peer review experiment was carried out depended on whether the first experiment regarding prior knowledge showed a statistically significant effect. If the effect was significant, then the study would remain focused on prior knowledge and a new study using the most successful technique would be carried out. However, if there was not a significant effect, then the study would shift focus to peer review. Based on the lack of statistical significance in the first experiment, the study did, in fact, shift focus to peer review. The details are discussed in depth in the methodology chapter.

My research questions are:

- Can activating prior knowledge in online discussion boards in introductory community college philosophy classes lead to higher order thinking?
    - *Subquestion*: Can activating prior knowledge (via pre-discussion multiple choice questions) in online discussion boards in introductory community college philosophy classes lead to higher order reasoning (in relation to Aristotle's eudemonia)?
    - *Subquestion 2*: Can activating prior knowledge (via an open-ended discussion question) in online discussion boards in introductory community college philosophy classes lead to higher order reasoning (in relation to Descartes' certainty)?

- Can using a peer review technique in online discussions boards in introductory community college philosophy classes lead to higher order thinking (either relating to Aristotle's eudemonia or Descartes' certainty)?

This structure of this dissertation is as follows. In Chapter 2, I review the literature on online learning generally, then literature related to my use and operationalization of higher order reasoning, prior knowledge, and peer review. Then in Chapter 3, I discuss the study design in detail. The theoretical framework, as noted, is epistemological holism (which I compare and contrast with constructivism and post positivism). I also restate the research questions and explain their implications in more detail. Data was collected in the form of written student responses from online discussion boards as well as multiple choice quiz answers. Data was analyzed quantitatively for relationships between variables. Student responses on discussion boards were scored, checked for interrater reliability, and found to be reliable. I go on to discuss my positionality (my potential bias as a researcher) and potential threats to validity. In Chapter 4, I discuss my findings. No statistically significant relationships were found between the teaching techniques and higher order thinking. However, there were statistically significant relationships between both age and gender and higher order

thinking. Specifically, women engaged in more higher order thinking than men on Aristotle multiple choice questions. And older students engaged in more higher order thinking than younger students generally. In Chapter 5, I discuss the implications of these findings. First, the lack of statistical significance between the techniques and higher order thinking might be explained by the ceiling effect, since a significant portion of students did engage in higher order thinking. Second, some research (Machado, 2011) suggests that women tend to be more inhibited than men on online discussions. This research, along with the fact that the Aristotle questions were arguably easier, may help to explain my result that women engaged in more higher order thinking on Aristotle multiple choice questions. Third, some research (Huschle, 2013) suggests that older, experienced students contribute greatly to the effectiveness of online discussions. This research may help to explain my result that older students engaged in more higher thinking generally.

Chapter 2: Literature Review

Before addressing higher order thinking, prior knowledge, and peer review, I review the general research on online discussions to see what is lacking. Then, I elaborate the theoretical framework, epistemological holism. I discuss its relation to post positivism and constructivism.

Wijekumar and Spielvogel (2006) created what they called "intelligent discussion boards" that used an intelligent tutoring system. They applied the system with two undergraduate courses, using quantitative and qualitative analysis to conclude that those using their system tended to create more posts relevant to the discussion. While the authors target problems that I feel are relevant and important, solutions already exist that address some of these problems without having to employ an entirely new program. For example, one problem they note is that students tend to paraphrase each other's posts rather than create a unique contribution to the discussion. But a recent feature in the course management system Blackboard allows the instructor to change a setting so students must post first before they can even access the discussion. The authors also note that further research with larger numbers of students, and in multiple domains, is necessary.

Song and McNary (2011) analyzed students' interaction patterns on online discussion boards, using descriptive and regression analyses to draw their conclusions. They found a relationship between course design and student interaction on the discussions. Additionally, they found that social interactions generally, among other things, determine the quality of online learning, but little is known about the relationship between specific communications and learning. Finally, they argue from their results that

it is most likely the quality of posts, rather than the number of posts, that predicts a student's grades.

Machado (2011) coded over a thousand discussion board posts and weblogs to discover gender-related patterns. She admits her findings are not generalizable, but, among other things, she found that there was greater variability in frequency of posts on discussion boards among women than among men. She calls for more research on gender in relation to online learning. As noted in the previous chapter, my results regarding gender may support her findings.

Kay (2006) argues that our understanding of online discussions as a tool for learning could be improved by creating a theory-driven metric to determine effectiveness. She created such a metric, then applied it by conducting a study at a secondary school with teenagers. I mention Kay's research to show the way my approach differs. My approach is less comprehensive, and focuses simply on whether two techniques (activating prior knowledge and the use of peer review) may or may not lead to higher order thinking. This is not to diminish the significance of Kay's project; in fact, the results of my research may contribute to understanding the effectiveness of parts of Kay's metric, such as that related to knowledge level.

However, some of Kay's (2006) results should be mentioned. Her results suggest that the discussion question's quality did not have an impact on how many students responded to it coherently. This counts as some evidence against scaffolding discussion questions for prior knowledge. But in the study, most questions (95%) were created by students, so it is possible that if there were more opportunities to respond to a sophisticated, scaffolded question from an instructor, then the results would have been

different. Additionally, her study was on teenagers and I am studying adult community college students. Also, as discussed below, Sautter's (2007) research suggests that how a discussion prompt is a structured is directly relevant to the development of higher order thinking on that discussion.

It is worth noting one other area of focus in the research on online discussions: the instructor's role. Gerber, Scott, and Clements (2005) argue that the instructor should tailor the level of abstractness of his responses to particular students; they also found that, especially with less abstract topics, when an instructor takes a more challenging role in his responses it is associated with higher student engagement. As noted above, her results are limited in generalizability, but Kay (2006) did find that an instructor with too strong of a presence in the discussion can shut down student engagement. Matheson, Wilkinson, and Gilhooly (2012) argue that an instructor should help students to ask the right questions. Song and McNary (2011) stress the importance of instructor guidance to keep students on track, but also the relationship between course design and online instructor interaction with students. However, Huschle (2013) reminds us that, in addition to the instructor, mature and experienced students play a big a role in an effective online discussion. In Chapter 3, I discuss the implications this research has regarding limitations to my study.

Having discussed some of the general literature on online discussion boards and learning, I now turn specifically to research on higher order thinking.

Higher Order Thinking

I begin this section with research explicitly attempting to find relationships online between prior knowledge and/or peer review and higher order thinking. Then, I move into more specific definitions of higher order thinking.

Land and Dornisch (2002), in their study on online discussions, collected data by frequency counts of student posts and examples of posts with evidence of reflection in a class on web development. There were 35 graduate and upper level undergraduate students involved (ten males and 25 females) over the course of two semesters. They coded student discussion responses for four levels of reflection. Level one reflection entails little or no critical analysis of a students' own or other students' posts. Level two reflection entails recognition of another perspective's difference from one's own, but no analysis of that difference. Level three reflection entails recognition of another perspective's difference from one's own, elaborated on with personal experiences, alternative perspectives, suggestions, or examples. Level four reflection entails reflection on how a new perspective extends or refines an original idea. In the second semester, they found that students whose prior knowledge was elicited were more likely to engage in level four reflection. However, they did not elicit prior knowledge, but simply noted whether a student referred to prior knowledge in his or her post. I elaborate below on the way this discussion of reflection ties into higher order thinking.

Matheson, Wilkinson, and Gilhooly (2012) wondered whether combining patchwork text and discussion boards as a method of assessment could assist in promoting critical thinking and collaborative working among students. They mention that they are interested in higher order thinking as well, but they do not distinguish between it

and critical thinking. Below, this distinction is made and clarified. Like Meyer (2003), whom I discuss below, Matheson, Wilkinson, and Gilhooly (2012) adapted a model for evaluating student discuss responses from Garrison, Anderson, and Archer (2001). Based on qualitative coding of discussions with this model, Matheson, Wilkinson, and Gilhooly (2012) were successful in promoting collaboration among students and inducing critical thinking. I cite their study to provide one example of the way this model has been used.

I now turn to definitions of and discussions of higher order thinking online independent from prior knowledge or peer review.

As noted above, Machado (2011) discusses gender differences in student posts on discussion boards. However, what is relevant to this section of my dissertation is her discussion of different types of reflection, which she draws from Schon (1987). There is reflection-in-action, which refers to contemporaneous thought, or the process of modifying one's thinking as it occurs; there is reflection-on-action, which refers to retrospective thought, or learning from past experiences; finally there is reflection-for-action, which refers to anticipatory thought, or anticipating the effect one's thought might have on others, the class itself, and the community as a whole. However, she notes that there is little research actually documenting whether any of these types of reflection can be developed as a result of scaffolding online discussions appropriately. We will see later how these categories may or may not apply to higher order thinking.

Although Sautter's (2007) study was conducted on college marketing students in both online and face to face settings, given his focus on higher order thinking some of his conclusions are worth discussing. Specifically, the author notes that structured discussion matters most when the goal is higher order thinking, but structure matters less when the

goal is creativity or openness. Considering that the dependent variable of my study is higher order thinking, this suggests the discuss prompts should as structured as possible in the experimental condition.

I now turn to discussions of higher order thinking independent of online learning *and* prior knowledge or peer review.

Ennis (1993) discusses critical thinking in relation to Benjamin Bloom's taxonomy (represented in Figure 1). Ennis argues that while analysis, evaluation, and creation (which sit at the highest levels of Bloom's taxonomy) are somewhat relevant to critical thinking, they fall short of establishing a coherent definition. However, since we are after higher order thinking, and not critical thinking, I leave this as a starting point. Resnick (1987) discusses higher order thinking and lower order thinking, arguing that the two types of thinking work together, and should not really be separated. Like Ennis, Resnick gives us a starting point. However, the most comprehensive analysis of higher order thinking comes from Lewis and Smith (1993).

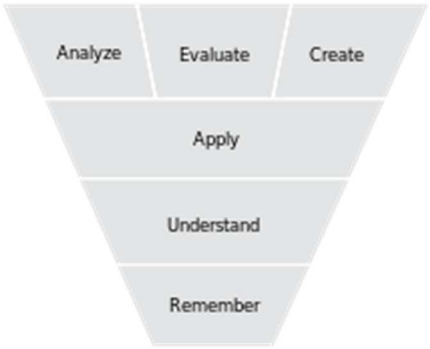

Figure 1. Bloom's Taxonomy.

Lewis and Smith (1993) first attempt to disambiguate higher order thinking, and to separate it from other terms like critical thinking. They investigate the matter in three

directions. First, they discuss higher order thinking as it has been used by philosophers and psychologists. Philosophers, they argue, drawing from the example of Socrates (Plato, 1956), tend to see higher order thinking as being very systematic and disciplined, a means of guarding against fallacious arguments and societal corruption. Psychologists, on the other hand, are more concerned with the thinking process itself, and tend to emphasize problem solving rather than systematic reflection. However, Lewis and Smith argue that while both the philosophical and psychological approach to higher order thinking are necessary to the concept, they are not sufficient.

The second aspect of their investigation attempts to distinguish higher order from lower order thinking skills. Differences in students' prior knowledge might make it so that one needs higher order thinking, and one lower order thinking, to solve the same problem. Sometimes this is referred to as the relativity of higher order thinking, and it suggests that some information about a student's prior knowledge is necessary to determine whether higher order thinking is happening. Lewis and Smith (1993) distinguish between learned behavior (lower order thinking) and reasoning (higher order thinking), drawing from Maier (1933, 1937). On this view, learned behavior would be learning multiplication tables through repeated practice, whereas reasoning would be applying knowledge of multiplication to a unique problem the learner has not yet faced. The authors also discuss Bartlett's (1958) view, who introduces the concept of *gap filling*, actively filling a gap in missing knowledge, to characterize higher order thinking. Lewis and Smith also discuss thinking as hierarchical, with higher order thinking being the crown of the hierarchy.

