

USING PROBLEM-CENTERED LEARNING TO FOSTER
ARGUMENTATION IN INTRODUCTORY SOCIOLOGY

A Dissertation
presented to
the Faculty of the School of Information Science and Learning Technologies
at the University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
HOLLY R. HENRY
Dr. David H. Jonassen, Dissertation Supervisor
Dr. Joi L. Moore, Dissertation Co-Advisor

MAY 2013

The undersigned, appointed by the dean of the Graduate School,
have examined the Dissertation entitled
USING PROBLEM-CENTERED LEARNING TO FOSTER ARGUMENTATION IN
INTRODUCTORY SOCIOLOGY

Presented by Holly Henry

A candidate for the degree of

Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

Professor David H. Jonassen (deceased)

Professor Joi L. Moore

Professor Edward E. Brent

Professor Gail E. Fitzgerald

Professor James M. Laffey

Dedication

This manuscript is dedicated, along with my sincere gratitude and profound respect, to the memory of my advisor, Dr. David Jonassen. To say that Dave is well-known by educational researchers across a broad spectrum of sub-specialties is stating the obvious, and yet it is still an understatement. He was an exceptionally prolific researcher over the course of a long career during which the educational video technology of his own dissertation gave way to Computer Based Training (CBT) which was almost certainly eclipsed by the advent of online classrooms via the Internet which seem to have circled back around to educational video.

Others have written about all of that – about his profound influence on the field of Educational Technology in particular and educational research in general; his scholarly record; his unparalleled achievements as a researcher. While there is no doubt that those achievements influenced me as much as they influenced the field (would I have chosen the University of Missouri if I hadn't been familiar with Dave's work?), such remembrances are best written by those who have been in that field longer and have a more firm grasp on its history than I do.

Instead, my sole "claim to fame" when it comes to David Jonassen is that I'm the last of the "Dissertations Chaired" section in his 83 page CV (<http://web.missouri.edu/jonassend/CV-JONASSEN.pdf>). I'm on page 82, though he hadn't updated the title to reflect a committee member's concern at the proposal defense that necessitated a change. And sadly, I wasn't able to defend in 2012, as he'd listed it on his CV, since we lost Dave before the date he'd scheduled with the committee for that defense. "Sorry for crunching everything so much," he wrote when we were sending the dissertation out to the committee. "I have this need for closure and want to make sure that you finish with as little hassle as possible." And so my contribution to the

literature remembering Dave's contributions to our field is to tell you a little about how Dave turned novices into experts in his own practice as a teacher and a researcher.

The short (aka tl;dr) version is that Dave mentored me through the entire process of earning the doctorate, afforded me opportunities on a wide range of research projects to develop my skills, influenced my thinking about my professional and academic field, and introduced me to *La Fin du Monde*. The long version is too long to prevent my dissertation from becoming two volumes. What follows is merely a snapshot of David Jonassen as teacher and mentor.

Everyone who has ever met Dave knows that his brain worked faster than most of us ever dream of, so he might not ever be the person you'd identify as a patient teacher. And the way he signaled that he was hearing you was this, "yeah, yeah, yeah" or "right, right, right" that seems to be a part of the campus culture around here, but his was ten times as fast as everyone else's version so that he often fit in many more yeahs or rights than anyone else I've ever heard. Hear he did, though, and he always assimilated at that same lightning speed.

So it wasn't in the way he conversed that demonstrated his patience, and it led many a student to assume he wasn't patient at all. However, he often surprised me with articles he'd run across that reminded him of an idea we'd discussed months before. And of course he was always ready with a suggestion of whose work to check out right in the moment, too, along with an anecdote or two about having discussed that person's research over a beer at a conference. Dave seemingly knew everyone in the field and remembered every conversation he'd ever had.

This trait perhaps most of all tended to either endear him to students or alienate them from him. If you have ever snuck a peak at his ratemyprofessors.com profile, you can see this tendency toward the extremes in his impressions on students. Some students interpreted this familiarity as nothing more than ego-driven name dropping. The rest of us understood that his

interconnectedness with the breadth of learning technology research and seemingly all of the related fields that influenced and informed it were part of what made him such a great teacher. He was like a walking, talking cognitive flexibility hypertext that could help you make sense of just about any complex theory and how it related to other theories you've studied. And just when you started to be overwhelmed, he'd shrug and say, "but it's just a theory" (his shorthand for noting that theories were always ripe for empirical research to see whether and in what conditions they held up under scrutiny).

As you might imagine, Dave walked his talk when it came to his constructivist philosophy of learning, and it was here that he expressed his patience with me as a student. He gave me the opportunity right out of the gate to see how the sausage is made and he encouraged me to participate in it; to shoot high and miss and shoot again. And again. And still again. Even in this document, I shot high and missed a bit, but I learned enough on the way to write this dissertation.

Dave was fond of quoting Karl Popper: "All of life is problem-solving!" he'd say, and he set about proving it to us in his courses and in our participation in his various research and consulting projects. The first thing he introduced you to would be the problem you'd be expected to help solve. In his Problem Based Learning Environments course, he put us in groups and assigned us to a faculty member he'd recruited. The faculty member would have an example of a problem that he or she would want students to be able to solve. It was our problem to figure out how to design an environment to support those students in solving the problem that, like the problem Dave had just given to us, made the problem the central focus of the learning experience with the accompanying knowledge transfer types of activities in service to that central focus. The study in this dissertation examines the efficacy of the learning environment our team, Wenting

Jiang, Pete Kinser, Hsiao Wen Liao, and I, created for the class. Dr. Ed Brent, the faculty member who was our client in that course, helped me to fine tune the content and the types of assignments that his students would perform in the semesters that followed Dave's PBLE course, but the basic design of the environment is our solution to the problem laid before us in that course, informed by the types of theories Dave felt best explained the type of problem we were working on and how to solve them. Without him, then, not only would this dissertation not exist, but neither would an entire series of learning environments designed to teach students in a variety of disciplines how to solve the problems that would be typical of practice in their fields. By having taught us, Dave is still helping to teach those students in some small way as well.

I am also grateful to Dave for giving me responsibility as early and as often as possible on research projects. Of course the tasks started small, but I was always given a fair amount of autonomy and ownership over how to carry them out. He really encouraged full participation from graduate students, asking for our input on design and methodology even when we were brand new to a project. Of course, if you met Dave, you already know what an imposing figure and enormous presence he was in nearly any size room, so you can probably imagine how effectively he conveyed his skepticism whenever he felt our ideas were headed off track, often without so much as a word. As new grad students, most of us backed off quickly upon seeing him eye us over the top of his glasses across the table with one of those skeptical looks. However, we learned over time that skepticism didn't necessarily mean that we were wrong, per se, but that we hadn't convinced him. We also learned how to make more effective, well-supported arguments, both because we respected and wanted to impress him...and because we wanted to avoid that look.

There is always, out of necessity given the power imbalance inherent in the relationship, something of a distance between professor and student. However, while Dave held much of his personal life in reserve behind that wall, he was still very social with us as grad students, regularly scheduling after-work happy hours or dinners; even inviting the entire department to his home to welcome the new cohort. In these settings, Dave shared his famous laugh more frequently than he did in our research team meetings or the classroom, along with stories of mountain climbing, chats with famous researchers, or what he liked about a favorite beer. He asked a lot of questions and encouraged the sharing of all of our life stories and the paths that had led us to the middle of Missouri. I remember fondly some of the casual conversations we shared and regret that we never thought to take photographs during those occasions, though I am grateful for the snapshots that remain in my memory. Life's best moments are truly ephemeral, aren't they?

As a notable researcher in our field, Dave has also attracted so many grad students who have become notable researchers themselves. To those who preceded me in the collection of dissertations that proudly lined his office shelves: know that he was particularly proud of you and often cited your past and current work in classes he taught. He sometimes held one of you up as an example to us about how we might translate an idea into a research or instructional design. He celebrated your successes and delighted in recounting them. To me, as the last Ph.D. he finished, you are as notable as he is and I have probably read most of what you have published as have many of the students from his courses.

My final words are for Dave himself, offered posthumously though as they must be. Dave, I will carry your teaching and your friendship with me wherever my path leads. Thank you for everything you contributed to our field and everything you contributed to my education within

it. Thank you for the opportunity to work with you side by side and especially for the occasions we got to laugh together. Thank you for everything. Thank you. I will do my best to pay it all forward.

Acknowledgements

One of the things I've learned in my graduate studies is how much of an advantage there is in educational attainment among children whose parents read to them from a young age versus those who don't. I offer gratitude beyond words and much love to my parents, Ron and Emily, for reading to me and buying me that red dictionary along with the red bike for my fourth birthday; for supporting me and challenging me to keep pursuing my dreams in the face of adversity; for instilling in me their value of education; and for raising me to be curious, persistent, and strong-willed enough to accomplish this goal.

Dr. Joi Moore has been a role model and a valued supporter from the day I started at the University of Missouri. When we lost Dave just before the originally scheduled defense, Joi stepped in on top of everything else she had recently taken on in the department and kept me from feeling orphaned. In addition to her service on my committee, she has been my teaching mentor, my guide for focusing on the human side of human-computer interaction, and my conference roommate. She has always given me solid advice and often provided the inspiration for me to keep going when I was tempted to give up. Thanks so much, Joi.

I am particularly grateful to Dr. Edward Brent whose passion for promoting critical thinking skills in introductory students in sociology was the spark for this entire project. He lent invaluable expertise in developing the instructional materials and shaping the design of the case study website and all of the assignment instructions. Thanks, Ed. The soul of your sociological imagination shines through this entire project.

Dr. James Laffey was the most supportive devil's advocate I could ever have hoped for, challenging me to push my own envelope at every stage of my doctoral education. He was also

the first person to welcome me to Columbia and opened the doors of his home to me when I first arrived. While this project doesn't reflect his expertise in the social aspects of learning and technology as we both might have hoped before it took shape, his influence is still pervasive throughout this document and his ideas about the ways in which students influence each other's learning are ideas I put into practice in my professional life daily. Thank you, Jim.

I was not alone in designing the case study website used in this study. This was truly a team effort under the supervision of Drs. Jonassen and Brent. There would have been nothing here to research without the outstanding work of my teammates Wenting Jiang, Pete Kinser, and Hsiao Wen Liao to realize the vision of how to translate cognitive flexibility theory into a real learning environment. Thank you to each one of you for your invaluable contributions.

In addition, I want to thank Dr. Brent and his entire team at IdeaWorks for developing SAGrader and the semantic network underlying the scoring for the assignments in this project; for providing additional data whenever I requested it; and for answering all of my questions at virtually a moment's notice along the way. I hope that, in some small way, this project helped advance your goals for SAGrader.

There are not words to express how much I appreciate Dr. Rose Marra's willingness to provide support in the face of her own unimaginable loss. Rose, thank you for your incredible generosity, your encouragement, and the gift of your friendship. I hope that we will be able to collaborate often in the future.

My gratitude encompasses the entire faculty and staff at the School of Information Science and Learning Technologies. The faculty's scholarship and dedication to teaching with technology every day makes me proud to be a part of this department. The staff who must have grown weary of my terrible habit of forgetting about paperwork have always been not only

patient with my foibles, but also warm and supportive of my efforts. They have helped me clean up my administrative messes on more than one occasion and their dedication to helping the graduate students in the department succeed is behind every one of us who does so. Special thanks to Amy and Toni who always made me feel like I was part of the family.

Many members of my cohort here have become not only colleagues, but also true friends without whom I would not have made it this far. Special thanks to Dr. Krista Galyen who provided a direction on something I had no idea how to tackle. My appreciation to Krista again and to Dr. Camille Dickson-Deane for the particular joy of travelling this road together and being a joint force in what is, at its core, often a solitary endeavor. Thanks to Drs. Matthew Schmidt, Matt Easter, Sean Goggins, I-Chun Tsai, Pei-Ju Liu, Kyungbin Kwon, Moon-Heum Cho, and Young Hoan Cho for the “I’ve just survived where you are now” advice, and to Gordon Graber, Dr. Andrew Tawfik, and Xianhui Wang for peer learning on more assignments and research projects than I can count.

Many thanks to Zac March who provided invaluable assistance in helping me to balance dissertation-writing with full time work as well as constant support and encouragement along the way. It has been my pleasure to work for him for the past three years and only wish it could last a bit longer.

...

The only people I know from the first college I attended who went to graduate school are Dr. Sue Spagno-Ramlo and Dr. Dara Wegman-Geedey. They, too, provided inspiration and motivation during my grad school years, and I am ever so grateful for their faith in me.

Now, I’m sure there are others from my first undergraduate institution who attained graduate degrees. I just don’t know of them. Since I couldn’t afford to finish the BA I’d started

there, I lost touch with many of the people I knew there as I moved back home and started working.

When I was in my thirties, I returned to school as a non-traditional student to complete the bachelor's degree that had been a lifelong goal with no intention of carrying my education beyond that milestone. While many non-traditional students suffer from being treated like second class citizens at their institutions, the faculty who provided evening courses in the Information Technology program at the University of Cincinnati provided a winning combination of professional practice education grounded in solid scholarship that afforded me a path from non-traditional bachelor's degree completion student to graduate student. I will never forget the Saturday morning statistics instructor who began an answer to my question about correlation with, "When you go to grad school and study statistics there..." before I even considered continuing after the BS. All dreams were welcome there and I am grateful to the faculty there for encouraging mine. I'm also grateful to Luann Sieve who went on from that non-traditional bachelors program in which we met to get her MA and talked me into doing the same.

When I was accepted as a master's student at Teachers College Columbia University, I was still focused on improving my professional practice rather than scholarship. It was through the mentoring and model of Dr. Howard Budin that I gradually evolved from practitioner to scholar, though it is a process that I consider to be a lifelong one. I am grateful for the opportunity to explore the field of learning technology with the guidance of the talented and dedicated faculty in the Math, Science, and Technology department at Teachers College that served to whet my appetite to continue with doctoral study. I appreciate, too, the way that Dr. John M. Broughton from my outside department broadened my understanding about what education means and how it both influences and is influenced by the culture in which it is situated.

I also want to thank my colleagues who were doctoral students at TC when I was pursuing my MA. By allowing me to understand their experiences of the dissertation process, and by encouraging me to pursue it, they may be most directly responsible for my choice to continue past my master's degree. While they will understand that there have been times in the process when I have thought I might have been better off if they had not encouraged this crazy notion, I am grateful to Dr. Lin Lin and Dr. Brenda I. López Ortiz for their support and encouragement, both throughout my master's program and in my transition to the doctoral program. Thanks also to Dr. Christopher M. Westgate, a colleague in my master's program, who took the leap before me and cheered me on to do the same. I am also obliged to Dr. Steven W. Goss for supervising my master's thesis work and recommending me for work in the field after my MA. That work let me know I was on the right path with the area in which I chose to study.

Finally, I owe a huge debt of gratitude to friends and family near and far whom I have too often neglected in this process. You know who you are even though the dozen people who may read this document in my lifetime will not, so I will not mention each name on this page; rather, know that I am speaking it in my heart. Thank you. I am who I am because you were a part of my life and I am grateful for your presence in it each and every day.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	vii
List of Figures and Tables.....	xi
Figures.....	xi
Tables.....	xi
Abstract.....	xiii
Chapter 1: Rationale for the Study.....	1
Theoretical Framework.....	2
Problem-solving.....	2
Cognitive flexibility hypertexts.....	3
Argumentation.....	4
Purpose and Impact of the Study.....	6
Research Questions.....	7
Chapter 2: Literature Review.....	8
Problem-Solving.....	10
Cognitive Flexibility Hypertexts.....	14
SAGrader and Revisions.....	17
Argumentation.....	20
Demographic Considerations.....	28
Gender differences in argumentation.....	28
First generation college students.....	31
Summary.....	34

Chapter 3: Methodology	35
Research Questions	35
Participants	35
Instructional Methods	36
Case website.....	38
SAGrader.	40
Development of the environment and associated treatments.	41
Treatments.....	43
Summarize treatment.	45
Construct treatment.....	46
Counter-argue treatment.	47
Data Collection.....	47
Design.....	48
Methods.....	48
RQ1: Learning outcomes.	48
Overall scores.	48
Components.	49
RQ2: Revisions.	50
RQ3: Website usage.....	52
Statistical analysis of usage statistics.....	52
Analysis of page navigation.....	54
Limitations of the study.....	56
Chapter 4: Results.....	58
RQ1: Learning Outcomes.....	59
Overall scores.....	59

Components.....	64
RQ2: Revisions.....	67
RQ3: Website Usage	70
Statistical analysis of usage statistics.....	71
Analysis of page navigation.....	74
Chapter 5: Discussion	85
Effects on Argument Quality.....	85
Effects on the Writing Process	87
Effects on Cognitive Flexibility Hypertext Usage.....	90
Chapter 6: Summary	93
Summary of Results	93
Implications for Practice.....	94
Implications for Future Research	97
Conclusion.....	99
References.....	101
Appendix A: Holistic Argument Scoring Rubric.....	109
Appendix B: Assignment Instructions	110
Case 1 Instructions	110
Case 1A Argumentation: Make a Choice and Argue for Your Case	110
Description.....	110
Questions	110
Case 1C Argumentation: Counter an Argument	111
Description.....	111
Questions	111
Case 1D Argumentation: Discuss the pro's and con's of each potential renter	112

Description.....	112
Questions	112
Case 2 Instructions	112
Case 2A: Choose and argue	113
Description.....	113
Questions	113
Case 2C: Counter an argument.....	114
Directions.....	114
Questions	115
Case 2D: Discuss.....	115
Description.....	115
Questions	116
Case 3 Instructions	117
Description.....	117
Appendix C: Component Items	119
VITA.....	125

List of Figures and Tables

Figures

Figure 1. Screenshot from the case website.....	39
Figure 2. Screenshot from the Case 1 assignment in SAGrader.....	40
Figure 3. Distribution of the first revision scores for the Case 3 assignment for all treatment groups.	60
Figure 4. Distribution of the final submission scores for the Case 3 assignment for all treatment groups.	61
Figure 5. Simple plot of first revision vs. final scores for all three treatment groups.....	62
Figure 6: Proportion of significant items achieved by group.....	67
Figure 7. Average ranks for pairwise comparisons of website usage measures. Lighter gray lines represent statistically significant differences.....	73
Figure 8. Legend for network analysis visualizations.....	75
Figure 9. Centrality visualization for the summarize treatment group	76
Figure 10. Centrality visualization for the construct treatment group	77
Figure 11: Centrality visualization for the counter-argue treatment group.....	78
Figure 12: Circular layout for the summarize treatment group.....	79
Figure 13: Circular layout graph for the construct treatment group	80
Figure 14: Circular layout graph for the counter-argue group.....	81
Figure 15: Dual-circular layout graph for the summarize treatment group	82
Figure 16: Dual-circular layout for the construct treatment group	83
Figure 17: Dual-circular layout for the counter-argue treatment group.....	84

Tables

Table 1. Between subjects effects on first draft scores for Case 3.....	64
---	----

Table 2. Significant component items	66
Table 3. Tests of between-subjects effects by treatment group on score and number of revisions	69
Table 4. Games-Howell pairwise comparisons.....	69
Table 5. Tests of normality	72

Abstract

This mixed methods case study explored the efficacy of a problem-based learning environment and three different instructional methods designed to foster argumentation in an introductory sociology course. While no statistically significant differences were found in the overall assignment scores among the three treatment groups, there were significant differences in the effort they expended to achieve those scores, suggesting that a treatment in which students were instructed to generate counter-arguments to a provided model essay was most efficient while a treatment in which students were guided to construct an argument step-by-step encouraged more time-on-task.

Chapter 1: Rationale for the Study

A survey of leaders practicing in the field of sociology (Persell, Pfeiffer, & Syed, 2007) identified learning to “think sociologically” as the top goal for students in introductory sociology. Likewise, Atkinson and Hunt (2008) argue that instructors should encourage students to “think and act like sociologists” through the use of assignments that allow the student to engage in sociological practice. This parallels the notion forwarded by Brown, Collins, and Duguid (1989), which proposes that knowledge is not simply an abstraction, but rather part and parcel of enculturated practice within a domain.

One challenge to such an approach is that many students in an introductory sociology course may not even be majors in the discipline, let alone plan to become practitioners in the field. Consequently, fully situating class activities within those that may be typical of practitioners may limit students’ appreciation of the relevance of sociological thinking in their own lives and planned careers. When discussing their concepts of sociological thinking, survey respondents in the Persell et al. (2007) study raised themes related to the interplay between social structures and individual lives. In addition, other learning objectives highlighted by the study include broadly applicable concepts, such as critical thinking, social construction of ideas, and even “the importance of trying to improve the world” (Persell et al., 2007). This suggests that learning activities which encourage students to apply sociological thinking to everyday life activities could be an effective way to enculturate students within the discipline regardless of their intended major or career path, particularly if they are situated in contexts that reflect ordinary life.

This study examined the efficacy of a set of learning activities designed around the sociological concept known as social stratification. Social stratification refers to the ranking of people and groups within a society and the rules, formal and informal, that govern how resources

and status are accorded based upon those ranks. Ranking inherently results in unequal distribution of resources and status, and sociologists study both the nature of the ranking systems themselves and the effects of the inequality that results from them (Kerbo, 2006). In this way, structural inequalities that result from the systems of social stratification in a given society represent problems that are typical of those with which sociologists concern themselves while also being readily exemplified in the daily lives of non-sociologists. By applying a “sociological lens” to these problems, students in an introductory sociology course can better appreciate the subject matter they are being asked to learn and better see its relevance to their own lives.

Theoretical Framework

In order to teach students to think like sociologists, they must be guided to understand how sociologists approach the problems in their domain (Jonassen, 2000). Thus, problem-solving served as the overarching framework under which the learning activities were designed. The nature of the type of problem that structural inequality represents suggested the best theoretical frameworks for both representing the problem in the learning activities, namely Cognitive Flexibility Hypertexts, and the products that students would produce to demonstrate their problem solutions: argumentation to justify their choices. This section provides a brief overview of these frameworks.

Problem-solving. There are many types of problems, ranging from those in which all of the constraints, boundaries, and parameters are established and known to those in which it is impossible to even ascertain all of the data that would be required in order to obtain a definitive answer or “solution.” Social stratification and resulting structural inequality can be classified as an ill-structured and highly complex problem that is effectively situated within the domain of problems that sociologists consider central to their domain.

According to Jonassen's problem solving taxonomy (2000, 2011a, 2011b), social stratification is an example of a dilemma. Reorganizing the operational rankings in society affects many different people and groups in different ways, and there is no solution that would please everyone. One's ranking in the social order only carries meaning in terms of the ranks of others, so raising the status of one group could cause another group's status to be lowered. Resource distribution is an issue that affects individuals, families, businesses, governments, and economies as is easily witnessed in debates over federal, state, and local government budgets, and in the way that the resulting tax structure and prices of goods and services affect family budgets.

The idea of "solving" such a problem confounds experts in countless fields every day, so how can students be expected to tackle it in class? First, the problem must be exemplified in a scenario in which the student is given agency to make a specific choice among alternatives that represent the complexity inherent in the problem. However, the student must argue why that choice is the best one using both the evidence provided in the problem and the appropriate concepts and theories from the study of sociology to demonstrate that the student has considered the conflicting perspectives.

Cognitive flexibility hypertexts. In summarizing their respondents' views on critical thinking as a learning objective for introductory sociology courses, Persell et al. (2007) noted that:

For many respondents, the understanding that there are multiple perspectives on any given question or issue was key to their conception of critical thinking. Complex and critical thinking involves approaching social issues and problems with a nuanced view that takes multiple perspectives into account and raises new questions.

This echoes Jonassen's view that ill-structured problems may have "multiple interpretations, perspectives, and solutions" (2011a, p. 208). In school, though, many instructional designs

involve simplifying complex problems in order to make them more accessible to novices. Spiro and his colleagues (Spiro, Coulson, Feltovich, & Anderson, 1988) argue that such simplifications often result in persistent student misconceptions, particularly in contexts that deal with complex systems, such as medicine (in their discussion) or social stratification (in this one) where it is important to understand how action to correct one symptom may cause many others.