The third aspect of Lewis and Smith's (1993) investigation into higher order thinking focuses on critical thinking and problem solving. The authors note that critical thinking has been defined in multiple ways, including as problem solving, evaluation, and a combination of the two. However, sometimes critical thinking is differentiated from problem solving. Distinct from problem solving, critical thinking can be seen as determining the accuracy or authenticity of knowledge, such as accepting or rejecting a statement (that is, a form of evaluation). Another view of critical thinking the authors discuss is not just evaluation, but also the construction of arguments or hypotheses.

Lewis and Smith (1993) argue that higher order thinking must involve both evaluation and problem solving. They discuss a hypothetical doctor who is trained in problem solving, but who may be confronted by ethical issues that go beyond problem solving, and require critical thinking. The authors argue that both of these, problem solving and critical thinking, are aspects of higher order thinking. They propose the following:

> Higher order thinking occurs when a person takes new information and information stored in memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers in perplexing situations (p. 136).

They give examples that fit this definition, including deciding what to believe, making a prediction, or solving a nonroutine problem.

In my research, this is the definition of higher order thinking I employ. As noted, I measure higher order thinking in two ways: multiple choice answers and discussion board responses. Answering a multiple choice question correctly involves making a prediction based on previously learned information. Having the ability to accurately elaborate a

sophisticated philosophical concept in one's own words on a discussion board also suggests prediction, but perhaps also the solving of a nonroutine problem.

This above research also suggests that students with greater prior knowledge of a given concept will be more likely to engage in higher order thinking about that topic. A prediction of my research, therefore, is that students whose activated prior knowledge is richer in relation to the topic at hand will be more likely to indicate higher order thinking. I discuss this prediction in Chapter 5.

However, while multiple choice question responses are easy enough to assess, the assessment of students' written discussion board responses is less easy. As noted, in my study student discussion responses were scored, checked for interrater reliability, and found to be reliable. This scoring was grounded in Meyer's (2003) research, who adapted a model from Garrison, Anderson, and Archer (2001). Meyer coded online discussion boards specifically for evidence of higher order thinking. She used coding to place student discussion responses into one of four categories: triggering (posing a problem), exploration (searching for information), integration (constructing a possible solution), resolution (critically assessing the solution), and social (basically an "other" category). Unlike Meyer, however, I modified the categories to reflect Lewis and Smith's (1993) work above. My categories are: understanding of the philosophical argument, predicting the implications of the argument, and connecting the argument accurately to one's existing beliefs on the matter. Although Meyer found some evidence of higher order thinking in her study, since her coding is different, her results are less relevant to my study than her methods.

See Appendix 1 for the specific way I operationalized these categories for student peer review of each other's discussion posts and for interrater reliability scoring. In the next chapter, I provide examples of student discussion responses that illustrate, and fail to illustrate, higher order thinking according to the categories

Prior Knowledge

Because research on prior knowledge in relation to online discussion boards is incomplete (Kay, 2006; Rovai, 2004; Song and McNary, 2011), in this subsection I will draw on research from both prior knowledge per se, and prior knowledge studied in relation to online learning (in some cases this research focuses on discussions specifically, in other cases it focuses on online learning generally).

Dochy, Segers, and Buehl (1999) define prior knowledge as the entirety of a person's actual knowledge that "(a) is available before a certain learning task, (b) is structured in schemata, (c) is declarative and procedural, (d) is partly explicit and partly tacit, (e) and is dynamic in nature and stored in the knowledge base" (p. 146). While one might fault this definition for being circular, it gives us a starting point. The authors conducted a review of the different ways prior knowledge has been assessed and measured by different researchers, with a focus on external validity (that is, experiments with larger sample sizes that are non-artificial and lengthy). They found that prior knowledge can be validly measured using standard measures, like multiple choice tests, open questions, cloze tests, completion tests, recognition tests, and matching tests.

Dochy, Segers, and Buehl (1999) also mention some characteristics of prior knowledge, some of which seem to be problematic for learning. On the positive side,

prior knowledge can give students clarity in studying, such as leading them to find material faster while searching for it in a book. However, some educators falsely assume that students' prior knowledge is generally accurate, when in fact often students' prior knowledge is riddled with misconceptions (and the literature on cognitive biases today adds fuel to this fire). In fact, prior knowledge can sometimes hinder learning if a student is unwilling to let go of a misconception that has been activated via some pedagogical technique. Still, students with no prior knowledge of a subject are handicapped in that area.

Again, Pritchard, Lee, and Bau (2008) see prior knowledge as a pedagogical technique as referring to constructing an association between new knowledge and prior/existing knowledge (the activation of prior knowledge). Rovai (2004) discusses the implications for constructivism within a learning environment specifically, such as customizing curricula to students' prior knowledge as a means of promoting dialogue. For example, rather than ask students to read the same paper, a teacher might ask students to read different papers that they relate to personally, but all with a similar theme.

So we see two trends in the literature: assessing and defining existing prior knowledge, and the activation of prior knowledge. As noted, I am interested in the latter, but as we will see there is a connection between existing prior knowledge and activating it.

Pritchard, Lee, and Bau (2008) discuss previous research showing that prior knowledge played an insignificant role in students' test score gains in Newtonian mechanics. The authors question this result, wondering why students' test scores did not increase with increased prior knowledge. However, they note that different learning

theories suggest that the rate of learning will have different dependencies on a given student's prior knowledge. They conducted their own study in which they created mathematical learning models based on various learning theories, one of them being constructivism. Since constructivism suggests that students gain knowledge by constructing an association between it and existing knowledge, this implies that, "the more prior knowledge one has, the faster learning will occur and, conversely, that if the relevant knowledge needed for the construction is unknown, then learning will be slowed," (p. 2). However, this does not seem to take into account misconceptions noted by Dochy, Segers, and Buehl (1999) above. A student with more prior knowledge may be weighed down by misconceptions in that prior knowledge. In fact, this may explain the results of the research they cite showing that prior knowledge played an insignificant role in test score gains.

Alvermann, Smith, and Readence (1985) conducted a study in which they activated prior knowledge in middle school children via multiple choice questions and free response questions. They were wondering what happens when a student encounters prior knowledge that runs counter to a text that contains true information. They found support for the notion that prior knowledge may interfere with reading comprehension under some conditions. Not surprisingly, when students activate a belief they hold strongly, even true information is not likely to override that belief. The authors distinguish usefully between existing prior knowledge and activated prior knowledge.

This study attempted to activate prior knowledge in students with multiple choice questions and open ended questions (see the specific questions I used in Appendices II

and III). This type of prior knowledge activation is supported by Dochy, Segers, and Buehl (1999) as well as Alvermann, Smith, and Readence (1985).

It stands to reason that activating prior knowledge can both reveal existing prior knowledge as well as spur students to use that newly surfaced prior knowledge. This study spurred students' prior knowledge in an attempt to lead them to higher order thinking. Some of the issues from the literature, such as the idea that prior knowledge can in some cases be a hindrance, are discussion in Chapter 5. For a more detailed discussion of how prior knowledge was implemented in my study, see Chapter 3.

Peer Review

Peer review seems to have a more intuitive definition than does prior knowledge. After all, we are familiar with peer review in the scientific process, specifically the blind review process in journal submissions. The idea behind a blind review is to eliminate bias and to see whether a submission has merit that speaks for itself. A commonsense definition of peer review, then, is when peers check over each other's work for mistakes, usually anonymously.

Rovai (2004) discusses peer review/discussion extensively in relation to the implications for constructivism within a learning environment. While some of us may have memories of peer review gone wrong when we were teaching assistants—such as differential student effort—the literature suggests that there are legitimate uses of peer review, especially within an online environment.

Although Cathey, Panza, and Potthast (2006) do not provide much information on their methodology, they argue that online discussion can help enhance peer-to-peer

engagement, specifically in philosophy classes. They note that, “failing to critically engage with others dooms us to self-deception in our reasoning, as false beliefs remain active, and perhaps even multiply, until they are challenged” (Cathey, et. al., 2006, p. 281).

Kulkarni, Socher, Bernstein, and Klemmer (2014) attempt to address some of the problems with peer review of short answer questions in an online environment, one of the main problems being that students spend too much time grading. Basically the authors created an algorithm that predicts the student grade and estimates the confidence level, which in turn estimates the number of peer reviewers required. Students go on to identify features of the answer using a rubric, other students determining whether these labels are accurate. They call this the *identify-verify approach*. The results were 80-90% more accuracy in grading with only 54% of the grading time. Although my classes are smaller than the large-scale classes these authors discuss, they make general points about peer review that are relevant to my research.

Cathey (2007) used a semester long peer review assignment in an undergraduate sociology course. Although the course was face to face, the peer review was done on an online discussion board. From the beginning students were split into groups. When an assignment was due, each member of one group had to post their paper online, and each member of the other groups had to comment, where each commenter was assigned to a specific paper. When a new assignment was due, another group would post their papers, and the other groups would comment, and so forth. Students were instructed to comment systematically, first justifying whether the first paragraph provided a clear description, then the second paragraph, and so forth, then justifying whether the paragraphs related

sufficiently to each other. Cathey assigned a grade to each commentary, but she did not allow the commentaries to influence her grading of the assignments themselves. She graded the commentaries on the extent to which they matched her own perceived effort. In the end, although students reported putting in more work, the results did not suggest an increase in writing skill or student understanding.

In an article written about Scott Klemmer's work on student peer review, Fox (2013) lists seven habits for effective peer review that Klemmer identifies. The habits are assignment-specific rubrics and training, iteration before release (pre and during), assessing self after peers, using staff grades as "ground truth", aggregating grades adaptively, using cued prompts to offer written feedback, and closing the loop; giving assessors' feedback.

Some of these habits will be implemented in my study and others will not. My study will be conducted on the class management system Blackboard, and will therefore be limited by the technological capabilities of the program. The study will also be limited to what I am able to implement and accomplish considering the general time constraints of completing a dissertation. While I will be providing students with something similar to assignment-specific rubrics that draw from the literature on higher order thinking (see Appendix I) when they peer review, I will not have the time or resources to put students through the training period Klemmer recommends. Because the classes with which I will be working are small and short in duration, I will also not be able to go through the iterations or to aggregate grades adaptively (this latter habit seems more relevant to larger online classes anyway). Peers will assess their own answers after assessing other students' answers. Staff grades will be used as "ground truth", or as the default correct

answers. However, the "staff" in the peer review portion of the study will only be myself (though recall that my scoring was found to be reliable). My study will include cued prompts to offer written feedback and it will provide assessors' feedback.

For a more detailed discussion of how peer review will be implemented in my study, see Chapter 3.

Theoretical Framework: Holism

Considering that this study has an experimental design, this might suggest that the study is grounded in post positivism, the view that there is an objective truth to be known and it is most likely to be known by eliminating confounding variables and isolating an experimental treatment. While this study is post positivist in this sense, it also makes use of constructivist pedagogy. However I hesitate to say that it is grounded in constructivism, since some constructivists take the view that there is no objective truth to be known (Rovai, 2004)—my experimental treatments can be said to be more or less effective in developing higher order thinking, objectively. On the contrary, there *is* an objective truth to be known, in my view. However, Rovai (2004) does note that not all constructivists deny objective truth. Still, it is my view that the name constructivism itself is misleading in this regard.

Rovai (2004) argues for a more pragmatic view of constructivism in which knowledge is objective but just happens to be the product of social processes. He calls constructivist views that suggest there is no objective reality unfairly restrictive. But in one sense, one might say that such views *are* constructivist at the core, and that the

essence of constructivism, indeed an important way it differs from other epistemological theories, is a lack of objective truth. After all, it is in the name: we *construct* reality.

We can scale back this argument against Rovai a bit to make a practical argument. We can say that constructivism, regardless of its core nature, is known to most as a theory that denies an objective truth. I refer back to Mertens (2010) who notes that constructivists see research as being not independent of human minds. Even Rovai (2004) suggests that this is a viable interpretation of constructivism.

The theory of epistemological holism (Sher, 2015), unlike constructivism, suggests that there is a whole, that there is something beyond human minds. Let us unpack holism by seeing an interpretation of it from one of its most vocal proponents. Sher (2015) aims to develop a substantive, holistic, theory of truth that "addresses itself to humans' desire to know and understand the world in its full complexity on the one hand and to their intricate yet limited cognitive resources on the other" (p. 1). She defends this holistic view of truth, and consequently knowledge. Holism can be contrasted with foundationalism, which suggests that truth must have a solid foundation (such as logic, or scientific facts). Sher does not fully reject foundationalism, since she believes truth has some foundation, but she believes foundationalism is too strict and, therefore, self-defeating, since the base itself cannot be justified in principle. As she puts it: "We arrive at our holistic methodology by relaxing the strict ordering requirement of the foundationalist methodology" (p. 5).

Sher believes that the foundational project can be carried out in many ways, referring to her project at one point as *foundational holism*. She lists some advantages to pursuing this route to truth, though I mention here only the advantages relevant to this

dissertation. One advantage holism affords is that we can use resources from any domain of knowledge to help solve a problem. Another advantage is that there are multiple routes to reality, including circuitous ones. Another advantage is that justifying claims happens in steps, and partial justification can be still be worthwhile. Since we want to know the world in its full complexity, Sher argues, "this requires stretching our cognitive endowments, devising multiple means for reaching its less than accessible regions, improvising, experimenting, tinkering, exercising our imagination" (p. 6). What this implies for truth is that, "it's seriously possible that there are multiples routes of correspondence between true cognitions and reality, and that some of these routes are quite intricate" (p. 6).

It is important to note that Sher's project is one nested tightly within epistemology in the philosophical tradition. She is responding to and arguing against many other theories of truth in philosophy, constructivism among them, and I cannot hope to do justice to the nuances of her arguments here. However, her overall framework of holism is what I think can best ground my research, rather than constructivism.

While constructivism captures the intuition that Socrates had when using his Socratic method, and that seminar leading professors have when they see knowledge being critiqued and evaluated socially, it does not capture the intuition that there *are* truths to be known beyond the social, that there *is* an objective reality beyond human minds. If teachers were constructivists to the core, then there would be no right or wrong answers to their assignments, no better or worse essays, just constructed realities. The entire practice of grading is an argument against constructivism. Holism in a sense embraces constructivism, in that it acknowledges multiple paths to truth (one of them

might be social construction), but it nevertheless says there *is* a truth beyond minds and there *is* a reality beyond minds and there *is* knowledge beyond minds. Holism is more in line with our intuitions as teachers.

Truth be told, prior knowledge, peer review, and higher order thinking may not look too different through the lens of holism than constructivism. The main difference is in holism's emphasis on an actual truth, rather than a strictly socially constructed one. Thus, rather than let constructivism ground my research, I will ground it in holism. After all, I am using an experimental method to figure out the answer to whether online discussion boards can be designed to lead to higher order thinking. Presumably, higher order thinking happens or it does not, it is not something that can be reduced to social construction alone.

I now turn to methodology.

Chapter 3: Methodology

The purpose of this study is to determine whether the use of two popular teaching techniques (the activation of prior knowledge and peer review) on online philosophy discussion boards can lead to higher order thinking. Specifically, I ran a series of controlled experiments in online philosophy classes over the 2015 summer and fall semesters to investigate this issue. I used qualitative scoring of discussion board posts to complement the quantitative analysis.

In this chapter I discuss the participants and setting of this study. Then, I discuss data collection and analysis. Next, I restate my research questions and discuss them in light of the preceding material. Finally, I discuss my positionality (my potential bias as a researcher) and deal with threats to validity.

Participants and Setting

The participants involved in this study were community college students enrolled in an online introductory philosophy class at East Hills Community College (pseudonym) located in an urban city in Southern California. The classes were fully online and conducted via the course management system, Blackboard. There were 36 students in the study.

Data Collection and Analysis

The study used an experimental design using four community college philosophy online classes that was implemented through the course management system, Blackboard.

I solicited student participation by offering them a small percentage (1.5%) of extra credit toward their grades, as well as a ten dollar gift card, if they participated. Students who did not want to participate were offered the same percentage of extra credit toward their grades for completing a different task, as well as the gift card. Students formally gave their consent by filling out an electronic consent form (see Appendix IV). There were three experiments in the study, outlined below. To ensure random assignment in the first study, students who participated from either of the two classes involved were randomly assigned to experimental and control discussion boards external to their respective class shells in Blackboard. However, in the second and third studies, due to a change in my teaching schedule, students were drawn only from a single class (one class per study) to be randomly assigned to experimental and control discussion boards external to those Blackboard class shells (one external shell per study).

Although the discussion boards were external, the link appeared within the class and linked to another Blackboard class shell that looked the same as the original class shell. So, although students were technically leaving the current class shell, from their perspective it may not have felt like they left at all.

Detailed descriptions of the experiments are below, but Figure 2 gives a concise overview of the designs.

Table 1. Study Design Overview.

Studies/Experiments	Experimental	Control
Experiment 1a (Aristotle)	Activation of prior knowledge prompt using multiple choice questions (Group 1)	Standard prompt (Group 2)
Experiment 1b (Descartes)	Activation of prior knowledge prompt using an open-ended question (Group 2)	Standard prompt (Group 1)
Experiment 2a (Aristotle)	Activation of prior knowledge prompt using multiple choice questions (Group 3)	Standard prompt (Group 4)
Experiment 2b (Descartes)	Activation of prior knowledge using an open-ended question (Group 4)	Standard prompt (Group 3)
Experiment 3a (Aristotle)	Peer review with a standard prompt (Group 5)	Standard prompt only (Group 6)
Experiment 3b (Descartes)	Peer review with a standard prompt (Group 6)	Standard prompt only (Group 5)

Brief Overview of Experiment 1: The independent variables for the first experiment in the study were the activation of prior knowledge in a discussion prompt (experimental group) and a standard discussion prompt (control group). Two ways of activating prior knowledge were tested in different randomly-assigned groups: a series of multiple choice questions and an open ended question (see Appendices II and III). The first group began as the experimental group and ended as the control group, and vice versa for the second group.

The dependent variable was higher order thinking in relation to particular philosophical arguments—Aristotle's argument for eudemonia and Descartes' argument

for certainty. As discussed in chapter 2, this understanding was measured by the scoring of student discussions as well as a set of post-discussion multiple choice questions (see Appendices I and III). The scoring was checked for interrater reliability; I sought two people with philosophy degrees as well as philosophy teaching experience to assist in this matter. Both philosophers were sent five anonymous student discussion board posts as well as guidelines for scoring them (see Appendix 1). Once they sent me their scores, I checked them for interrater reliability. The scoring was found to be reliable ($\alpha = .962$).

Experiment 1a: The first group (in this case the experimental group) was encouraged to activate prior knowledge through the use of multiple choice questions prior to a standard discussion prompt on Aristotle's eudemonia. The second group was given only a standard discussion prompt on Aristotle's eudemonia (in this case the control group).

Experiment 1b: The first group (in this case the control group) was given only a standard discussion prompt on Descartes' certainty. The second group (in this case the experimental group) was encouraged to activate prior knowledge through the use of an open-ended question on a discussion prompt on Descartes' certainty.

Hypothesis for Experiment 1: Given that access to prior knowledge should positively affect the rate of learning (Pritchard, Lee, & Bao, 2008), my hypothesis was that students whose prior knowledge have been activated (in the experimental groups) are more likely to engage in higher order thinking in relation to the philosophical arguments.

Experiment 2 was exactly the same as experiment 1, except that the former drew from only one online fall class, whereas the latter drew from two simultaneous summer

classes. The groups used in this class (who, as in experiment 1, alternated between experimental and control groups) are referred to as the third and fourth groups.

Brief Overview of Experiment 3: Like eliciting prior knowledge, peer review is often identified as a component of constructivist pedagogy. This experiment investigated whether the use of student peer review (the independent variable) was more or less likely to lead to higher order thinking in relation to the philosophical arguments (the dependent variable).

Experiment 3a: The fifth group (in this case the experimental group) was given a standard discussion prompt on Aristotle's eudemonia; members of the group were asked to peer review each other's responses to the discussion. The sixth group (in this case the control group) was given only a standard discussion prompt.

Experiment 3b: The fifth group (in this case the control group) was given only a standard discussion prompt. The sixth group (in this case the experimental group) was given a standard discussion prompt on Descartes' certainty; members of the group were asked to peer review each other's responses to the discussion.

Hypothesis for Experiment 3: Considering the already successful role of peer review in some forms of online learning (Kulkarni, Socher, Bernstein, & Klemmer, 2014), my hypothesis was that students who used peer review (in the experimental groups) would be more likely to engage in higher order thinking in relation to the philosophical arguments.

Once again, the data was collected in the form of student multiple choice question responses and student discussion boards posts. However it should be noted that the only data in experiment 3 is the multiple question responses; this is because the peer review,

the intervention, happened after the students' initial posts on the discussion boards. Once the student discussion board posts being used for the study were scored according to the criteria in the literature review, and checked for interrater reliability (see Appendix I), the data was analyzed quantitatively along with the multiple choice responses. The data was analyzed using the statistical analysis software PSPP as well as SPSS.

It will be useful here to include examples of both a discussion post that was scored as indicating higher order thinking, and one that was not. Recall that the three categories are understanding of the philosophical argument, predicting the implications of the argument, and connecting the argument accurately to one's existing beliefs on the matter (see Appendix 1 for more details). Students could get nine total points in their scores, three for each category. If they got at least seven out of nine, they were said to be engaged in higher order thinking. Before giving examples I should note that, regarding the multiple choice questions, students were said to be engaged in higher order thinking only if they answered all the multiple choice questions correctly. Now, I will start with the example of a student *not* engaged in higher order thinking on the discussion board. This student was responding to the Descartes experimental discussion prompt (see Appendix 2).

> In my Arts class when I was younger, we were taught that there were three primary colors: red, yellow and blue. In order to get the secondary colors,purple, green and orange, we would need to mix two primary colors. For violet, mix blue and red; green is a mixture of blue and yellow; while orange is yellow plus red. I was certain that these secondary colors are really made up of two of the primary colors because first, the teacher told us it was, second it was written in the books, and lastly we actually tried and mixed the colors.
>
> Certainty is being sure about something. Sometimes it is hard to be certain while it is easy to doubt the things we see, learn, and read about. Being

> 100% sure or certain is possible, though. Depending on the issues you are dealing with, it may be possible not to doubt it. For instance, we need oxygen to live. But I wouldn't stake my life that everything I know is certain. I have a curious mind and doubt many things as well. Based on Descartes' argument about certainty, he believes that if we take out all the doubts we have, what remains should be "certainty". I believe that Descartes would agree to me that a mixture of two primary colors will create a secondary color. In addition, according to him, to perceive it as true means it is also "clear and distinct." Because I believe that blue mixed with yellow makes green, and I did mix those colors which produced the result I perceived, then my claim is clear and distinct. Descartes' "I think, therefore I am" and as a thinking human being, therefore my knowledge about colors is certain.