How will the introductory sociology student be able to consider the myriad of conflicting perspectives available in a given social stratification problem? One instructional design model that shows promise in encouraging learners to view a problem from multiple perspectives is the cognitive flexibility hypertext in which information is provided in a non-linear fashion as multiple representations of knowledge with multiple connecting points between concepts demonstrating the interrelatedness of the material. A key principle of the cognitive flexibility hypertext is the notion of “criss-crossing the landscape,” in which the design of the learning environment promotes viewing the concepts multiple times from multiple angles so that learners achieve a sense of the complexity and interconnectedness of the material (Spiro, Feltovich, Jacobson, & Coulson, 1992).

However, Spiro and his colleagues (Spiro et al., 1988) promoted Cognitive Flexibility Theory for “learning beyond the introductory stage for a subject area” (p. 375). Introductory sociology students aren’t the same as medical students, after all, and as previously mentioned, few are planning to seek mastery-level knowledge in the field in the future. Does this preclude the use of cognitive flexibility hypertexts for introductory users? Jonassen (2011a) suggests that they can be suitable for a broader range of users, but that the design should sequence cases from simple to complex and choose cases that are appropriate to the experience of the learners.

Argumentation. Since dilemmas such as social stratification, even when presented on a smaller scale (for example, as a decision choice between two individuals), have no definitive

solution, the practice of sociologists in working with social stratification issues involves “social science reasoning”; that is, making arguments to support the decision or recommendation based upon sociological theory and relevant data (Persell et al., 2007). As previously noted, justification is a typical task for ill-structured problems in general and dilemmas in particular (Jonassen, 2000). Jonassen sees benefits to learning for fostering argumentation on both well- and ill-structured problems (see, for example, Nussbaum and Sinatra’s (2003) study on using argument in physics and Chiu’s (2008) in algebra), but argues that it is of particular importance for ill-structured problems “because [they] do not have convergent answers or consistent solution criteria,” thus requiring “learners [to] construct arguments to justify one’s own assumptions, solution paths, and proposed solutions because there are no certain rules and principles to apply and there may be many possible solutions” (2011a, p. 323). Thus, argumentation is an activity that, in the context of a sociological dilemma or decision problem, requires students to think and act like sociologists.

Effective, evidence-based argumentation is a challenging task. According to The Nation’s Report Card (National Assessment of Educational Progress, 2008c), some 40% of high school seniors test at a level lower than “satisfactory” in the persuasive writing task (p. 44), suggesting they may be ill-equipped to handle argumentative writing in college. Interventions that target this shortcoming among first year students could improve their performance throughout their college careers. However, it also suggests that many first year students will need additional guidance about how to structure an argument. Thus, it is important for the design of argumentative learning activities to scaffold students to construct arguments.

Purpose and Impact of the Study

The primary purpose of this study was to examine which of three types of practice activities (summarization of the evidence available for an argument, generation of original components of argument, or generation of components to counter a model argument) resulted in the highest quality of original argument. In part, this study sought to validate and extend the significant results from Jonassen and Cho (2011), namely that students given practice activities in which they generate counterarguments to a provided model argument construct higher quality arguments than those who are not given such a model, by testing them in another domain.

The larger and somewhat more diverse sample from this study afforded the ability to explore whether any significant results differed based on demographic characteristics, such as gender, and the level of education attained by parents of students in the sample. Additionally, because data about revisions were available due to the incorporation of an automated essay grading tool called SAGrader in the design, this study expanded upon the previous research by examining the relationship of revisions to the different treatments and the quality of the participants' final essays.

Results from the study could inform instructional design practice in several ways. Given that a significant number of high school graduates may not be adequately prepared to construct effective academic arguments (National Assessment of Educational Progress, 2008c, p. 44), the efficacy of the instructional design that was tested could suggest ways to increase student practice and scaffold their learning in some of the introductory courses that form a large part of many first year postsecondary curricula. Since cognitive flexibility theory was designed as a model for mastery learning (Spiro et al., 1988), the results of this study could inform its applicability to introductory learning settings. In terms of problem-solving, the study provided a test of Jonassen's (2011a) model of problem-based learning environments for complex, ill-structured

problems in general as well as the applicability of problem-based learning to the specific domain of sociology.

Research Questions

This study examined learning outcomes for students who have used a learning environment and activities designed for this study in their introductory sociology course. The research questions for the study, presented in order of importance, were:

1. Which instructional intervention (summarize, construct, or counter-argue) results in the highest quality of argumentative essays?
 - a. Are there significant differences in performance between genders among the instructional interventions?
 - b. Are there significant differences in performance between first-generation and non-first-generation college students among the instructional interventions?
2. Are there significant differences among the treatment groups in the number of revisions that participants make to their argumentative essays?
 - a. If significant differences exist, what proportion of the overall performance variance is represented by revision attempts vs. the treatment itself?
3. Are there significant differences among the treatment groups in the manner in which they used the case website?

Chapter 2: Literature Review

One of the learning objectives cited by practitioners in the Persell et al. (2007) study as important for introductory sociology students to understand is “the centrality of inequality,” or the ways in which social stratification shapes and is shaped by social structures in which individuals are located. Instructors of sociology have employed a variety of exercises designed to help students understand stratification and inequality in their classrooms. Some have simply tried variations on lecturing about these concepts in the abstract. For example, because many students may not have yet participated meaningfully in the workforce, Brislen and Clayton (2005) superimposed grades (A-F) in place of incomes on the U.S. wealth distribution curve to analogize the concept.

Others have used games and simulations to illustrate these concepts in order to make learning of the material more engaging. Instructors had students work puzzles with varying numbers of missing pieces (Wills, Brewster, & Fulkerson, 2005) and draw pictures using crayon boxes with varying numbers of colors (Wetcher-Hendricks & Luquet, 2003) to make unequal wealth distribution a more visceral experience for their students. Similarly, in another classroom exercise, instructors designed and implemented a coin flipping game among pairs of students to demonstrate wealth distribution (Renzulli, Aldrich, & Reynolds, 2003). Sociopoly (Jessup, 2001) and a successor called USA Stratified Monopoly (Fisher, 2008) use custom-designed rules for the familiar and traditional Monopoly board game to demonstrate mechanisms that maintain social class structures in the US and how those structures contribute to social class stratification. In each of these interventions, instructors had students follow up with class discussion, written reflection assignments, or both to construct meaning from their experiences of playing the games.

Variations on budgeting exercises abound, from using photocopied dollars divided up during the class to groups of students, eventually ripping the last dollar to spread among the

groups designated poorest (Harlow, 2009); to determining the prevailing wage for low-skilled jobs and creating a budget based on that income (Abelev, Vincent, & Haney, 2008); to budgeting for a family from an assigned social class (Hattery, 2003; McCammon, 1999). Each of these instructional interventions is designed to help convey to students the reality of unequal wealth distribution and its effects.

Experiential exercises are also prevalent in the teaching of stratification. Hattery's (2003) students were assigned the direct experience of spending no more than \$4.35 per day on food, navigating on campus in wheelchairs, sleeping in an on-campus "shantytown," and engaging in service learning projects that put them into direct contact with poor and homeless people in the community.

Still other classroom activities have been designed to put them in the role of sociological practice. Some have included direct observation, such as riding an instructor-assigned bus route, taking field notes about the different neighborhoods and communities on the ride, and producing a written paper detailing those observations (Nichols, Berry, & Kalogrides, 2004). Others have analyzed artifacts, such as restaurant menus (Wright & Ransom, 2005), for signals of class distinctions. Instructors who taught both a higher level course and an introductory one had the students from the higher level course model a documentary research project called "True Colors" which exposed racial and ethnic inequality, and then present those findings to the students in the introductory course (Pence & Fields, 1999). Such interventions begin to approach the goal of moving students closer toward sociological practice advocated by practitioners in the field (Atkinson & Hunt, 2008; Persell et al., 2007), but, with the possible exception of the students in the advanced course in Pence and Fields (1999), they are not focused on problem-solving practices.

The remainder of this chapter discusses the theoretical and empirical frameworks that governed both the instructional design and the experimental design for this study. The

instructional design represents a series of problem-solving activities incorporated in a cognitive flexibility hypertext and a web-based essay assessment tool known as SAGrader with the goal of fostering argumentation skills for addressing dilemma problems in the domain of sociology. The chapter is organized in sections covering each of these key constructs: problem-solving, cognitive flexibility hypertexts, SAGrader, and argumentation. In addition, prior research into the specific demographic groups examined in the study is discussed.

Problem-Solving

In order to think like a practitioner in any discipline, it is not sufficient to simply memorize the salient facts and procedures typical of the domain, but rather to understand what types of problems are represented in the discipline and how its practitioners apply their domain knowledge to solve them (Jonassen, 2000). Thus, activities which seek to encourage introductory sociology students to think like sociologists must allow them to apply the facts and concepts from lectures and textbooks to solving the types of problems that sociologists typically work on.

As noted in the previous chapter, Jonassen (2000, 2011a, 2011b) would classify social stratification as a dilemma because it is ill-structured, complex, inter-disciplinary, and cannot be “solved,” per se because it involves a multitude of conflicting perspectives that cannot fully be addressed. Instead, “solutions” take the form of decisions that must be justified using available evidence to demonstrate that the learner has weighed the conflicting perspectives.

While instructors of sociology have employed problem-solving approaches in their classrooms previously (e.g. Ross & Hurlbert, 2004; Sernau, 1995), there is scarce empirical research examining the learning outcomes of these approaches. Ross and Hurlbert (2004) report only student satisfaction data when reporting on their civil unions problem-based learning exercise. However, they do assert that the exercise met their “substantive goals for the course, including the examination of family diversity, an analysis of defining the term family, and the inter-connected nature of social institutions at a macro-level analysis” (p.88) and discuss the

components of the student work from within the exercise that exemplify those goals. They also claim that the exercise fostered active learning, information literacy, and increased awareness of “the concerns of a subordinated group in the United States” (p. 88), though again these outcomes are based on the observation of the instructor rather than empirical analysis.

Sernau (1995) likewise focuses on the content of the problem-based learning activity in his Social Stratification course and describes its success in terms of student satisfaction from course evaluations. The exercise itself is very near to engaging in authentic sociological research practice in that it involves a research project which includes both a literature review and observations or interviews from the field along with related writing assignments. Sernau does not discuss whether the majority of his students are sociology majors, though he notes that some of his former students “have received undergraduate research grants, have made presentations at regional conferences, and have continued their research in graduate school” (p. 372). This observation is anecdotal rather than empirical, but it reflects the instructor/author’s view that the problem-based activity he has designed helps prepare students for such roles.

A qualitative study of high school students in a Social Studies course (Osana, Tucker, & Bennett, 2003) examined the justifications students gave for their responses in force-decision problems focused on equity, however, the purpose of the study was to examine adolescent decision making rather than the effects of the activity on learning outcomes. Osana et al.’s results showed that the subjects had insufficient maturity and depth in their conceptions to address the complexities of such an ill-structured problem. The students tended to arrive at relatively simple solutions and consider the objectives met; they did not “actively generate alternative distribution schemes and thus had few or no options to choose from” (p. 378).

When it comes to problem-solving research studying ill-structured problems outside of sociology, most of the empirical studies focus on argumentation, which will be covered in detail later in this chapter. However, studies have been done on scaffolding and supporting students’

efforts solving ill-structured problems which have informed aspects of the design of the learning environment used in this study.

Notable among these is the study conducted by Ge and Land (2003) to assess relative impacts of question prompts and working with peers when solving an ill-structured problem. Participants in the study were assigned to one of four treatment groups. One group worked on the problem individually and without question prompts. Another worked individually but was given the prompting questions. A third group worked with peers but without question prompts. The fourth group both worked with peers and received question prompts. The students' solution reports were assessed by three raters. Students in both question prompt groups outperformed those in the groups that did not receive the questions at statistically significant levels. The peer groups did not show statistically significant results compared to the groups working individually overall, though they did outperform the students working individually on an individual measure of problem representation. Qualitative data, in the form of observations and interviews, was also gathered on a sub-sample of the participants. Analysis of that data revealed that the "had an effect of directing student attention to important information they might have overlooked, thus facilitating awareness of what is known and not known" (p. 31). The questions also prompted students to "articulate why they had selected particular solutions" (p 32). The students in the peer groups expressed that they benefitted from each other's perspectives while working in groups, though observation data showed "that there tended to be more agreement than disagreement among the members, and few constructive suggestions were made to each other" (p. 33).