Although this student articulates her points well and makes some valid connections to her prior knowledge, she is not engaged in higher order thinking primarily because she misunderstands Descartes' crucial point: his argument is not that thinking makes us human but that thinking is the primary piece of evidence we have to know that we exist as something. Specifically, this student got a two on understanding the argument, because she does accurately discuss a couple of Descartes' early premises. But she got a one out of three on predicting the implications of the argument, since she never accurately reaches the end of the argument to begin with. And she got a one out of three on connecting the argument to her beliefs because Descartes would not agree with her views of certainty—she does not accurately connect her views to Descartes' argument.

Here is an example of a student who *is* engaged in higher order thinking on the discussion board. This student was responding to the Aristotle control discussion prompt (see Appendix 2).

> Aristotle's argument for eudaimonia is the realization of one's full potential. This principle goes beyond our standard definition of happiness, enjoyment, and pleasure. We often label a person's innate talents and abilities as their "calling" – something they are almost predestined to

> do. This could include a person "called" to serve others like Mother Teresa, a phenomenal athlete like LeBron James, a prolific composer like Wolfgang Amadeus Mozart, or more common people like a Philosophy professor who is fulfilled by enriching young minds, an accountant who can almost hear numbers sing to her, or a mother who can't imagine doing anything other than loving, raising, and nurturing her children. But realizing and living out your calling is only one aspect of it. To realize complete eudaimonia, one must live a full, self-sufficient life, unrestricted by mental, physical, spiritual, or financial issues – or at least they stand a better chance without these obstacles.
>
> At the restaurant Carlos is living one small aspect of eudaimonia, in that he is surrounded by friends and is enjoying social interaction. He appears "happy" in that moment, but we really don't have enough information about Carols to know his level of eudaimonia. Having a bad day on the job isn't enough to rule out the possibility that he is achieving the full potential of his life, but having a good evening out with friends isn't enough to confirm that he is either.
>
> I think Aristotle would like to talk to Carols to learn more about him, his background, his hobbies, his career, his financial and spiritual status, his lifestyle and living situation, and his passions, and would then determine to what extent he has been able to detect and realize the full potential of his existence, in order to achieve eudaimonia.

First of all, the student reveals her full understanding of the argument in the first paragraph (three out of three), as well as her ability to predict the implications in specific individuals and their callings in life (three out of three). Regarding connections, the student understands that Carlos being apparently happy at one event is not even enough to determine if he is experiencing eudemonia. Aristotle, the student says, would want to know if Carlos has truly reached his full potential. These are decent connections, but Aristotle would probably have enough info to say that Carlos has a serious obstacle to eudemonia if he is not happy with his job (two out of three). Overall, the student got an eight out of nine, clearly illustrating higher order thinking according to the categories.

Now, to see exactly how the data helped to answer my research questions, let us take a look at the questions once more.

Restatement of Research Questions

- Can activating prior knowledge in online discussion boards in introductory community college philosophy classes lead to higher order thinking?
    - *Subquestion*: Can activating prior knowledge (via pre-discussion multiple choice questions) in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Aristotle's eudemonia)?
    - *Subquestion 2*: Can activating prior knowledge (via an open-ended discussion question) in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Descartes' certainty)?
- Can using a peer review technique in online discussions boards in introductory community college philosophy classes lead to higher order thinking (either relating to Aristotle's eudemonia or Descartes' certainty)?

How did my study answer these questions? I will address my findings in Chapter 4, but let us first consider what it would take to support my hypotheses.

Beginning with experiments 1 and 2, if a quantitative analysis showed that the experimental groups' (whose prior knowledge was activated via pre-discussion multiple choice questions or an open ended question) performance on the post-discussion multiple choice questions was significantly better than that of the control groups, and the student discussion board scores were higher, then it would be reasonable to answer the first subquestion above with a "yes". These results would also provide some evidence to answer "yes" to the more general research question: can activating prior knowledge in online discussion boards in introductory community college philosophy classes lead to higher order thinking?

Regarding experiment 3, if a quantitative analysis showed that the experimental groups' (who used a peer review technique) performance on the post-discussion multiple choice questions is significantly better than that of the control groups, then it would be reasonable to answer my final research question with a "yes". Again, note that experiment 3 did not involve scoring of student discussion posts. This is because students posted their first discussion posts *before* having peer reviewed each other's posts, so their first discussion posts could not indicate whether peer review led to higher order thinking.

Validity

I begin this section with a discussion of my positionality (or potential bias as a researcher), then I move on to validity per say, particularly in relation to the quantitative aspect of my study.

As noted in Chapter 2, one area where my positionality might influence the study is in the instructor's role. One of the most relevant points from the research is that an instructor who takes a more challenging role in his responses to students on online discussions can enhance student engagement (Gerber, Scott, & Clements, 2005). However, if the instructor has *too* strong of a presence, this can shut down student engagement (Kay, 2006). Some argue that an instructor's role is to help students ask the right questions (Matheson, Wilkinson, & Gilhooly, 2012). Others stress the relationship between instructor's role and course design (Song & McNary, 2011). Considering that I taught the classes as I ran the experiments for this study, this research on the instructor's role in online discussions suggests that I should have been cognizant of my interaction patterns with students on the discussions. Moreover, considering that this study is largely

quantitative, my role as professor could potentially have been a confounding variable. I therefore made a conscious effort to not be too challenging with students on the discussions, striving to be as consistent as possible across classes. That is, I made an extra effort to pay special attention to devoting an equal amount of time to either discussion (both experimental and control). If I had more interactions on, say, the experimental discussions than the control discussions, then that could have confounded the results as well.

Like all of us, I must beware of the confirmation bias, the tendency to confirm one's existing beliefs regardless of evidence and arguments. Of course on some level I *want* my hypotheses to be true. However, this bias is mitigated in part by the nature of the study design being quantitative. No matter how strong my confirmation bias is, there is hard statistical data that provides an indication of whether or not my research questions were answered. There was also the concern that my bias would surface when assessing student discussion responses for evidence of higher order thinking. But this concern was mitigated by interrater reliability.

Mertens (2010) discusses several threats to validity with experimental and/or quasi-experimental designs. She distinguishes between threats to internal and external validity. I now discuss some of these threats and what I did to combat them, beginning with internal validity—internal validity being when changes to the dependent variable are due to the independent variable, rather than confounding, or unintended, variables.

One threat to internal validity is history, which refers to outside events that happen throughout the course of a study that can influence results, since it might be the outside event that led to any changes observed in the dependent variable rather than the

experimental treatment. Another threat is maturation: any changes that participants go through during the study (like students getting bored or tired over time) might also serve as confounding variables. However, these two threats can be controlled by the use of a control group, provided that people in both groups experience the same history and/or maturation. Considering that the students in the control and experimental group pairs in my study were part of the same or similar online classes, it stands to reason that they experienced the same or similar history and/or maturation. Mertens (2010) discusses the threats of statistical regression (using extreme groups as participants) and differential selection (when people with different characteristics are in experimental and control groups). However, these two threats, as Mertens herself notes, can be combated by random assignment of participants to control and experimental groups, and random assignment was a feature of my study design. Mertens discusses the threat of testing, which arises when a pre-test in a study leads students to be sensitized to what kind of information to expect on the post-test. However, this is not an issue for my study since in the one experiment that involved a pre and post-test, the tests were different (the first activates prior knowledge and the second assesses higher order thinking). Related to this last threat is instrumentation, which happens when there is a change in the test taking instrument from pre to post (one may be easier than the other). But again, the pre and post-tests in my study measured different things, so it is hard to say that one is "easier" or "harder" than the other. Finally, Mertens discusses the threat of experimental mortality, which happens when more participants drop out of the experimental than control group (or vice versa) over the course of the study. As Mertens notes, theoretically random assignment combats this threat since one would expect drop-outs to equalize over time

between experimental and control groups. While I paid special attention for this threat throughout my study, there was only one student who dropped out of the study. Since my study involved students in a class, participants were likely more motivated to complete the study as a partial motivation for completing the class.

I now turn to Mertens' (2010) list of threats to external validity—external validity being the extent to which a study's results can be generalized to other people and/or situations. The first threat to external validity is an insufficient description of the independent variable. However, I made an effort to describe the details of the independent variables in my study and how they draw from the literature. In the appendices are examples of the ways I operationalized the independent variables. Another threat is multiple treatment interference, but since in each experiment of my study the students are receiving only one major treatment (activation of prior knowledge or peer review), this threat does not seem to be an issue. The next threat is the Hawthorne effect, which is the idea that getting singled out to participate in a study alone, rather than the experimental treatment, is enough to influence the dependent variable. Related to this threat is novelty, another threat which suggests that the experimental treatment might produce positive results only due to the fact that it is novel. However, these last two threats are mitigated by the fact that my study is with full, existing classes, meaning that students were less likely to feel that they were "singled out" since they were already part of the class. Also, the experimental treatment was less likely to seem novel to students since it took place as part of the curriculum for the class that they were already in. Moreover, the experimental treatments themselves were partially built into the class structure so this also likely made the treatments seem less novel. The experimenter effect

is a threat which suggests that the results of an experiment will not generalize to other situations without the initial experimenter. This threat was not likely a problem in my study since, as discussed earlier, I strove to make my role a consistent one on the discussion boards. Furthermore, it is the experimental treatments themselves that mattered most, not the instructor's implementation of them. Other threats are pre-test and/or post-test sensitization, the idea that taking either of these tests may sensitize students to the treatment. However, the pre-tests at least seem to have been less of an issue for my study since the pre-tests were actually part of the treatment. Post-test sensitization was mitigated by the fact that all students, in both control and experimental groups, took the post-test as it was a part of the general curriculum. And the post-test included questions that are *not* part of the study as well, making it more likely that the students did not know which questions were part of the study when they took the test. Another threat is the interaction of history and treatment effects, which happens when the contextual factors of one study cannot be duplicated in another. However, since my study was of an online class in Blackboard, it can easily be duplicated by another instructor who has knowledge of Blackboard. Another threat is the measurement of the dependent variable—if the dependent variable is measured in one way in one study, it may be measured in another way in another study and have different effects. However, I combated this threat by using two different measurements of the dependent variable (higher order thinking) in the first place, which drew from literature where the variable was operationalized in similar ways. Finally, there is the threat of the interaction of time and the treatment: if a post-test is given, say, one day after the treatment rather than one week, there may be different results. I admit that I could not fully combat this threat, but I

will note that, as a part of an existing class, the implementation of the treatments was timely in that they fit coherently with the structure of the rest of the class.

Besides threats to internal and external validity, there is an ethical concern with experimental studies in that the control group does not get a potentially beneficial treatment, while the experimental group does. However, my study is partly designed to explicitly deal with this ethical concern: in each experiment, the experimental and control groups were switched halfway through, meaning that all participants in the study had the chance to experience an experimental treatment.

One final threat to validity (or we might call it a limitation), is the sample size of participants. The study included 35 participants. This smaller sample size limits the generalizability of the study to other situations. For a generalizable study of this kind, there would need to be approximately 100 participants (this sample estimate is based on the population size, variance, and a few other factors).