Similarly, Bixler and Land (2010) explored the use of cognitive ("focusing on selecting or organizing concepts") and metacognitive ("focusing on reflection and self-monitoring") prompts (p. 5) in an online learning environment for an undergraduate information technology course. The study employed a post-test to compare participants who worked in the environment with the prompts versus a control group that worked in the environment without the prompts.

Their performance used the same rubric as the earlier studies by Ge (2001) and Ge and Land (2003) to assess the problem solution reports submitted by the students. Students in the treatment group outperformed those in the control group at a statistically significant level on all four individual measures of problem solving performance (problem representation, developing solutions, making justifications, and monitoring and evaluation), with problem representation showing the strongest statistical effect of the treatment. The study did not compare the effects of cognitive versus metacognitive prompts.

Shen (2008) also compared a treatment group that received question prompts with a control group that did not in an online learning environment associated with an ethics course for undergraduate engineering majors. Like the previously mentioned studies, this experiment observed a significant improvement in total essay scores for the treatment group over the control. However, in an activity in which students were asked to generate questions relevant to the case they were solving, there was no significant difference between the group that received links in the form of question prompts and the group that received traditional descriptive links.

Another support that has been studied in various ill-structured problem solving contexts is providing a library of cases that solvers may access in their work. Hernandez-Serrano and Jonassen (2003) compared students who had access to a library of cases (stories of professional experiences related to the issues in the problem compiled from practitioners in the field) with students who had access to “fact sheets (expository descriptions of the issues contained in the stories, with episodic information deleted)” (p. 107). Using a pre-test and post-test comprised of multiple choice and short answer questions, they found a statistically significant performance improvement among the participants who had access to the case study library compared to participants who only had the fact sheets, suggesting that the contextualization of the relevant information in story form was helpful to student performance. Thus, while the design in the

current study does not incorporate a case library, it uses this idea of contextualizing subject matter content within the framework of a story or narrative.

Cognitive Flexibility Hypertexts

As noted in the introduction to this study, social stratification is a complex, ill-structured problem with a myriad of conflicting perspectives. Even the simplest example of a choice to rank a pair of individuals reveals underlying structural inequalities that result from the ranking process. Cognitive flexibility theory (Spiro et al., 1988) stresses a focus on giving due consideration to the connections and relationships among different components of such complex systems so that solutions will not suffer from “reductive biases” (p. 376); that is, misconceptions that result from oversimplification.

Cognitive flexibility hypertext is an instructional design model for representing a problem through multiple representations of knowledge with multiple connecting points between concepts demonstrating the interrelatedness of the material. By using a design which promotes viewing the concepts multiple times from multiple angles, learners gain a sense of not only the concepts, but also of how they relate to and affect one another (Spiro et al., 1992). In fact, Jonassen derives his model of using “cases as alternative perspectives” in problem-solving learning environments (2011a, pp. 208-221) from cognitive flexibility theory, arguing that it is particularly suitable for “conceptually complex, ill-structured problems” (p. 209).¹

Research on cognitive flexibility hypertext environments has shown some strength in its application in different ill-structured domains. In a study conducted with paid undergraduate volunteers given problems about technology’s impact on society and culture, Jacobson and Spiro

¹ In the interests of full disclosure, it should be noted that an earlier version of the learning environment employed in this study is pictured in *Learning to solve problems: A handbook for designing problem-solving learning environments* (Jonassen, 2011a) as an example of a problem-solving learning environment modeled on cases as alternative perspectives and designed to foster argumentation.

(1995) implemented two control conditions in which participants were given traditional drill activities while the experimental group received a treatment the researchers called “Thematic Criss-Crossing Hypertext” (p. 305). The two control groups were differentiated by the instructions they received during the initial task of reading the case material. One group was given the same instructions and content as the treatment group, which included mentioning that each case had multiple themes, while the other was instructed to focus on the most relevant theme for each case. As the researchers had hypothesized, the treatment group outperformed both control groups on the problem-solving essay assessments that participants wrote during their second and fourth sessions. In addition, the treatment group’s mean score on the second essay improved from that on the first while the control groups’ mean score actually declined from their first essay to their second. However, and also as hypothesized, their control groups performed better on recall tests than did the treatment group, suggesting that such environments may not be appropriate for applications in which traditional recall forms of assessment will be used.

Researchers in another study (Demetriadis & Pombortsis, 1999) used a multimedia electronic book for the control condition with a cognitive flexibility hypertext as the experimental one in a set of learning activities about computer networking. Demetriadis and Pombortsis employed a pretest-posttest-delayed posttest research design in which each of the posttest instruments contained a problem solving question which the researchers called “flexibility items” in their description. The results showed no significant differences between the control and treatment groups for basic knowledge acquisition, but did show significantly better performance on the flexibility items by the treatment group compared to the control. These results reinforce those of Jacobson and Spiro (1995) in the sense that the learning exposed by such treatments suggest that appropriate assessment instruments are necessary in order to determine learning outcomes from cognitive flexibility hypertext environments.

Cernusca (2007) explored the efficacy of a cognitive flexibility hypertext design focused on biblical criticism in multiple iterations in a design based research study. In one of the iterations, student performance after using the hypertext environment was contrasted with their performance after an instructional module comprised of lectures and small group discussions. Mean essay scores for the cognitive flexibility hypertext activity were higher than those for the traditional classroom activity at a statistically significant level. However, the students had been divided into four groups based upon different methods of biblical criticism, and one of those groups had lower mean essay scores than the other three groups at a statistically significant level. The researcher interpreted this to represent a weakness of the instructional design for that particular activity. However, since the next design phase involved an entirely different set of activities, it is not possible to determine possible causes for the poor performance of that particular group.

Dabbagh (2002) explored the use of heterarchical versus hierarchical structures in a web-based representation of an ill-structured problem space in instructional design. A heterarchical design may better meet the needs for cognitive flexibility theory for “rearranged instructional sequences, for multiple dimensions of knowledge representation, for multiple interconnections across knowledge components, and so on” (Spiro et al., 1992, p. 67). In Dabbagh’s (2002) study, one group of students used a heterarchical version of the problem representation while another group worked from a hierarchical version. All of the information within each representation was the same; only the navigational structures differed between the two versions. A simultaneously running application gathered responses from users immediately after they made navigational selections. In addition, Dabbagh observed the two groups of students as they worked through the problem and used the website and then interviewed each group after they finished. The heterarchical group visited fewer links than the hierarchical group but spent more time on task overall and a longer time on each individual page that they visited. Students using the

heterarchical site also tended to visit pages that they were not sure would be useful about twice as often as the group using the hierarchical site. Both groups categorized similar links from the sites as the most useful in their work. Data about learning outcomes cannot be considered generalizable since the researcher only compared two problem-solving groups and did not use rigorous assessment methods to evaluate the problem solutions, but the group that used the heterarchical website structure articulated the issues of the case and their recommended solution more clearly than the group that used the hierarchical structure according to the researcher.

In summary, cognitive flexibility hypertexts have empirically demonstrated positive effects on learning outcomes in ill-structured domains when essays are used as the assessment instrument. Significant effects have not been observed on instruments designed to assess basic recall information; in fact, control groups sometimes outperform treatment groups on such instruments.

SAGrader and Revisions

SAGrader (Ideaworks Inc., 2008) is a commercially available system for automatically grading student essays. The scoring engine uses a combination of “fuzzy logic, ... a semantic network, ... and rule-based expert systems” (Brent, Atkisson, & Green, 2010, p. 5) to score essays for a variety of popular course content. The scoring engine can be customized to accommodate new courses.

A key affordance of SAGrader is its ability to give students feedback on successive drafts of their assignment responses and allow them the opportunity to revise based upon feedback. At the instructor’s discretion, students may be limited to a specified number of revisions or allowed to revise at will until the assignment due date.

MacArthur (2007) asserts that revision is important both because skilled writers employ it extensively in their writing process and because it provides additional opportunity for instructors to provide feedback which increases what students learn about writing. Revision is a key aspect

of the cognitive framework of the writing process proposed by Flower and Hayes (1981). The researchers used a “think aloud” protocol in which participants described what they were thinking as they composed an essay, and found that the major components of the writing process planning, which included tasks such as setting goals for the essay, generating ideas to achieve those goals, and organizing the ideas; translating, which involved deciding how to word the ideas and put them on the page in the form of sentences and paragraphs; and reviewing, which encompassed evaluating what they had already written and revising it to better meet their goals.

This framework has been used by most of the prominent authors in the field of writing research, including Rijlaarsdam and van den Bergh (2006), who extended it by observing in their own research that much of the variance in the quality of writing product could be explained by the frequency of each of the activities, including revision, provided that the model also accounted for when in the process each of the activities occurred:

It turned out that none of the activities were effective during the whole process; some contributed during one or two phases in the same direction (positively or negatively), others, in reverse directions: positively in one phase, negatively in another (“Observation 1: What Constitutes a Writing Process?”).

Specifically with regard to revision, the research group found (van den Bergh, Rijlaarsdam, & Breetvelt, 1993) that three processes they considered to be part of the revising process had positive effects on the quality of the written product: “rereading the last passage, evaluating a passage, and changing of (a part of) a sentence,” while three others had negative effects on the writing outcomes: “rereadings of the whole previous text, evaluation of formal aspects, and changes of words” (p. 145).

Their finding that word-level changes negatively influenced quality is somewhat inconsistent with an earlier study by Bridwell (1980) in which analysis of multiple revisions among the essays of twelfth graders found positive relationships with final quality for changes to what she classified as “lexical level” (p. 203) and above, but negative relationships with what she characterized as “surface level” (p. 203), such as grammatical, spelling, and punctuation changes.

The participants in her sample who “revised most at the word level tended to do so at other levels” (p. 211), which may account for some of the difference in her findings compared to van den Bergh, Rijlaarsdam, and Breetvelt (1993). In a review of prior research on revisions, Fitzgerald (1987) notes that several studies report that surface level revisions predominate among all ages, but that older students tend to make more higher-level revisions than younger ones. Thus, the fact that the students in the van den Bergh, Rijlaarsdam, and Breetvelt (1993) were ninth graders while those in the Bridwell (1980) study were twelfth graders could further explain the inconsistency in their results.

With few exceptions, the studies that Fitzgerald (1987) reviewed found a positive correlation with revision and product quality. Studies of students younger than high school age found either no improvement or even degradation of quality with successive revisions. Additionally, one study she cited found that “revisions of unskilled college writers often resulted in worse drafts” (p. 493). Studies of revisions performed after feedback or instruction rather than students’ own decisions of whether to revise were generally correlated with improved quality, though one study found no improvement.

Among twelfth graders who took the writing assessment for The Nation’s Report Card (National Assessment of Educational Progress, 2008c), those who reported that their teachers had asked them to write more than one draft of paper assignments (National Assessment of Educational Progress, 2007) showed significantly higher scale scores on the writing exam than those who reported that they had not previously been assigned drafts. The effect was even greater among students who reported that they “sometimes” or “almost always” revised their papers.² This is consistent with findings from earlier implementations of the test and survey as cited in

² The detail on scores based upon answers to the Student Background Questionnaire are not available in the main report. These percentages were obtained from the NAEP Data Explorer (National Assessment of Educational Progress, 2008a) at <http://nces.ed.gov/nationsreportcard/naepdata/>.

Fitzgerald (1987). This data demonstrates that, whether instructor-mandated or simply a choice made by individual students, learners who revise papers before submitting them will generally be better writers than their peers who do not engage in revision.

It is not surprising, then, that students who use SAGrader's feedback to revise their essays improve their scores with successive revisions. Ideaworks has published whitepapers indicating that students improve their scores an average of 20% from first submission to last based on SAGrader feedback (Brent, Carnahan, & McCully, 2006b) and that statistically significant gender differences found in first submissions with SAGrader are reduced to a statistically insignificant level with final submissions (Brent, Carnahan, & McCully, 2006a), suggesting that the system provides useful feedback to students who use that feedback to revise their essays. A later study (Brent et al., 2010) showed a 21% improvement in mean scores for all essays from first draft to final submission and a 30% improvement among essays with one or more revisions. Both figures were found to be statistically significant using t-tests.