However, in Chapter 5 my conclusions are stated in proportion to the sample size; my conclusions are considered alongside conclusions that other researchers have reached in this domain. This study identifies correlations that may be useful for future researchers to investigate in more depth.

Chapter 4: Findings

Existing research within the field of online learning suggests a relationship between the activation of prior knowledge and higher order thinking (Rovai, 2004), as well as between the use of peer review and higher order thinking (Kulkarni, Socher, Bernstein, & Klemmer, 2014). However, there is not much research on the relation between these variables in *online* contexts like discussion boards (Meyer, 2003). And there is even less research regarding the two teaching techniques and online *philosophy* discussion boards.

This study built on the prior research while attempting to break new ground regarding the two techniques and online philosophy discussion boards. Specifically, this study investigated whether two teaching techniques, the activation of prior knowledge and the use of peer review on online philosophy discussion boards, were more or less likely to lead to higher order thinking.

The first section of this chapter discusses the demographic data of the students involved in the study. The second section discusses the findings in relation to research question 1 (can activating prior knowledge in online philosophy discussion boards lead to higher order thinking?), and the third discusses findings in relation to research question 2 (can using a peer review technique in online philosophy discussion boards lead to higher order thinking?). Research questions are restated in sections below respectively.

Unfortunately, the general findings do not indicate statistical significance between the teaching techniques and higher order thinking. So this study cannot conclusively answer the research questions posed. However, as noted previously, since a large proportion of students in the study did engage in higher order thinking, this general lack

of statistical significance may be due to the ceiling effect and not due to the lack of effectiveness of the teaching techniques in engendering higher order thinking. Additionally, the results of the study did show some interesting trends between the demographic data (gender and age) and higher order thinking. I discuss these trends in the fourth section of this chapter.

Demographics of Student Participants

The demographic data is divided based on the three experiments conducted in the study. Table 2 shows the data for experiment 1. The student with the incomplete (a student who dropped out of the class before completing the study) was naturally not included in the analysis. Students 1 through 9 comprised group 1, which, like all the groups, alternated between being the experimental and control. Students 10 through 17 comprised group 2. (See Table 1 in the previous chapter for an overview of the groups and the associated experiments.)

Table 2. Experiment 1 Student Demographics.

	Gender	Age	Academic Goal	Interest in Philosophy
Student 1	Female	18 to 24	AA or BA	Kind of
Student 2	*Incomplete*	*Incomplete*	*Incomplete*	*Incomplete*
Student 3	Female	Over 24	AA or BA	Kind of
Student 4	Male	Over 24	AA or BA	Kind of
Student 5	Female	Over 24	AA or BA	Kind of
Student 6	Male	Over 24	AA or BA	Very
Student 7	Female	Over 24	Not sure	Very
Student 8	Female	18 to 24	AA or BA	Very
Student 9	Female	Over 24	AA or BA	Kind of
Student 10	Male	Over 24	MA or PhD	Very
Student 11	Female	18 to 14	AA or BA	Kind of
Student 12	Female	Over 24	AA or BA	Very
Student 13	Female	Over 24	AA or BA	Kind of
Student 14	Female	Over 24	MA or PhD	Kind of
Student 15	Female	18 to 14	AA or BA	Not very
Student 16	Male	18 to 14	MA or PhD	Very
Student 17	Male	Over 24	MA or PhD	Very

Table 3 shows the data for experiment 2. Students 18 through 22 comprised group 3. Students 23 through 27 comprised group 4.

Table 3. Experiment 2 Student Demographics.

	Gender	Age	Academic Goal	Interest in Philosophy
Student 18	Female	18 to 24	N/A	N/A
Student 19	Male	18 to 24	N/A	N/A
Student 20	Male	18 to 24	N/A	N/A
Student 21	Male	18 to 24	N/A	N/A
Student 22	Female	18 to 24	N/A	N/A
Student 23	Female	18 to 24	N/A	N/A
Student 24	Male	18 to 24	N/A	N/A
Student 25	Male	Over 24	N/A	N/A
Student 26	Male	18 to 24	N/A	N/A
Student 27	Female	Over 24	N/A	N/A

Table 4 shows the data for experiment 3. Students 28 through 32 comprised group 5. Students 33 through 36 comprised group 6.

Table 4. Experiment 3 Student Demographics.

	Gender	Age	Academic Goal	Interest in Philosophy
Student 28	Female	18 to 24	MA or PhD	Very
Student 29	Female	18 to 24	MA or PhD	Kind of
Student 30	Female	Over 24	AA or BA	Not very
Student 31	Female	18 to 24	AA or BA	Kind of
Student 32	Female	Over 24	AA or BA	Very
Student 33	N/A	N/A	N/A	N/A
Student 34	N/A	N/A	N/A	N/A
Student 35	Male	Over 24	MA or PhD	Very
Student 36	Female	Over 24	AA or BA	Kind of

Findings Related to Research Question 1

The findings related to research question 1 come from experiments 1 and 2. Once again, the following is the first research question, including subquestions.

- Can activating prior knowledge in online discussion boards in introductory community college philosophy classes lead to higher order thinking?
    - *Subquestion*: Can activating prior knowledge via pre-discussion multiple choice questions in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Aristotle's eudemonia)?
    - *Subquestion 2*: Can activating prior knowledge via an open-ended discussion question in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Descartes' certainty)?

Recall that both subquestions 1 and 2 targeted two different methods of activating prior knowledge to determine whether one might be more or less likely to lead to higher

order thinking. Beginning with subquestion 1, the data shows that the correlation is positive, r(11) = .53, but not significant between the activation of prior knowledge via pre-discussion multiple choice questions and students' higher order thinking $F(1, 21) = .017$, $p = .896$.

Regarding subquestion 2, the data shows that the correlation is negative, r(11) = -.18, but not significant between the activation of prior knowledge via open-ended questions and students' higher order thinking $F(1, 23) = .865$, $p = .362$.

So overall, the answer to research question 1 is that more research is still needed. I speculate more on the implications in Chapter 5.

Findings Related to Research Question 2

The findings related to research question 2 come from experiment 3. Once again, the following is the second research question.

- Can using a peer review technique in online discussions boards in introductory community college philosophy classes lead to higher order thinking (either relating to Aristotle's eudemonia or Descartes' certainty)?

Regarding the Aristotle peer review group, the correlation is negative but not significant, r(6) = -.18, between the peer review technique and students' higher order thinking $F(1, 23) = .236$, $p = .632$. Regarding the Descartes peer review group, the correlation is also negative, r(6) = -.21, but not significant between the peer review technique and students' higher order thinking $F(1, 24) = 1.697$, $p = .205$. Overall, the peer review portion of this study involved the least amount of participants. This is unfortunate, because with a larger sample these trends might be worthy of being paid more attention.

As with the first, more research is still needed to determine the answer to this research question. I discuss the implications of this result in Chapter 5.

Influence of Demographics on Higher Order Thinking

As noted, I compiled some demographic data from students. Unfortunately, some of the surveys were not complete so I do not have complete data. But, there are some interesting trends when the demographic data are statistically analyzed with the higher order thinking scores (independent of whether participants were in the experimental or control groups). In some cases, although there was no statistical significance overall, the effect sizes (or impacts) were large even within the small sample.

Gender did have an effect on higher order thinking scores on the Aristotle prompt. Recall that in this study higher order thinking was measured by both multiple choice question responses and discussion board scoring. The effect between gender and discussion board higher order thinking scores was not significant, but the effect between gender and multiple choice higher order thinking scores was significant within the small sample, $F(1, 14) = 7.14$, $p = .018$. Specifically female participants ($M = .74$, $SD = .15$) demonstrated more higher order thinking on the multiple choice questions on Aristotle than male participants ($M = .52$, $SD = .18$). Additionally, there was a confidence interval range of 4.07 to 5.45 and the effect size within the small sample was large ($n^2 = .34$). I speculate more on these results in Chapter 5.

Age did have an effect on higher order thinking scores on both the Aristotle and Descartes prompts, but only on discussion board higher order thinking scores (not on the multiple choice). The effect between age and discussion board higher order thinking

scores on Aristotle was significant within the small sample, $F(2,31) = 3.36$, $p = .048$. The effect between age and discussion board higher order thinking scores on Descartes was significant within the small sample, $F(2, 32) = 3.46$, $p = .044$. Specifically, students who were over 24 years of age ($M = 250.22$, $SD = 446.48$) demonstrated more higher order thinking on the Aristotle discussion board than students between the ages of 18 and 24 ($M = 187.89$, $SD = 402.42$). Students who were over 24 years of age ($M = 235.49$, $SD = 436.55$) also demonstrated more higher thinking on the Descartes discussion board than students between the ages of 18 and 24 ($M = 187.74$, $SD = 402.5$). Additionally, the effect size within the small sample was large in both cases. There was a confidence interval range of 4.07 to 5.45 on the Aristotle discussion and the effect size was large ($n^2 = .18$). There was a confidence interval range of 4.97 to 6.60 on the Descartes discussion and the effect size was large ($n^2 = .18$). I discuss the implications of these results in Chapter 5.

Academic goal and interest in philosophy both had no effect on higher order thinking scores.

Chapter 5: Reflection, Implications, and Conclusion

Existing literature calls for more research on the relationship between two teaching techniques (the activation of prior knowledge and the use of peer review) and higher order thinking, specifically with online learning and even more specifically with online learning of philosophy. Grounded theoretically in epistemological holism, this study investigated whether the two aforementioned techniques can lead to higher order thinking on online philosophy discussion boards. After scoring student discussion boards and multiple choice responses, no statistically significant relationship was found between these techniques and higher order thinking. However, this lack of significance may be indicative not of the failure of these techniques, but of the ceiling effect. There were statistically significant correlations between two student demographic variables (gender and age) and higher order thinking.

In this chapter, I first reflect on the study findings and their meaning in relation to existing literature. Then, I discuss the implications of this study for both research and teaching.

Reflection on Findings

In this section I first restate my research questions and review the study design. Then I briefly review my findings. I reflect on each finding as I go along, tying it in with existing literature and research when possible. I also attempt an explanation for each finding.

Once again, my research questions are:

- Can activating prior knowledge in online discussion boards in introductory community college philosophy classes lead to higher order thinking?
    - *Subquestion 1*: Can activating prior knowledge via pre-discussion multiple choice questions in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Aristotle's eudemonia)?
    - *Subquestion 2*: Can activating prior knowledge via an open-ended discussion question in online discussion boards in introductory community college philosophy classes lead to higher order thinking (in relation to Descartes' certainty)?
- Can using a peer review technique in online discussions boards in introductory community college philosophy classes lead to higher order thinking (either relating to Aristotle's eudemonia or Descartes' certainty)?

These questions were addressed using an experimental design. The independent variable in each experiment was one of the teaching techniques, either the activation of prior knowledge or the use of peer review. The dependent variable was higher order thinking. There were three experiments conducted with four different online community college philosophy classes. The first two experiments focused on prior knowledge, and the last on peer review. Students were given a small amount of extra credit for participating. Groups were alternated, so all students experienced some sort of intervention. All experiments drew from discussions boards whose content was either Aristotle's argument for happiness (eudemonia) or Descartes' argument for certainty (his famous "I think, therefore I am" argument). The experimental groups were focused on prior knowledge or peer review. Prior knowledge was activated either via open ended questions or multiple choice questions, drawing from Dochy, Segers, and Buehl (1999) as well as Alvermann, Smith, and Readence (1985). Peer review was conducted in the peer review system in the Blackboard Learn course management system (Blackboard, 2016), the implementation drawing from Fox (2013) and from Kulkarni, Socher, Bernstein, and Klemmer (2014). Control groups were given standards prompts.

Theoretically, my claims about higher order thinking are grounded in Lewis and Smith's (1993) definition, which is as follows:

> Higher order thinking occurs when a person takes new information and information stored in memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers in perplexing situations (p. 136).

Drawing from this definition as well as from Meyer's (2003) research and Garrison, Anderson, and Archer's (2001) research, I used multiple choice questions and the scoring of discussion boards to measure higher order thinking. Discussion board scoring was checked for interrater reliability and found to be reliable. Once collected, the data was analyzed statistically for relationships between the dependent and independent variables. Student demographic data were also analyzed. The demographic variables are academic goal, interest in philosophy, age, and gender. See the appendices for the details on the implementation, including the specific discussion prompts, multiple choice questions, and the details of how discussion boards were scored.

Before conducting the experiments, my hypothesis, regarding both primary research questions, was that the respective teaching techniques (prior knowledge and peer review) would lead to higher order thinking. The findings indicate no statistically significant relationship between the techniques and higher order thinking. However, of the 35 students involved in the study, 25 of them engaged in higher order thinking generally (on the multiple choice questions, on the discussion boards, or both). So it is possible that these results are explained by the ceiling effect.

The ceiling effect can have different meanings depending on the field of study. But I mean it in the following sense. The ceiling effect describes a situation where the

dependent variable is no longer effected by the independent variable. It suggests that above a certain level, a treatment or intervention (the independent variable) is no longer effective. A medical example is often used to illustrate the effect: you can increase the dosage to your heart's desire with some pain relief medication, but there is a point at which the increased dosage becomes ineffective.

Again, the independent variables in my study were the teaching techniques and the dependent variable was higher order thinking. Considering that the level of higher order thinking for all participants was already relatively high (25 out of 35), this might suggest that the techniques were simply not going to push students higher in a way that was statistically significant. Thus, it may be that the techniques are, in fact, effective (as prior literature suggests), but that they did not reveal their effectiveness in this study since most students were already engaging in higher order thinking to begin with. The techniques cannot fix what was not broken. On this view, it is no wonder that the results did not indicate statistical significance.

I now turn to the influence of student demographics on higher order thinking. As noted, the demographic variables I collected were academic goal, interest in philosophy, age, and gender. There were no statistically significant relationships between either academic goal or interest in philosophy and higher order thinking. However, there were statistically significant relationships between age and gender and higher order thinking. These latter relationships were not just statistically significant, but the effects also had a large impact. Therefore, it is worth discussing these demographic results in more depth.

Regarding age, older students (students over the age of 24) tended to engage in more higher order thinking. To some, this finding might simply verify what we know

through common sense: maybe older students are better prepared than younger students. But the finding also coincides with the literature. Huschle (2013), for example, found that mature and experienced students contribute greatly to the effectiveness of online discussion. Additionally, this finding corresponds with relatively recent neuroscientific evidence regarding how the brain develops. Consider a study by Larsen, Hartmann, and Nyborg (2008). The authors point out that early work on cognition suggested that cognitive ability declines in young adulthood and into adulthood. Consequent research suggested that cognitive ability does not decline as one gets older, but rather stabilizes. However, the authors themselves found reasons to doubt these results, arguing that their research suggests that cognitive ability actually improves in several ways in young adulthood through adulthood. The authors' research, therefore, coincides with my finding that older students engaged in more higher order thinking than younger students. Higher order thinking, after all, is an example of one important cognitive ability.

Another interesting possibility relates to prior knowledge. Recall Lewis and Smith's (1993) research, which suggests that students with greater prior knowledge of a given concept will be more likely to engage in higher order thinking about that topic. Do older people have better, more accurate, prior knowledge than younger people, generally speaking? At first this possibility might not seem to be the case: after all, global warming deniers can be younger college students or older pillars of the community. However, if people's cognitive ability improves as they get older, then we might expect an older person's general prior knowledge to, in fact, be more accurate than a younger person's. Because my study does not provide data that would clearly shed light on this issue, I speculate more in the section below on the implications for future research.

Regarding gender, my findings are more nuanced. Gender had an effect on higher order thinking scores, but only on the Aristotle discussion prompt. The effect between gender and discussion board higher order thinking scores was not significant, but the effect between gender and multiple choice higher order thinking scores was significant. Specifically, females tended to do better (illustrate more higher order thinking) on the Aristotle multiple choice questions than males. How can this finding be interpreted? Machado (2011) found some evidence that females tend to be more inhibited than males on online discussions. Machado's finding may therefore help to explain why female students did better on the multiple choice questions: they felt less inhibited than they do on the online discussions. Their inhibitions were removed, and they were able to focus more on the content of the questions. Or at least this is one possibility. The reason why females in my study did not also do well on the Descartes discussion might be explained by the fact that Descartes' argument for certainty is harder and more abstract than Aristotle's argument for eudemonia. But this analysis still does not explain why men, if they were less inhibited, did not do well on the Aristotle discussion board responses. Thus, more research is needed here, as I discuss below.

There are other ideas from the literature that could be discussed. One is the relativity of higher order thinking. The relativity of higher order thinking is the idea that whether a student engages in higher order thinking will depend on her prior knowledge of the subject at hand. Another idea from these authors is that prior knowledge can in some cases inhibit learning, if a student's prior knowledge is itself flawed. But I also discuss these ideas in the sections below since the data in my study do not directly shed any light on them.

Implications for Further Research

The primary limitation of this study is its small sample size. When planning the study, I had expected to get more participants, but by chance the last two classes I pulled from did not generate as many participants as the first two classes. So perhaps the best way to improve this study would be to conduct it again with more participants. This improvement is particularly relevant to peer review, as I conducted only one study on peer review and it included a total of only nine participants. Recall Klemmer's seven habits for effective peer review (Fox, 2013): assignment-specific rubrics and training, iteration before release (pre and during), assessing self after peers, using staff grades as "ground truth", aggregating grades adaptively, using cued prompts to offer written feedback, and closing the loop; giving assessors' feedback. Also recall that, due to the limitations of the Blackboard course management system, I was only able to implement peer assessment, the use of staff grades as "ground truth", cued prompts, and the use of assessors' feedback. Future research could improve this study further by using all seven of Klemmer's habits for effective peer review.

There are additional implications for research. Recall that I activated prior knowledge on the online discussion boards in one case with an open ended question, and in another case with multiple choice questions. Regarding the open ended question, it may be that the control group's question was too similar to the experimental group's question. Here is the experimental open ended question for the Aristotle discussion prompt:

> Think back to a time in your life when you were happy. What were you happy about? Why were you happy? If you could define happiness, how would you do it?

> Now that you have addressed these questions, how do you think Aristotle would react? Given Aristotle's argument for eudemonia from the chapter, how is your idea of happiness similar to or different from Aristotle's eudemonia? What *is* Aristotle's argument for eudemonia?

And here is the control open ended question for the Aristotle discussion prompt:

> Imagine that a person, Carlos, is out at a restaurant hanging out with his friends. Carlos is socializing, laughing, and joking—he appears to be having a great time. One of Carlos' friends, Tom, remembers that yesterday Carlos was complaining about how he didn't like his job. At dinner, Tom asks Carlos how he feels, and Carlos answers that he feels totally happy.
>
> Now that you have addressed these questions, how do you think Aristotle would react? Given Aristotle's argument for eudemonia from the chapter, how is Carlos' apparent happiness similar to or different from eudemonia? What *is* Aristotle's argument for eudemonia?

I address only the Aristotle prompt for brevity, but the arguments I am making here apply equally to the Descartes prompt. See the appendices for all prompts. Initially, the difference between the prompts was supposed to be that the experimental prompt targets the student's own thoughts about happiness, bringing to the surface his prior knowledge on the subject. The control prompt, on the other hand, points to someone else being happy. However, I see now how a student completing the control prompt might also be drawn to their prior knowledge on happiness, even if the question does not target it directly. For example, as a student reads about Carlos, she may be thinking of how her views of happiness do or do not coincide with Carlos' experience. In other words, a similar mental process that draws from prior knowledge may be implicated by both prompts. So if this study is conducted again, it should make a stronger effort to distinguish control and experimental discussion prompts. For example, maybe the control prompt should not address a scenario at all, but only ask for a student's knowledge of

Aristotle's and Descartes' arguments, per se. This methodological issue may also explain why the prior knowledge aspect of this study did not reveal statistical significance with respect to higher order thinking.

Another way this study could be improved is to use philosophical arguments that are more equal in difficulty and abstraction on both the control and experimental discussion prompts. Aristotle's argument for eudemonia (happiness) is generally much easier and less abstract than Descartes' argument for certainty. Descartes' argument for certainty, in fact, is notoriously difficult for students to understand when they first encounter it, based on my ten plus years of teaching philosophy. But there is also a common sense argument to be made here. Happiness is something that many people think about, either directly or indirectly, since they may face moments in their lives when they are unhappy. Certainty, on the other hand, is not necessarily something that people are naturally driven to investigate in their lives, especially not at the level that Descartes investigates it. Descartes goes so far as to doubt his entire existence, not just whether textbooks are accurate, or whether what people are telling you is true. All of this is to say that a future study could choose two arguments that are more equal in terms of difficulty and abstraction. For example, a future study might replace Descartes' certainty argument with Peter Singer's charity argument (Singer, 1972). Singer's charity argument asks readers to consider human suffering in the present world, and whether or not we are obligated to help. Like happiness, human suffering seems like a topic that is pretty straightforward (at least, the fact that people around the world do, in fact, suffer), and one that the average student has probably contemplated (especially if she is religious and believes in a God who is all good).

The relativity of higher order thinking is another idea worth considering for future research. Lewis and Smith's (1993) research suggests that students with greater prior knowledge of a given concept will be more likely to engage in higher order thinking about that topic. From this research, I predicted that students whose activated prior knowledge is richer in relation to the topic at hand will be more likely to indicate higher order thinking. However, based on their discussion board responses in my study, no students illustrated prior knowledge of either Aristotle's or Descartes' arguments. A future study could make an effort not just to activate prior knowledge in participants, as this study did, but also to assess their prior knowledge levels relative to the arguments in question, perhaps with prior multiple choice questions. If this effort is made, a future study might either support or deny Lewis and Smith's research.

Additionally, Dochy, Segers, and Buehl (1999) point out that sometimes prior knowledge can inhibit learning, if a student activates prior knowledge that is flawed. It is possible that the lack of statistical significance between the activation of prior knowledge teaching technique and higher order thinking is due to some students' flawed prior knowledge. From my scoring of student discussion posts for this study, no serious flaws in prior knowledge were revealed that, in my view, would inhibit a student's ability to engage in higher order thinking. However, this study did not explicitly make an effort to distinguish flawed prior knowledge from accurate prior knowledge. A future study, therefore, could make this effort. For example, rather than just scoring the discussions, as was done in my study, a future study could analyze student discussion responses for the quality of prior knowledge suggested by the post itself.

I now turn to implications for teaching.

Implications for Teaching

Because both the use of peer review and the activation of prior knowledge are already tried and true teaching techniques, it would have been revealing indeed if this study suggested that these techniques were problematic. Although the study did not reveal statistically significant differences between the techniques and higher order thinking, this result is most likely due to the ceiling effect, the small sample, the flaws in the design discussed in the previous section, or some combination.