Argumentation

Recall from the introductory chapter that incoming college freshmen may not be adequately prepared to perform well on argumentative essays (National Assessment of Educational Progress, 2008c, p. 44). However, argumentation is a part of practice for sociologists (Persell et al., 2007), and for a variety of other disciplines that deal with complex, ill-structured problems in general and dilemmas in particular (Jonassen, 2000).

The ability to effectively consider alternative perspectives is, as previously noted, central to the ability to effectively solve dilemma problems. While higher levels of education may moderate the effect somewhat, a series of studies has shown that people may have a tendency to see arguments that support their own positions on a given topic as stronger than arguments to the contrary point of view (Kuhn & Udell, 2007). These researchers conducted one study with participants from a middle school, a high school, and a graduate school to see whether they would

be more likely to choose warrants that strengthened the argument they were instructed to make rather than those that weakened the opposing argument. While their results were not statistically significant, they noted a modest trend toward increased likelihood for choosing warrants that weakened the opposing argument from middle school to high school to graduate school. It also cannot be concluded whether this modest trend correlates with age rather than education level because no participants from outside an educational environment were studied. In a second study, graduate students were divided into two groups. One group was provided the argument and a warrant which supported it while the other was provided the argument and a warrant which weakened the opposing view. Both groups were then asked to rank the strength of the warrants. No statistically significant difference was found between the groups, though the group viewing warrants which supported the argument gave modestly higher ratings ($M=2.86$) than the group viewing warrants weakening the opposing argument ($M=2.49$). In a third study, the researchers attempted to address the question of whether education or age was the primary influence on the results of their first study by replicating it with sample selected from a community group. In their report, they do not detail a comparison of these participants with regard to education though their description of the group as including “several doctors and lawyers, as well as librarians, sales clerks, office workers, homemakers, and retirees” (p.96) suggests that there would have been differing education levels within the group of participants. Instead, they note that the overall performance of this mixed group was comparable with the high school group in the first study ($M=2.83$ vs. $M= 3.04$, respectively). Again, the results were not statistically significant.

The instrument in Kuhn and Udell’s (2007) first three studies represented a forced choice scenario and, as the authors note, the choice involved did not necessarily represent desired performance in an actual argument. Consequently, they carried out a fourth study in which participants were given the same argument positions as in the preceding studies, but instead of choosing between given warrants, they were asked to respond to the question, “What is the best

argument to make?” (p. 97). Three groups were included in the study: a middle school group, a university group, and a community group. Students in the middle school group were given the opportunity to seek guidance on the first question about what they were expected to do, but otherwise received the same treatment as the adult groups. The adult groups were more likely to address the opposing argument at a statistically significant level when compared to the middle school group. In order to assess whether younger students did not provide counterarguments to weaken the opposing position due to a lack of knowledge about how to do so, Kuhn and Udell conducted a fifth study with middle school students in which they used the instrument from their fourth study, but explicitly instructed the participants to generate counterarguments. Most of the participants were able to do so, though they did sometimes “also [include] arguments supporting the favored position, suggesting that students of this age do not distinguish sharply between the two kinds of arguments” (p. 101).

Inexperienced learners tend to use superficial criteria to critically evaluate the arguments of others (Brem, Russell, & Weems, 2001). In a study of 81 female high school students involving the critique of web sites, the researchers instructed participants to critique “both the scientific argument [being made on the site] and the reporting conduit [the credibility of the site’s authorship or ownership]” (p. 201), but the vast majority of the critiques failed to report their critique of the latter. Those who did tended to rely on criteria such as affiliations, titles, and related credentials; to assume that sites with many authors were more credible than sites with single authors; and to assume that scientific degree credentials precluded motives other than pure science for presenting information. Students also tended to judge the validity of claims on the assigned web sites based on the amount of support the sites gave for their claims rather than the specifics of what that support contained.

Individual characteristics, such as epistemological beliefs, assertiveness, and extraversion, can also impact the level of learner willingness to engage in argumentation

(Nussbaum & Bendixen, 2003). Based on the responses of 238 undergraduates to a series of instruments designed to measure argumentativeness (tendency to approach an argument or avoid one), need for cognition (enjoyment of “effortful and complex thinking” (p. 581)), epistemic beliefs, and extraversion, the researchers found that participants with a tendency to believe that knowledge is certain and those who tended to believe that the most robust knowledge is that which is most simple and least confusing were the participants most likely to score high on the tendency to avoid argumentation. Participants with high scores on the need for cognition instrument tended to score highly on tendency to approach an argument, though this was most likely to be true when they also scored high on an assertiveness measure from the extraversion scale, and low scores on the instrument did not demonstrate a strong relationship with argument avoidance. Instead, participants with low levels of assertiveness and/or high levels of warmth were most strongly associated with argument avoidance.

If we want introductory sociology students, who may bring a variety of differences in personal traits, epistemological beliefs, and experience with argument, to succeed in thinking like sociologists in an argumentation activity, we must provide guidance in what constitutes a valid sociological argument. Previous studies have examined the use of constraints during argumentation activities (Cho & Jonassen, 2002) as well as the impact of various types of goal instruction (Nussbaum, 2005; Nussbaum & Kardash, 2005) on the resulting quality of student argumentation.

Cho and Jonassen (2002) found that groups who were constrained by a system which required them to classify their messages according to categories describing what the message intended to contribute to the discussion (hypothesis, data, principles, or unspecified for initiation messages and for, against, or “and” for response messages) developed arguments with more components of argumentation (specifically, claims and grounds) than groups not constrained by the treatment system. They also found that ill-structured problem-solving groups produced more

argumentation components than the groups solving well-structured problems. While such a methodology is not directly informative for situations in which students will be developing individual argumentative essays rather than engaging in argumentative discussions with one another, Cho and Jonassen did find that their results have an effect on transfer to just such an individual essay assignment where the groups that had used the treatment system outperformed those who had not. This suggests that providing some form of training or scaffolding about the components of good arguments may help students to produce those components in their essays.

Similarly, Nussbaum (2005) found that providing goal instruction that directed students toward focusing on a particular aspect of argumentation (or to simply “explore” the issue, for the control group) in an online discussion had significant effects on generating claims for two of the conditions studied (“persuade others” and “provide as many reasons as you can”), though the counterargument condition did not. In collaboration with Kardash (Nussbaum & Kardash, 2005), Nussbaum performed a pair of studies in which similar instructions were given to participants who would write argumentative essays rather than engage in online discussions with their peers:

In one condition (*control condition*), participants were simply asked to write their opinion. In a second condition (*reason condition*), participants were asked to write their opinion but also asked to write as many reasons as they could think of that supported their opinion. The third condition was identical to the reason condition except that participants were additionally requested to produce both counterclaims (“reasons others may disagree with you”) and rebuttals (“why those reasons are wrong”). The third condition was labeled the *counterargue/rebut condition* (p. 159).

Among these groups, the third condition demonstrated significant effects in all measures other than supporting reasons for counterclaims while the other treatment condition did not show statistically significant differences from the control at all. In the second study, students in the control section were again asked to express an opinion in the form of an essay on the provided question while members the treatment section were instructed to each write a persuasive letter to their Congressional representative. Each of these sections was also split into two groups, one of

which received a document derived from participant responses in the first study listing pro and con arguments about the question. The instruction to persuade had a statistically significant negative effect on the resulting holistic scores on the essays and also on the number of counterclaim reasons generated. However, the holistic score effect was mitigated among the group who had received the document containing sample arguments. Overall, participants who received the document generated more counterclaims and fewer rebuttals than those who had not received it, but these effects disappeared in the presence of the instruction to persuade.

Of particular relevance to this research is a series of studies conducted in an ethics course for undergraduate engineering majors (Jonassen & Cho, 2011; Jonassen, Cho, Easter, Henry, & Kwon, 2010; Jonassen, Cho, Kwon, Henry, & Easter, 2009; Jonassen, Cho, Kwon, Henry, Shen, et al., 2009; Jonassen, Shen, et al., 2009; Shen, 2008) which formed the foundation of both the instructional design and research methodology described in the next chapter. A cognitive flexibility hypertext website was designed called “Engineering Your Ethics,” or E.Y.E. The website was centered around cases written in story format that represented ethical dilemmas for engineers. In each study, one or two teaching cases provided links to pages summarizing how ethical canons from the National Society of Professional Engineers and three different general theories of ethics would each apply in the given scenario. In addition, links to pages explaining the perspectives of various key players in the case were provided. The teaching cases were accompanied in each iteration of the study by a practice case and a transfer case, the latter of which served as the assessment for the activity.

In the first iteration of the study (Jonassen, Shen, et al., 2009; Shen, 2008), two teaching cases were provided. For the control group, the links were as described above. For the experimental group, the links took the form of question prompts as previously discussed in the section on cognitive flexibility hypertexts. In the practice case for this iteration, instead of links to application pages for perspectives, canons, and theories, students were asked to enter their own

interpretation of how these would apply via provided text boxes. Once they had submitted their own answers, they were provided answers from experts about how the perspectives, canons, and theories would apply to the case. For the transfer case, “students were asked to generate a list of questions that would be necessary to resolve the case, and then they wrote an essay to answer those questions” (Jonassen, Shen, et al., 2009, p. 240). In both groups, holistic essay scores rose with the number of clicks in the learning environment. In addition, participants who generated more questions tended to score higher on the assessment. The treatment group generated more questions that reflected argumentation components (i.e. “multiple perspectives, ethical theories, ethical canons, multiple solutions, or decisions” (Jonassen, Shen, et al., 2009, p. 244)) and more questions representing higher-order thinking skills as scored based on Bloom’s Taxonomy (Bloom, 1956).

In a follow-up study (Jonassen & Cho, 2011; Jonassen, Cho, Kwon, Henry, Shen, et al., 2009; Jonassen, Shen, et al., 2009), the researchers deployed three versions of the E.Y.E. learning environment. The teaching case for each version contained training activities for the students to complete. The control group was instructed to summarize the perspectives, canons, and theories at play in the teaching case. One treatment group, called the evaluate group, was given a series of questions designed to have them choose between two provided solutions and give evidence from perspectives, canons, and theories to support their conclusions. A second treatment group, designated the construct arguments group, was not given proposed solutions; instead, the questions they received prompted them to construct their own solution and accompanying support for the case. The second case, then, was used as a near transfer assessment and the third as a delayed transfer assessment. In each of these cases, all groups were instructed to construct a solution and provide justification for that solution. Using a revised rubric, the essays were scored for holistic quality. There were statistically significant differences between each of the two treatment groups and the control group for the near transfer case, but the two treatment groups did

not differ significantly from one another. Further, there were no significant differences among any of the three groups for the delayed transfer case.

A subsequent iteration of the study (Jonassen & Cho, 2011; Jonassen et al., 2010) focused on generating counterarguments, which were generally found to be low among previous participants. In this iteration, two treatment groups were used: one in which students constructed arguments just as the construct group had done in the prior study, and one in which the environment for cases 1 and 2 provided an exemplar student argument from the previous study and instructed the students to generate counterarguments to the exemplar. The third case served as the posttest for both groups and instructed them to write an argumentative essay about the scenario. Here, a simple three point rubric was used to denote the quality of counterarguments (underdeveloped, partially developed, or well-developed) for the first two cases. The holistic rubric from the previous study was used to assess the essays generated in case 3. There were no significant differences in the number or quality of counterarguments between the groups for either case 1 or case 2 activities, which were completed in the same class period, but the treatment group that generated counterarguments to the exemplar did outperform the group that constructed an argument to a statistically significant degree when the cases were combined, both in terms of number of generated counterarguments and the quality of the counterarguments that were generated. The counterargument group also outperformed the construct argument group on the holistic essay score in the third case, though there was no significant difference between the groups in the number of counterclaims (or other components of argumentation) in those essays.

In the final study in the series (Jonassen & Cho, 2011), four cases were used to foster rebuttal arguments. The first case was used as a pretest. In the second case, three different treatment groups were used. The construct group was asked to make a decision and then generate counterarguments to their own decision. The example group read two alternative arguments for each perspective, canon, and theory that countered one another. The combined group was given

one example argument for each perspective, canon, and theory and asked to provide a counterargument to those examples. All three groups were then guided to make rebuttal arguments. Rebuttals were assessed using a five point (0-4) quality rubric. Cases 3 and 4 were used as immediate and delayed posttests respectively and all three groups wrote argumentative essays in response to each case which were assessed according to the holistic rubric. The example group generated higher quality rebuttals at a statistically significant level compared to the other two groups. However, the construct group performed better on the immediate posttest than the other two groups. There were no significant differences among the groups on the delayed posttest.

Demographic Considerations

The current study proposes to extend the research conducted in the engineering ethics studies by situating a similar learning environment in a different context. The introductory sociology course differs from the engineering ethics course not only in subject area, but also in serving as a core curriculum elective taken by students in a variety of majors. Consequently, the new context not only affords the ability to explore whether similar learning activities succeed in a domain other than engineering, but also to examine whether any observed effect differs among demographic groups. In particular, the study proposes to examine effect differences between genders and between first generation students and their counterparts who have parents with college degrees.

Gender differences in argumentation. According to The Nation's Report Card (National Assessment of Educational Progress, 2008c), girls at the both the eighth and twelfth grade levels continue to outperform boys at statistically significant levels in the writing component of the assessment. While the score gap between males and females at the twelfth grade level shows as having narrowed to a statistically significant degree in the 2007 data compared to the 2002 results, the gap remains high (p. 40). Male students also have a higher incidence of scores below satisfactory on the persuasive writing category of the assessment. In

2007, 48% of twelfth grade boys who took the exam scored below satisfactory on the persuasive writing component compared to 32% of girls.³ The figures from the NAEP suggest that, of the 40% of high school seniors not adequately prepared to write argumentative essays as college freshmen (p.44), a disproportionate share of them are male. Thus, any interventions that particularly help male students to improve their scores could narrow this gap, while any interventions that particularly improve the scores of female students could widen it.

In the engineering ethics studies, there were not enough female participants to expose any differing effects on the various interventions between males and females. In the other previously cited studies, participants included a mix of genders, but gender differences were not examined in the analysis of the results. Prior research has exposed differences in dialogic argumentation performance between male and female participants, however, which raises a question of whether such differences occur in monologic argumentation exercises like those in this study.

Differences in dialogic argument between the genders have been demonstrated from an early age. For example, Howe and McWilliam (2001) studied gender and socioeconomic status (SES) differences among preschoolers in the UK in the context of play activities in nurseries. They coded verbal and nonverbal acts among the preschoolers as types of antecedent (the original request that sparked the disagreement), opposition (the form of the response to the request), or development (subsequent exchanges among the participants related to the same disagreement). They found that, in general, boys engaged in arguments more frequently than girls, especially when in groups of all boys. Gender differences were also associated with the complexity of the arguments, but those differences reversed between lower and higher SES groups. Among the lower SES groups, boys engaged in more complex arguments than girls, but higher SES girls had

³ The figures about performance on persuasion are not broken out by gender in the main report. These percentages were obtained from the NAEP Question Tool (National Assessment of Educational Progress, 2008b) at <http://nces.ed.gov/nationsreportcard/itmrlsx/search.aspx?subject=writing>

higher complexity levels than higher SES boys. The authors suggested that teachers' attention to the differing argumentation styles could potentially reduce these differences over time.

Recall from earlier in this chapter that Kuhn and Udell (2007) showed modest effects of education levels on the sophistication of argumentation skills, though they did not look at gender differences within their samples. Another set of researchers (Cala Carrillo & De La Mata Benítez Maria, 2004) examined the effects of gender and level of education on discourse and argumentation. By coding the verbal interactions of same-gender discussion groups of varying levels of educational attainment, researchers found that women at the "literacy" (basic) level of education used justifications more frequently than those at higher education levels. This would seem to preclude the theory offered by Howe and McWilliam (2001) that education would reduce differences between the genders over time. The result is also surprising given that, overall, groups at more highly educated levels of the study used justifications more frequently than the less-educated participants, echoing the modest education-level effects that Kuhn and Udell (2007) observed. However, it must be noted that the topics given to the participants in Cala Carrillo and De La Mata Benítez Maria's study (2004) involved "women's work" and "children's education" (p. 409). Given that interviews of participants showed that those at the "literacy" level came from households in which the division of housework chores was most unequal, the researchers concluded that the male participants in this category did not see a need to justify their position on "women's work."

Jeong and Davidson-Shivers (2006) designed a study to examine gender differences in online collaborative argumentation among students in a graduate level university course. Here, then, the participants had comparable levels of education to one another, perhaps slightly beyond the most educated participants in the Cala Carrillo and De La Mata Benítez Maria study. Students

were provided with a constrained set of message categories based on the Toulmin model⁴ and earlier research by the lead author. They were instructed to restrict each post they made to reflect only one of the categories and to label the posts with the appropriate category. In analyzing the data from the posts, the researchers did not find statistically significant gender differences in the frequency at which they challenged the claims of their opposite gender classmates, the mean number of critical responses to each other's posts, or the types of messages posted. They did, however, discover through post hoc tests that females in their study posted fewer rebuttals to the arguments of other females compared to the rates at which males rebutted other males, females rebutted males, and males rebutted females "with moderate effect sizes and *p* values approaching significance" (p. 562). The researchers posited that a larger sample size would be needed in order to further these findings. Jeong (2007) did so by combining the results from a similar exercise set in four different offerings of a different graduate course over a period of two years. With the increased sample size, he found that the number of rebuttals posted by male students was higher than that of female students at a statistically significant level. In addition, this result reflects the number of rebuttals overall, without regard to the gender of the classmate being rebutted. Consequently, this finding may have somewhat more applicability to monologic argumentation scenarios than Jeong and Davidson-Shivers (2006).

First generation college students. First generation college students have long been identified as a population that may be underserved in higher education settings. Specific federal programs have been targeted toward this population under the auspices of the Economic Opportunity Act of 1964 and the Higher Education Act of 1965, Title IV, Part A (Office of Postsecondary Education, 2011). From a sociological perspective, if higher education is to serve

⁴ The Toulmin model, described in *The Uses of Argument* (1958, 2003), established the components of argumentation used in many of the studies cited in this literature review, such as claims, warrants, rebuttals, etc. While the current study was not built directly on the Toulmin model, it is cited here as an indirect influence via other studies in which it was foundational.

as a mechanism for promoting upward social mobility (Pike & Kuh, 2005), one measure of how well it achieves that goal could be the socio-economic diversity of the student population (Bowen, Kurzweil, Tobin, & Pichler, 2005). Since higher levels of educational attainment are generally correlated with higher earnings (Ewert, 2012), first generation students may disproportionately come from lower socio-economic backgrounds than those students whose parents have obtained college degrees (Choy, 2001). Completing a degree when one's parents did not, then, could logically provide an avenue for one to move from the parents' socio-economic class to a higher one.

The National Center for Education Statistics has issued two extensive reports about first generation college students from a series of longitudinal studies (Choy, 2001; Nunez & Cuccaro-Alamin, 1998), both of which highlight lower persistence (i.e. students who had either received a degree or were still enrolled at the timing of each follow up) and bachelor's degree completion rates among first generation students compared to their peers. Further, first generation students represented a larger proportion of students who failed to complete their first year of undergraduate studies than students whose parents had obtained at least a bachelor's degree, even when controlling for other variables (Choy, 2001).

First generation students also perform lower than their compatriots in the writing component of The National Report Card assessment (National Assessment of Educational Progress, 2008c). Among students assessed in 2007, students with at least one parent who had some education after high school had an average scale score of 152 (p. 41), which is nearly identical to the overall national average of 153 (p.36). Students with at least one parent who graduated from college averaged 163, while students whose parents had only finished high school scored 141, and students who did not have at least one parent with a high school diploma had an average scale score of 134 (p. 41). The average scores for first generation students fall at the low end of the "basic" level of achievement reflected in the assessment (p. 37), suggesting that these

students tend to lag behind their peers in terms of adequate preparation and readiness for college-level writing.

Drawing on these statistics as justification, much of the literature about first generation students has focused on interventions to increase access to higher education or on retention efforts to improve completion rates of first generation students. While access was less of a direct concern in this study because the participants are already enrolled in postsecondary degree programs, evidence of skill and academic preparation gaps for first generation students as compared to their peers (e.g. Carter & Robinson, 2002; Choy, 2001; Pascarella et al., 1995) suggest that the effectiveness of given learning interventions may differ between these groups.

While little has been written about the performance of first generation students in specific learning interventions, the existing literature on retention efforts has shown some tendencies that inform learning outcomes. For example, such students are much less likely to network effectively with peers than students whose parents had college degrees (Carter & Robinson, 2002; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006; Pike & Kuh, 2005) or to seek out assistance when needed (Collier & Morgan, 2008; Nunez & Cuccaro-Alamin, 1998). First generation students, then, may benefit from more frequent feedback, particularly computer-generated feedback which would not involve seeking out other people to provide assistance. Such feedback could provide guidance about deficiencies in a low-risk fashion, allowing students who are reluctant to seek help to understand the shortcomings in their performance and correct them before turning in assignments for grades.

First generation students may also benefit more than their non-first generation peers from engaging in more effort in academic activities. Pascarella and his colleagues (Pascarella, Pierson, & Wolniak, 2004) discovered that “hours studied had a stronger positive effect on critical thinking, and number of term papers or written reports had stronger positive effects on writing skills...for first generation than for other students” (p. 274). The researchers found that first

generation students also benefitted more from courses in “the arts and humanities, mathematics, the social sciences, and the natural sciences and engineering than did other students” (p. 275) in outcomes such as writing skills, while demonstrating negative effects on such skills when taking more courses in “technical/preprofessional areas” (p. 275) compared to students whose parents had college degrees. Consequently, it was hypothesized that first generation students could demonstrate more substantial developmental improvement in their writing abilities in the intervention in this study than other students due to both the subject matter (social sciences) and the increased writing demands of the assignment.

Summary

Sociology is an ill-structured domain in which the types of problems addressed in its practice do not tend to have binary solutions. Since complex problems in ill-structured domains often require the justification of one of many possible solutions, building argumentation skills could improve learners’ performance in those domains.

It is difficult to represent such problems in a manner that allows learners to be exposed to not only the basic facts, but also the myriad perspectives of important stakeholders that must be addressed in order to pose a workable solution. Cognitive flexibility hypertexts, which encourage non-linear navigation that demonstrates the interconnectedness of theories, concepts, and multiple perspectives inherent in ill-structured domains may allow learners to appreciate the complexity of a sociological problem and consider the ramifications of proposed solutions for multiple constituencies.

Male students and first generation students may be less prepared in their first collegiate year for an argumentative essay assignment than their peers. While the current study does not represent a specific attempt to target these populations in the design of the learning intervention, it does identify these demographics as potentially confounding factors in assessing the efficacy of the treatment overall.

Chapter 3: Methodology

Research Questions

The research questions for this study, presented in order of how they are discussed in this chapter, are as follows:

1. Which instructional intervention (summarize, construct, or counter-argue) results in the highest quality of argumentative essays?
 - a. Are there significant differences in performance between genders among the instructional interventions?
 - b. Are there significant differences in performance between first-generation and non-first-generation college students among the instructional interventions?
2. Are there significant differences among the treatment groups in the number of revisions that participants make to their argumentative essays?
 - a. If significant differences exist, what proportion of the overall performance variance is represented by revision attempts vs. the treatment itself?
3. Are there significant differences among the treatment groups in the manner in which they used the case website?

Participants

The data for this study was gathered from several sections of an Introduction to Sociology course at a large, public University in the Midwestern U.S. in fall 2009. Students in all undergraduate majors may choose this course to fill a general requirement in Behavioral and/or Social Sciences. Consequently, the course typically experiences large enrollments across multiple sections during both resident academic semesters. Most students generally take the course during their freshman year, though some take it as late as their senior year.

Students in all three course sections were required to participate in the argumentation assignments for course credit, however, they were given the option of whether to have their data included in the study. Students responded electronically via a form in SAGrader to indicate whether or not they consented to have their data included. A total of 538 students consented to be included across the three course sections.

Of those students in the three sections of the Introduction to Sociology course who consented to have their data included, 506 completed all three argumentation assignments. Of these, 143 were in the section that served as the control group, 219 were in the section that constructed arguments from the cases, and 144 were in the section that generated counterarguments to a provided model. Out of those students who completed all three argumentation assignments, however, only 399 completed demographic data surveys (122, 217, and 60 respectively in the control, construct, and counter-argue groups).