A major component of this study was the scaffolding of discussion board prompts (see Appendix 2). The research already suggests the importance of scaffolding discussion prompts in an online class, particularly in relation to higher order thinking (Sautter, 2007). Like many educational researchers, I was teaching the classes on which I conducted the research. Much more so than before I began this study, I appreciate the importance of scaffolding an online discussion. In fact, my failure to scaffold appropriately, to distinguish between control and experimental groups regarding prior knowledge, as discussed in the previous section, may in part be the reason why my results were not statistically significant in this case. This study reaffirms the importance of the instructor carefully scaffolding any discussion prompt that she writes.

Recall that in my study women showed more higher order thinking than men on some measures. This finding is worth considering as it relates to teaching, especially alongside Machado's (2011) finding that females tend to be more inhibited than males on online discussions. However, until more research is done, strong conclusions should not be drawn about gender and higher order thinking in the online environment. The only conclusion I will draw here is that online teachers should pay attention to male and

female performance in their classes, and perhaps consider ways they might make their classes as fair as possible to both genders, considering the research.

Recall that there was a statistically significant relationship in my research between age and higher order thinking: older students engaged in more higher order thinking. This finding coincides with existing literature (Huschle, 2013) and neuroscientific evidence (Larsen, Hartmann, & Nyborg, 2008). Considering the triangulation of different lines of evidence in favor of this finding, it is worth taking seriously for teachers, in particular online teachers who tend to get older, more mature students than teachers of on campus classes. If older students are already engaging in more higher order thinking, this suggests that older students need less guidance and scaffolding than younger students, generally speaking. Teachers might, for example, target particularly knowledge older students as online tutors. Or teachers might offer more difficult assignments for extra credit for interested older students, so that they are as challenged by the class as younger student with a lower capacity for higher order thinking.

Conclusion

At about the time I began this doctoral program, I had recently designed a discussion prompt in one of my online classes that asked students to discuss a belief from their worldview. One student, I still recall, wrote about her lack of belief in reincarnation. Reincarnation is silly, the student said, because there is no evidence for it. However, she went on to claim that her belief in heaven and hell was perfectly reasonable, next to reincarnation. At the time, what I saw in this student's post was the confirmation bias, the

tendency to confirm what we already believe regardless of arguments and evidence. She already believed in heaven and hell, so the lack of evidence did not bother her. But the lack of evidence bothered her regarding reincarnation, because she did not believe it to begin with.

I wrote about the confirmation bias in my first year paper, but I realize now that my seeing the student's use of the bias was the beginning of a larger trend. It is not just cognitive biases that plague online discussions, but lack of critical thought. Since seeing that student's post years ago now, I have seen many other students' posts that cried out for deeper analysis. When I first went to my adviser to discuss my dissertation, I remember drawing from my experiences teaching online discussions, and using the phrase "online discussion board effectiveness". My adviser helped me realize that I could get more specific, and we ended up settling on higher order thinking, prior knowledge, and peer review. Hence, this study was born.

But I had another, related motivation for this study. Despite the presence of some lack of critical thought, I believe that online discussions have great potential for teaching philosophy. However, many of my colleagues do not agree. While I have not formally surveyed any colleagues, informal conversations have revealed their skepticism, often relating to what they see as a lack of interpersonal communication in the online setting. A lot more could be said here, as part of a greater conversation in the philosophical community. For one, online discussions need not be a replacement for direct, personal engagement. But online discussion has the potential to lead students into deep, engaging written conversations, often mining subtle points that get glossed over in live conversation. I include a discussion exchange below between a few of my students from

the study, to illustrate the potential I believe the online discussion format has for teaching philosophy.

> *Reply to discussion prompt from Jane (pseudonym)*: Certain, in my experience, is a statement or action that is unquestionably true. As Descartes would put it, certainty is beyond the point of doubt. I believe many examples of certainty could come through simple arithmetic exercises, like '1+1=2' or '3x4=12', since there's no other way to interpret those situations into any other logical and factually correct conclusion. I think it would be hard to have me mentally challenge this, given that the ideas of addition have been hammered into my head for the past 15 or so years of education.
>
> However, while thinking about the prompt, I tried to think about many different situations where something I'd perceive was 'certain', but actually had a chance of being incorrect due to a freak accident or luck. Where I would say that flicking a light switch would certainly turn on a lightbulb in my room, the switch has a chance of being shorted out, and the bulb could be burnt out in the moment that I try to prove this statement. Additionally, I could say that dropping my phone could crack the screen, but if I'm lucky enough to have it land a certain way on a specific angle, it might just leave a scuff. The fact is that there's often a margin of error in our assumptions, regardless of how ridiculous slim it can be, which can prove our notions of certainty wrong, especially with trying to test things with real life applications.
>
> Though I feel Descartes would definitely not settle with arithmetic being my easy way out of the question. I'm sure he would assert that because he thought 1+1 equaled 19, he was correct in his thoughts. Even with the idea that I'd believe it'd be hard for him to prove his own assertions, he would still doubt that my choice was precisely certain.
>
> *Reply from Erick (pseudonym)*: Your statement is really a hard one to argue against, but, thinking with Descartes mind and its beliefs, even the simplest arithmetic's would be a great material to a big argument! As an example, your 1+1=2 it seems so clear that this is a simple straight forward question right? But in his mind, if we are a thinking creatures that the mind comes first then the body itself, I have to agree that maybe, we are being tricked by something, or, like Descartes says, an evil demon confusing our minds. In fact, 1+1 = 3 could be the real answer. Crazy right? But he has his explanations based on what is illusion and what is reality. In my opinion, it seems reasonable! So, instead of being something hard to challenge in Descartes way of thinking, it would be pretty easy!

> The same idea can be extended to your bad luck about the light bulb or even your cell phone. Rene Descartes would say that maybe it was a dream and your reality is your dream instead. Or better, we all live in a Matrix type of think where somebody has the control of our thoughts. I know… Too much to accept, but at least, he made all of us think a lot and challenge our beliefs!
>
> *Reply from Matt (pseudonym)*: Hi Jane and Erick. It seems strange that Descartes even suggests the idea of some evil genius controlling us like puppets. How could that exist? We might ask. However I suppose this is no crazier than the idea of a God. A God is an all powerful good being, and the evil genius is an all powerful bad being. There's no denying that both beings are at least possible, given everything else we know about the universe.
>
> The skeptical conclusion of the Matrix does seem like a lot to accept. But my friend was telling me about this guy Nick Bostrom (I think he works at Oxford) who has actually convinced many people that we may actually be living in some sort of sophisticated simulation like the Matrix. Here is a link to his website where he discusses his theories.

Without going into too much detail, what is happening here is that students are not just responding to each other mindlessly, but helping each other to understand the concepts and connecting the discussion to outside sources. Jane's reasonable claims about math are respectfully challenged by Erick, and then Matt (responding logically to both Jane and Erick) connects it all to an Oxford scholar. To me, this example is one small taste of the promise of online discussions: meaningful, reasonable exchanges in which new ideas emerge and intertwine as part of the process. Importantly, when I first began teaching online, my students almost never engaged in this sort of discussion (as my reincarnation example above illustrates). If my students are more deeply engaged now it is, in my view, due in large part much of what I learned from conducting this study and from being in this doctoral program generally. If I can improve and use online discussion boards for meaningful philosophical discussions, then so can other philosophy professors.

For me personally, the lack of statistical significance between the techniques and higher order thinking does not do this study justice. I am such a better online teacher, and better teacher in general, as a result of conducting this study. I am still teaching online philosophy classes, and I have one in session right now. Drawing from this study, I have made several changes to these classes. For example, my criteria for grading student discussion posts now includes the criteria I drew from the literature regarding higher order thinking (see Appendix 1). Recently I spent some time, as I do every week, going through my students' posts on the discussions, responding when necessary and guiding them along. Although it is anecdotal, the evidence of increased student engagement is there, and I could not have pushed my students there without this study.

I am actually excited to experiment with peer review in my online classes. Without this study, I would not have known about Scott Klemmer's research, and the seven habits of effective peer review. I have done some peer review in my online classes, but I plan to get much more serious about it. The college where I teach will soon be migrating to a new course management system, so perhaps I can implement all of the seven habits in that system (which has not yet been chosen).

So, all in all, what began as the logical endeavor of a philosophy professor to investigate the confirmation bias in his students and prove his colleagues wrong, ended with a broadening of that same professor's horizons into peer review, prior knowledge, and higher order thinking.

Appendix I: Student and Professor Instructions for Peer Review

The following are the categories and instructions that will be given to students when they peer review each other's discussion posts. The categories will also be given to fellow philosophy professors when they peer review student discussion posts for evidence of higher order thinking.

Please evaluate the student's response to the question based on the following three categories. It is important to your own grade on this discussion that you provide a justification for your evaluations. It is also important to your grade to follow the instructions carefully. [These last two instructions will be given only to students.]

Understanding) On a scale of 1 to 3, how well do you think the student understands Descartes' argument about certainty (the "I think, therefore I am" argument)? Choose 1 for no understanding, 2 for partial understanding, and 3 for full understanding.

Implications) On a scale of 1 to 3, how well do you think the student understands the implications of Descartes' argument about certainty? In other words, does the student mention anything that goes beyond what Descartes' argument actually is, but that is faithful to the argument? For example, Descartes did not specifically say that you can't be certain of the computer screen in front of you, but it is nevertheless implied by his argument. Choose 1 for no understanding, 2 for partial understanding, and 3 for full understanding.

Connections) On a scale of 1 to 3, how accurately do you think the student connects his or her thoughts on certainty with Descartes' argument on certainty? In other words, does the student accurately identify the way that his or her answer is similar to or different from Descartes' argument? Choose 1 for an inaccurate connection to Descartes, 2 for partial accuracy, and 3 for full accuracy.

Now please add up the student's scores for each of the three categories to give him or her a score out of 9. For example, if the student got a 2 on understanding, a 3 on implications and a 2 on connections, her score would be 7/9. Once again, don't forget to provide justifications for your evaluations. [This last instruction will be given only to students].

Appendix II: Discussion Prompts

The following is the experimental discussion prompt for the discussion on Descartes' certainty:

> Think back to a time in your life when you were absolutely certain of something. What was the thing you were certain about? And why were you certain about it?
>
> To put the question another way, what is certain? What do you know to be true with 100% certainty? Anything? Everything? Nothing? Why or why not? Would you stake your life on it?
>
> One last way to phrase the question is: how much is it possible to doubt? Is there anything that is absolutely impossible to doubt? These questions might almost seem silly, but try to really think about them.
>
> Now that you have addressed these questions, how do you think Descartes would react? Based on Descartes' arguments about certainty from the chapter (specifically his famous "I think, therefore I am" argument), would he agree that what you've chosen is absolutely certain? What *is* Descartes' argument about certainty?

The following is the control discussion prompt for the discussion on Descartes' certainty:

> Consider the following scenario. Julia and her friend are walking down the street. A person crosses their path quickly. Julia tells her friend that the person who just crossed their path quickly was President Obama. Julia's friend is skeptical and doesn't believe her, but Julia insists that she is absolutely certain that she just saw President Obama.
>
> Based on Descartes' argument about certainty (his "I think, therefore I am" argument), how would he respond to Julia? Would he agree or disagree with Julia when she says that she is absolutely certain that she saw Obama? What *is* Descartes' argument about certainty?

The following is the experimental discussion prompt for the discussion on Aristotle's eudemonia:

> Think back to a time in your life when you were happy. What were you happy about? Why were you happy? If you could define happiness, how would you do it?

> Now that you have addressed these questions, how do you think Aristotle would react? Given Aristotle's argument for eudemonia from the chapter, how is your idea of happiness similar to or different from Aristotle's eudemonia? What *is* Aristotle's argument for eudemonia?

The following is the control discussion prompt for the discussion on Aristotle's eudemonia:

> Imagine that a person, Carlos, is out at a restaurant hanging out with his friends. Carlos is socializing, laughing, and joking—he appears to be having a great time. One of Carlos' friends, Tom, remembers that yesterday Carlos was complaining about how he didn't like his job. At dinner, Tom asks Carlos how he feels, and Carlos answers that he feels totally happy.
>
> Now that you have addressed these questions, how do you think Aristotle would react? Given Aristotle's argument for eudemonia from the chapter, how is Carlos' apparent happiness similar to or different from eudemonia? What *is* Aristotle's argument for eudemonia?

Appendix III: Multiple Choice Questions

The following are the multiple choice questions that will be used, along with instructor peer review of student discussion posts, to assess students' higher order thinking.

Descartes Quiz Questions:

Descartes believed that methodic doubt…

a. Is a very useful tool with which to rid himself of uncertainty.
b. Is a hopelessly cynical and pessimistic stance.
c. Is the first sign of insanity.
d. Should be avoided at all costs.

What is the one thing Descartes is certain of after he has established the uncertainty of everything else?

a. The existence of his own mind and body.
b. The existence of himself only as a "thinking thing".
c. The existence of other people.
d. Death.

Descartes was searching for…

a. Certainty.
b. A trade route to India.
c. Uncertainty.
d. Cartesian dualism.

Descartes Prior Knowledge Questions:

Of the following options, which best defines certainty in your view?

a. Something you can see right in front of you.
b. An idea that cannot be doubted.
c. An experience that cannot be doubted.
d. Something that has been proven to be true.
e. None of the above.

Of the following options, which is the most certain for you?

a. Your emotions.
b. The computer screen in front of you.
c. What happened yesterday.
d. God's existence.
e. None of the above.

Aristotle Quiz Questions:

According to Aristotle, eudemonia is…

a. The deep sense of happiness that comes from receiving the gift of diamonds.
b. The deep sense of happiness that comes from living a full and vigorous life.
c. The idea that reality is composed of the natural world.
d. The idea that admitting one's ignorance allows wisdom to develop.

Some see happiness as a temporary state, whereas Aristotle saw eudemonia as…

a. The first cause in his theory of causes.
b. A process.
c. A final state in which no further change occurs.
d. Inaction.

Aristotle claimed that a number of conditions were more likely to lead to eudemonia, including…

a. Knowledge and wisdom.
b. Living in a location with more sunny weather.
c. Luck and looks.
d. Trustworthiness.

Aristotle Prior Knowledge Questions:

Of the following activities, which makes you the most happy?

a. Traveling.
b. Socializing with friends.
c. Thinking about your accomplishments.
d. Reading, writing, or playing video games by yourself.
e. None of the above.

Of the following options, which do you think best defines happiness?

a. Feeling good in the present moment.
b. Feeling a sense of pride when looking back on your life.
c. Looking forward to the future.
d. All of the above.
e. None of the above.

Appendix IV: Consent Form

Here is the consent form that I will ask interested students to fill out electronically:

University of California, San Diego
Consent to Act as a Research Subject

Can online philosophy discussion boards be designed to lead to higher order thinking?

Who is conducting the study, why you have been asked to participate, how you were selected, and what is the approximate number of participants in the study?
Luke Cuddy is conducting a research study to find out more about how to design online philosophy discussion boards to lead to higher order thinking. You have been asked to participate in this study because you are a student in Professor Luke Cuddy's online philosophy class. There will be approximately 40-60 participants in this study.

Why is this study being done?
The general purpose of this study is to learn more about online discussion boards for educational purposes.

What will happen to you in this study and which procedures are standard of care and which are experimental?
If you agree to be in this study, the following will happen to you:
You will follow the instructions and complete the class assignments, just as you would do anyway as a student in the class. However, the experimental portions of the study are as follows. Before or during two of the discussion boards, you will be asked a few multiple choice questions, or an open ended question that requires a short written response, or you will be asked to peer review another student's discussion response according to provided guidelines.

How much time will each study procedure take, what is your total time commitment, and how long will the study last?
The study will be conducted during the class itself. If you participate in the study, the additional time requirement for answering questions or peer reviewing before or during the two discussion boards will be between five and ten minutes per discussion.

What risks are associated with this study?
Participation in this study may involve some added risks or discomforts. These include the following:

1. A potential for the loss of confidentiality, which will be minimized by keeping the research within a password encrypted file on multiple password protected computers and by using pseudonyms. Research records will be kept confidential to the extent allowed by law. Research records may be reviewed by the UCSD Institutional Review Board. Once the research is no longer needed, it will be destroyed.

2. There is the slight possibility that some subjects in the study will feel some mild emotional discomfort, embarrassment, or pressure for participating. However, participants can choose to opt out of the study at any time.
3. Participants may experience boredom. However, participants may stop their participation at any point in the process simply by informing the researcher that they do not wish to continue.

Because this is a research study, there may also be some unknown risks that are currently unforeseeable. You will be informed of any significant new findings.

What are the alternatives to participating in this study?
The alternatives to participation in this study is to not participate. There is no penalty for non-participation, and non-participation will not affect the relationship with the PI.

What benefits can be reasonably expected?
There may or may not be any direct benefit to you from participating this study. The investigator, however, may learn more about online discussion boards, and society may benefit from this knowledge.

Can you choose to not participate or withdraw from the study without penalty or loss of benefits?
Participation in research is entirely voluntary. You may refuse to participate at any time without penalty or loss of benefits to which you are entitled. If you decide that you no longer wish to continue in this study, you will be required to inform the PI, Luke Cuddy.

You will be told if any important new information is found during the course of this study that may affect your wanting to continue.

Can you be withdrawn from the study without your consent?
The PI may remove you from the study without your consent if the PI feels it is in your best interest or the best interest of the study. You may also be withdrawn from the study if you do not follow the instructions given you by the study personnel.

Will you be compensated for participating in this study?
In compensation for your time, you will receive a $10 gift certificate for participating in this research, as well as 1.5% of extra credit toward your grade in the class.

Are there any costs associated with participating in this study?
There will be no cost to you for participating in this study.

Who can you call if you have questions?
Luke Cuddy has explained this study to you and answered your questions. If you have other questions or research-related problems, you may reach Luke Cuddy at 619-739-1717.

You may call the Human Research Protections Program Office at (858) 657-5100 to inquire about your rights as a research subject or to report research-related problems.

Your Signature and Consent
You have received a copy of this consent document.

You agree to participate.

_____ _____
Subject's signature Date

References

Alvermann, D. E., Smith, L. C., & Readence, J. E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. *Reading Research Quarterly, 20*(4), 420-436.

Allen, I. E., & Seaman, J. (2013). Changing course: Ten years of tracking online education in the United States. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved October 13, 2013, from http://www.onlinelearningsurvey.com/reports/changingcourse.pdf

Bartlett, F. C. (1958). *Thinking: An experimental and social study*. London: Allen & Unwin.

Blackboard. (2016, April 26). Retrieved from http://www.blackboard.com/

Cathey, C. (2007). Power of peer review: An online collaborative learning assignment in social psychology. *Teaching of Psychology, 34*(2), 97-99.

Cathey, C., Panza, C., & Potthast, A. (2006). Thinking outside the room: Enhancing philosophy courses with online forums. *Teaching Philosophy, 29*(4), 279-297.

Chen, P.-S. D., Lambert, A. D., & Guidry, K. R. (2010). Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers & Education*, *54*(4), 1222–1232. doi:10.1016/j.compedu.2009.11.008

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*(2), 145-186.

Ennis, R. E. (1993). Critical thinking assessment. *Theory into Practice, 32*(3), 179-186.

Fox, T. (2013, November 20). *Peer Assessment in Online Education: Training Students to Know What's 'Good'*. Retrieved from http://www.calit2.net/newsroom/article.php?id=2261

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *The American Journal of Distance Education, 15*(1), 7-23.

Gerber, S., Scott, L., & Clements, D. H. (2005). Instructor influence on reasoned argument in discussion boards. *Educational Technology Research & Development*, *53*(2), 25–39.

Huschle, B., & Philosophy Documentation Center. (2013). Assessing learner outcomes in traditional and online medical ethics courses. *Teaching Philosophy*, *36*(1), 1–18.

Kay, R. H. (2006). Developing a comprehensive metric for assessing discussion board effectiveness. *British Journal of Educational Technology*, *37*(5), 761–783.

Kulkarni, C., Socher, R., Bernstein, M. S., & Klemmer, S. R. (2014). Scaling short answer grading by combining peer assessment with algorithmic scoring. In *Proceedings of the First ACM Conference on Learning @ Scale Conference*, 99–108. New York, NY, USA: ACM. doi:10.1145/2556325.2566238

Land, S. M., & Dornisch, M. M. (2002). A case study of student use of asynchronous Bulletin Board Systems (BBS) to support reflection and evaluation. *Journal of Educational Technology Systems*, *30*(4), 365–377.

Larsen, L., Hartmann, P., & Nyborg, H. (2008). The stability of general intelligence from early adulthood to middle age. *Intelligence, 36*, 29-34.

Lewis, A, & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131-137.

Machado, C. (2011). Gender differences in student discourse on discussion board and blogs: An instructor's quest to create a level playing field in a hybrid classroom. *Journal of Interactive Online Learning*, *10*(1), 36–48.

Maier, N. R. F. (1933). An aspect of human reasoning. *British Journal of Psychology, 24*, 144-155.

Maier, N. R. F. (1937). Reasoning in rats and human beings. *The Psychological Review, 44*, 365-378.

Matheson, R., Wilkinson, C. S., & Gilhooly, E. (2012). Promoting critical thinking and collaborative working through assessment: Combining patchwork text and online discussion boards. *Innovations in Education and Teaching International*, *49*(3), 257–267.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Mertens, D. (2010). *Research and evaluation in education and psychology*. Thousand Oaks, CA: Sage Publications.

Meyer, K. A. (2003). Face-to-face versus threaded discussions: The role of time and higher order thinking. *Journal of Asynchronous Learning Networks, 7*(3), 55-65.

Pritchard, D. E., Lee, Y-J., & Bao, L. (2008). Mathematical learning models that depend on prior knowledge and instructional strategies. *Physics Education Research, 4*(010109), 1-8.

Resnick, L. B. (1987). *Education and learning to think*. Washington, D. C.: National Academy Press.

Rovai, A. P. (2004). A constructivist approach to online college learning. *Internet and Higher Education, 7*, 79-93.

Sautter, P. (2007). Designing discussion activities to achieve desired learning outcomes: Choices using mode of delivery and structure. *Journal of Marketing Education*, *29*(2), 122–131.

Schon, D. (1987). *Educating the reflective practitioner: Towards a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass, Inc.

Sher, G. (2015). Truth as composite correspondence. In *Unifying the Philosophy of Truth*, edited by T. Achourioti, H. Galinon, J. Martínez Fernández, and K. Fujimoto. Springer.

Singer, P. (1972). Famine, Affluence, and Morality. *Philosophy and Public Affairs, 1*(3), 229-243.

Song, L., & McNary, S. W. (2011). Understanding students' online interaction: Analysis of discussion board postings. *Journal of Interactive Online Learning*, *10*(1), 1–14.

Wijekumar, K. K., & Spielvogel, J. (2006). Intelligent discussion boards: Promoting deep conversations in asynchronous discussion boards through synchronous support. *Campus-Wide Information Systems*, *23*(3), 221–232.