Of those participants who completed demographic data surveys, their ages range from 17-29 years of age and include students from across all undergraduate class years with the majority being freshmen of 18-19 years of age. The mean age of participants was 19.09. The gender distribution of the participants is 58.4% female, 41.4% male and .3% transgender. The majority of participants are white (89.7%) and non-Hispanic (96.2%). 80.7% of the participants are first-generation college students. These distributions are similar across treatment groups with the exception of gender where the control group is evenly split between male and female participants while the construct and counter-argue groups are divided 39 to 61% and 33 to 67% respectively.

Instructional Methods

One topic area covered by all instructors for this course is social stratification and inequality. Typically covered in lecture near the middle of the semester, the topic includes reading, lecture, and other learning activities designed to teach students about concepts related to

social stratification and inequality, such as social class, race, and gender; and the basic theories that inform the ways that sociologists study social stratification and inequality, such as Conflict Theory and Social Interaction Theory.

The majority of instructional design content for the course takes the form of textbook and ancillary reading material, lectures, essay assignments, and closed-book exams composed of multiple choice and/or short answer questions. Students attend lectures twice weekly, with some class periods dedicated to exams, for the duration of the sixteen week semester. The instructional methods employed for this study represent a moderate departure from those used throughout the rest of the semester in the course in terms of the web-based instructional media and the argumentation aspect of the writing assignments.

The argumentation assignments were introduced in a lecture during a regular class meeting of each section upon release of the case 1 assignment. The lecture consisted of a brief overview of how arguments are constructed and used in sociology along with an explanation of basic argument construction as a claim or thesis supported by evidence. In addition, instructors described the types of evidence that students would weigh as they examined the cases: the perspectives of the participants in the cases, the sociological theories that could be used to frame their arguments, and the effects of structural inequality, on which students had recently completed a course unit of study. Students were taught that participant perspectives should be examined for bias and logical fallacy when determining their evidentiary value. In this lecture, instructors also explained that effective arguments anticipate counterarguments and rebut them. With this common background, all three treatment groups were instructed to use the learning environment described below to complete the argumentation assignments.

The primary learning environment for the study is a series of static web pages accompanied by associated argumentation assignments in a system called SAGrader (Ideaworks

Inc., 2008), in which students enter essay assignments and receive feedback. This section will provide a brief overview of both aspects of the learning environment.

Case website. The learning environment is designed around three narrative cases that explore structural inequality. In the first, students are asked to take the role of the landlord and decide between two candidates for an apartment (Center for the Study of Problem Solving, 2009a). In the second, students assume the role of CEO for a company and consider the opinions of a group of vice presidents in the company as well as the relevant candidate submissions to decide which of three candidates is best for an open position (Center for the Study of Problem Solving, 2009b). In the third, students act as the new head of an admissions committee and consider a still broader range of perspectives, such as political, administrative, and academic, in addition to the qualifications of three candidates to determine which should be awarded admission to University (Center for the Study of Problem Solving, 2009c).

For all three cases, a set of file folders is presented as a visual metaphor to navigate the various documents and perspectives associated with the candidates and other parties involved. When working on a particular case, students may navigate back to prior cases but not to future cases. Navigation is non-linear with multiple connecting points to the Sociology-specific information pertaining to the cases in order to serve as a cognitive flexibility hypertext.

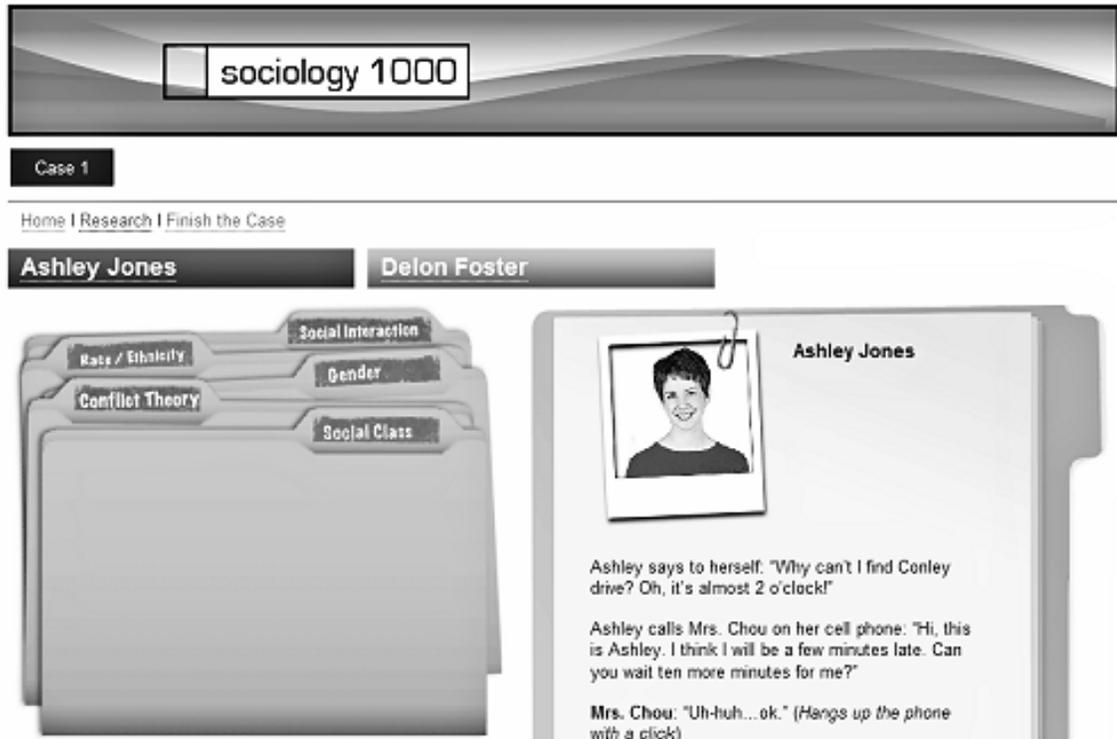


Figure . Screenshot from the case website

In addition, the domain-specific information is presented differently in each case. In the first case, sociological theories and concepts are given in paragraph form in such a way that the text can serve as a fairly direct example of how one could argue for a given position on that case. In the second case, these topics are covered as more general bullet points, though they still address specific people and perspectives in the case. In the third case, the topics are given as a series of questions to guide the students to consider how the theories and concepts apply in the case, but no specific arguments are given that pertain to individuals in the case. The use of questions was suggested by the use of question prompts in prior research to guide problem solving (e.g. Ge & Land, 2003). The variation in presentation of content information from the first case to the third is intended to serve as a form of fading scaffolding to guide students about

how to use sociological concepts and theories in argumentation that fades as they progress through the cases.

SAGrader. SAGrader is designed to provide automated grading and feedback for student essays. Students log into SAGrader to receive instructions about their assignments and enter their responses directly into the web-based forms provided. Upon submission, SAGrader provides an immediate score and specific feedback about items for which the student did not receive credit. In the Introduction to Sociology course, the instructor allows students to resubmit an assignment as often as they wish (up until the deadline) until they are satisfied with their score.

The screenshot displays a web-based assignment interface. At the top, the heading "Short Answer Questions" is visible. Below it, a text box contains the following instructions: "In this exercise you are to learn about Mrs. Chou who is looking for a renter for the upstairs apartment in her house, and two potential renters, Ashley and Debn. You will first need to research the case, looking at information about each renter and their interview. Then you will be asked to examine the conversation between Mrs. Chou and her daughter and identify various argumentation strategies and sociological concepts and perspectives they use as they consider the alternatives." Below this text is a button labeled "Research the Case".

Below the first section, there are two more question prompts, each followed by a large empty text box for the student's response:

- Identify at least three kinds of fallacies in reasoning in Mrs. Chou's or her daughter's comments and cite the comment that illustrates each of those fallacies.**
- Identify the final claim made by Mrs. Chou.**

Figure . Screenshot from the Case 1 assignment in SAGrader

For the argumentation assignments in this study, students received initial instructions in SAGrader and a link to the relevant case website. The assignment questions were also included on the assignment page in SAGrader. The link opened the case website in a new window so that the SAGrader window remained available to them throughout their work on the activity. Thus,

students could navigate through the case web pages in one browser window and switch back to their SAGrader browser window to answer the assignment questions.

Development of the environment and associated treatments. An initial pilot that included only the first two cases was implemented in a single section of the Introduction to Sociology course to test the usability of the learning environment. In these assignments, students answered specific questions about the cases that were designed to guide them toward building an argument. Students were surveyed about their experiences using the case website and the associated SAGrader assignments. Students generally reported spending more time on case 2 than they did in case 1, which was anticipated due to both the increased complexity of the case itself and the more detailed questions in the assignment for case 2. However, the majority of the survey respondents (53.45%) reported that they found both cases “about the same” in terms of level of difficulty. Among those who felt that one case was easier than the other, most unsurprisingly chose case 1 as easier than case 2 (130 respondents vs. 25).

Students were also asked if they used supplemental resources, such as their textbooks, lecture notes, or other Internet resources, to complete the assignments. Nearly half of the respondents reported using additional resources for both cases, while slightly less than one third of them indicated that they used no additional resources. Of those who reported using supplemental resources for one case but not the other, 21 of the respondents said they used them for case 1 but not case 2 and 61 said they used them for case 2 but not case 1. It should be noted that students were fully permitted and expected to use their textbooks and other course materials as needed to complete the assignments.

Most respondents (about 62%) felt that the argumentation assignments had helped them improve their knowledge of social stratification at least somewhat. A slightly smaller number (about 57%) believe that the assignments helped them improve their understanding of sociological theories. However, about 54% reported that they did not perceive improvement in

their argumentation skills as a result of completing these assignments. Since this pilot implementation did not include pre-tests on any of these areas of concern, we will not be able to determine the accuracy of these student perceptions. However, full analysis of performance data from the assignments would indicate what percentage of the students enrolled in the course successfully achieved the argumentation and content goals of the assignments.

Students were also asked whether they found the experimental assignments more or less interesting than the other assignments in the course. About 26% found the argumentation assignments more interesting while 47% preferred the non-experimental assignments. Approximately 27% had no preference for the experimental vs. the non-experimental assignments.

Students were also given the opportunity to respond to two open-ended comment questions about improving the case website and improving these assignments. Their responses demonstrate that many students were uncomfortable with the way that the information is presented in the case website. Some wanted all of the information on one page so that they wouldn't have to hunt for it (one respondent likened the experience to a "scavenger hunt"). Several wanted pointers to lecture notes or specific pages in the textbook on each HTML page on the case website. At least one respondent complained that the case website contained "too much information ... in relation to what was actually needed to complete the assignment."

Such student experiences are of concern if they impede the acquisition of structured knowledge that Spiro et al. argue is typical of an introductory course (1992). However, it is also important to note that these assignments were specifically designed to expose students to the ill-structured nature of the domain of sociology, particularly in its application. While an argument can be made that such exposure may be better reserved for more advanced students in the discipline, the instructor of the course believes that the cases are sufficiently limited to a degree that is well-suited to what the students are expected to understand from their study of related

course materials and completion of several more traditional assignments prior to the argumentation exercises. Consequently, a conscious decision was made to keep the non-linear structure of the case website for the study.

SAGrader uses a custom-built semantic network to assess student responses for assignments. The instructor guided the development of the code for the first two argumentation assignments for the initial pilot based upon existing, non-argumentative essay assignments previously used for the section of the course that focused on social stratification and structural inequality. Prior to deployment of the third assignment, in which students would be responsible for composing a full argumentative essay based on the case information, the semantic network was updated to include assessment of the quality of argumentation based the rubric (see Appendix A) used in Jonassen and Cho (2011).

A second pilot was completed with all three cases in order to verify the appropriateness of the SAGrader scoring. Students in every implementation of an SAGrader assignment are permitted to submit challenges to the scoring that are reviewed by the instructor. In addition to making any grade adjustments that the instructor deemed appropriate for affected students, the challenges in this pilot were used to refine the semantic network used in SAGrader to evaluate and score the submitted essays for each associated assignment prior to this study.

Treatments. In studying argumentation, researchers have found that students have a variety of difficulties learning to develop effective arguments as discussed in Chapter 2. Perhaps chief among them when it comes to ill-structured problems is the inability to address multiple points of view in an effective way. For example, in the series of studies by Perkins and his colleagues in the 1980s (Perkins, Farady, & Bushey, 1991), students at all levels of education as well as adults long out of school tended not to include counterarguments in their consideration of social dilemmas such as funding for schools and the nuclear arms race unless explicitly instructed to do so. While those at higher levels of education included a higher number of counterarguments

than those at lower levels of education, they still tended not to consider all possible sides of an argument.

Nussbaum and Kardash (2005) considered this when writing instructions designed to elicit more counterarguments in argumentative essays among college students. They found that the students who were given specific instructions to “discuss two or three reasons why others might disagree with you, and why those reasons are wrong” (p159), representing counterarguments and rebuttals respectively, not only generated more counterarguments, rebuttals, and support for those rebuttals than did their peers who were not so instructed, they also performed significantly better on the holistic quality scores for their essays.

Jonassen and Cho (2011) hypothesized that students might be more inclined to generate counterarguments to a provided solution than to their own solution. They compared one group of students who were asked to generate counterarguments to a provided solution (other-argument), which included a high quality of support but no counterclaims, to another group who were given instructions similar to the Nussbaum and Kardash (2005) study (self-argument). As Jonassen and Cho (2011) predicted, the students in their other-argument group generated more counterarguments than their counterparts in the self-argument group. The counter-argue group also outperformed the self-argue group in the overall quality of their essays.

Nussbaum and Schraw (2007) asserted that an effective argument requires more than merely enumerating possible counterarguments; rather, it requires one to integrate into one’s argument the possible support for those counterarguments and effective rebuttals to that support. In their experiment, they provided some students with guidance about what good arguments contained through the use of graphic organizers in the form of worksheets that laid out the structure of a good argument. Some of the students who received the graphic organizers and some of the students who did not also received instruction about how to create an effective argument. The participants in this graphic organizer condition outperformed other participants in generating

rebuttals, and outperformed the control group in generating counterclaims, but did not perform a group who received no graphic organizer but did receive instruction in how to create an argument in number of counterclaims. In their analysis of integration, the instruction had a main effect but the graphic organizer did not. Both graphic organizers and instruction had positive effects on holistic essay scores.

For this study, the treatments were similar to those in the Jonassen and Cho (2011) study with the addition of a control treatment designed in similar fashion to earlier studies in the research group (Jonassen, Cho, Kwon, Henry, & Easter, 2009; Jonassen, Cho, Kwon, Henry, Shen, et al., 2009) in order to see if their results would extend into a different discipline with a more diverse student population. Thus, three different versions of the assignments for the two teaching cases were designed to reflect the planned treatment groups. Based upon methods in the research group's earlier studies (Jonassen, Cho, Kwon, Henry, & Easter, 2009; Jonassen, Cho, Kwon, Henry, Shen, et al., 2009), the assignments encouraged students to either summarize salient information from the two teaching cases (to serve as a control group); and upon methods based upon used in Jonassen and Cho (2011), to construct their own arguments for their selected candidate in each case (construct group), or to argue against a provided model argument (counter-argue group). In this study, based upon Nussbaum and Schraw's (2007) idea that modeling a good argument would improve student argumentation performance, Jonassen and Cho's (2011) model argument was adapted to include consideration of counterargument and rebuttal. All three treatment groups received the same version of the case website and the same assignment for the third case to serve as the post-test for the study. The assignment instructions for each of the three treatments for all three cases are provided in Appendix B.

Summarize treatment. The intention of the summarize treatment was to guide students to identify the types of evidence they would need to construct an argument without instructing them to apply that evidence to an argument of their own making. Thus, in the first case, the

students in this treatment group were instructed to provide at least two supporting and two opposing warrants for each renter. They were also asked to classify the warrants as “evidence, opinion, or something else.”⁵

In the second case, the students in the summarize treatment group were again asked to identify supporting and opposing warrants for each candidate. This time, for the opposing warrants, they were instructed to list one based on opinion and one based on evidence. They were also asked to identify components of argumentation, such as warrants and rebuttals, specifically from the transcript provided to them of senior management meeting to discuss the job candidates in the case (<http://csps.missouri.edu/soc1000/FA09D/case2-meeting.html>).

Construct treatment. In contrast, the assignment instructions for students in the construct treatment were designed to lead them through the process of building their own arguments. The first thing they were asked to do in each of the two teaching cases was to choose which candidate they would support. In the first case, then, they were asked to provide supporting warrants for their candidate (like the summarize treatment group, the construct treatment students were asked to classify the warrants using the same wording). They were also asked to provide warrants that challenged the other candidate and classify them. They were also asked to identify a rebuttal statement supporting their candidate made by one or the other of the characters in the case and identify the opposing argument that the statement rebutted.

Likewise in the second case, while the assignment instructions were similar to those of the summarize treatment group, they were written from the perspective of selecting and supporting a particular candidate. They selected warrants “supporting [their] choice,”⁶ identified opinion-based warrants opposing their candidate, and selected evidence based warrants opposing each of the other candidates they did not select.

⁵ Quoted from the text of the assignment question (See Appendix A for details).

⁶ Ibid.

Counter-argue treatment. In the counter-argue treatment, students were shown an example essay designed to model the level of argumentation quality that would be expected of students in the transfer case. The instructions were designed to guide the students to argue against the position taken in the example essay. In the first case, this predetermined the students' candidate choice since the case included only two. The essay supported the candidate named Ashley. Counter-argue treatment participants were instructed to provide warrants that "support your choice to rent to Delon" and "against renting to Ashley."⁷

In the second case, their instructions were quite similar to those of the construct treatment because students were able to choose between two remaining candidates other than the one supported in the example essay. However, they were reminded in the general instructions that their "task is to make a counterargument, arguing in favor of either of the other two candidates and against hiring John Meyer."⁸

Data Collection

All instruments were deployed in the SAGrader course website for each treatment section of the course. Each treatment section had its own instance in SAGrader so that data was stored separately for each group. Data was stored on the SAGrader server and exported in Excel spreadsheet format with an internal participant ID and no personally identifiable information. Instructors of three sections of the Introduction to Sociology course agreed to participate in the study.

Website usage statistics were gathered anonymously using Google Analytics. Treatment groups received identical sets of pages in separate server directories. A tracking code was added to each page on the site that accumulated data into a single account for all three treatment groups.

⁷ Ibid.

⁸ Ibid.

The directory was used to distinguish one group's activities from another in the aggregated data set.

Design

A quasi-experimental, non-equivalent groups post-test design was used when applying the treatments. Three sections of the course formed the three treatment groups. The three treatment methods (control, construct, and counter-argue) were randomly assigned, one to each of the participating sections.

Methods

The following section describes the methods used to answer the research questions.

RQ1: Learning outcomes.

1. *Which instructional intervention (summarize, construct, or counter-argue) results in the highest quality of argumentative essays?*
 - a. *Are there significant differences in performance between genders among the instructional interventions?*
 - b. *Are there significant differences in performance between first-generation and non-first-generation college students among the instructional interventions?*

Two stages of analysis were performed to address the question of whether and how learning outcomes might have been affected by the treatments applied in this study. First, overall scores were examined by treatment group, gender, and first generation student status. Then performance was examined at a granular level by examining which components students achieved, partially achieved, or missed by treatment group, gender, and first generation student status.

Overall scores. Data was gathered, as described previously, on participant gender and status as a first-generation college student. Case 3 was considered to be the post-test as the same

instructions and requirements were provided to all three treatment groups. Student participants were instructed that their first submission would account for 25% of their final grade to encourage them to make a serious attempt at a first draft rather than entering minimal data and merely allowing the SAGrader feedback to guide their work. Using this first draft score, then, would reflect performance without any potential confounding effect that might result from the influence of the SAGrader feedback. In addition, since students could revise as often as they wished, it was assumed that final scores would tend to show less variance and be clustered more closely toward the perfect score mark (i.e. students would continue to revise until they achieved a high enough score to get the grade they desired).

The data was screened for assumptions of normality and homogeneity of variance. In screening for normality, first revision scores were shown to have a slight negative skew against the hard limit of a perfect score. However, since sample sizes larger than 30 were used, ANOVA results can be considered to be robust. Levene's test for homogeneity of variance was nonsignificant indicating that the assumption was met. Since only one participant identified as transgender, the case was omitted as an outlier for gender because results would have been constant and thus not meaningful. No outliers were identified using Mahalanobis distance, so no other cases were discarded.

Factorial ANOVA was performed using the first revision score for case 3 as the dependent variable with group, gender, and first generation student status as the independent variables. Hochberg's GT2 was selected as a post-hoc test of significance because the sample sizes are different among the groups, genders, and first generation statuses.

Components. Component data was gathered in ordinal form. Students received a 1 for each component item that they successfully achieved, 0.5 for each component that they partially achieved, and 0 for each component that they missed. With ordinal data, the distributions are considered to be nonparametric and thus cannot be assessed using ANOVA or MANOVA.

Consequently, the Kruskal-Wallis H Test was run to assess differences among the three treatment groups while Mann-Whitney U Test was run to determine differences between genders and between first generation and non-first generation students. While Creswell (2008) suggests that the use of these tests requires a continuous dependent variable, Cohen, Manion, and Morrison⁹ indicate that it is also appropriate when using an ordinal dependent variable (2007).

Since these tests do not allow the use of multiple independent variables, separate Kruskal-Wallis H Tests by treatment group were performed on the following subsamples of participants by group:

1. All male participants
2. All female participants
3. All first generation participants
4. All non-first generation participants

There are no data screening procedures for the non-parametric tests. However, the single transgender participant was omitted as an outlier for the same reasons stated earlier. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons.

RQ2: Revisions.

2. *Are there significant differences among the treatment groups in the number of revisions that participants make to their argumentative essays?*
 - a. *If significant differences exist, what proportion of the overall performance variance is represented by revision attempts vs. the treatment itself?*

⁹ A sampling of additional sources that support this choice include (Leech, Barrett, Morgan, Clay, & Quick, 2005), (Lund & Lund, 2012), (Guerra-López, 2008), and (Chan & Walmsley, 1997).

To verify that the number of revisions had an overall influence on the final scores in the study, a simple linear regression was performed using the Case 3 final score as the dependent variable and the number of revisions as the independent variable. Then, in order to evaluate whether the treatment had an effect on number of revisions and its relationship to overall performance, a multivariate analysis of variance (MANOVA) was performed with treatment group number as the independent variable using the number of revisions and the SAGrader final score as dependent variables. The data set was first screened for missing data and multivariate outliers. No cases were missing data. Outliers were screened using the criterion of Mahalanobis distance with three variables (group, Case 3 final score, and number of revisions) for which $X^2(3) = 16.266$ at $\alpha = .001$. Using this criterion, three cases were identified as multivariate outliers and discarded from the analysis.

Box's Test of the Equality of Covariance Matrices was significant at $p < .001$ with the full data set. With unequal numbers of participants in the treatment groups, this result suggested that MANOVA results from this data set would be unreliable. Consequently, because Box's Test is only relevant when there are unequal samples, the data set was further screened by selecting 60 cases at random from each of the larger groups (summarize and construct) to match the size of the smaller counterargue group.

Since the planned MANOVA contained only one independent variable (group number), no screening test was performed for multicollinearity. In screening for multivariate normality, scores were shown to have a slight negative skew against the hard limit of a perfect score. However, since sample sizes larger than 30 were used, MANOVA and ANOVA tests of significance can be considered to be robust.

The data was also screened for linearity, and homoscedasticity with no unusual results. The Games-Howell procedure, which affords additional accuracy in the event that population variances are not equivalent given the significant result on Box's Test, was selected as a post-hoc

test to verify the findings. Significant results were also verified using a post-hoc, one-way ANOVA.

RQ3: Website usage.

3. *Are there significant differences among the treatment groups in the manner in which they used the case website?*

As previously noted, data associated with website usage was gathered anonymously using Google Analytics. Since individual sessions were not associated with individual participants, activity on the site cannot be correlated with performance data. However, because participants in the different treatment groups used identical sets of web pages in different directories, the usage data can be used to compare aggregate differences among the groups.

Two types of comparisons were employed on the website usage data from the third case in order to examine differences in web usage among the treatment groups. The first comparison examined whether there were statistically significant differences in page views, average time per page, average page depth per visit, and average visit duration among the groups. The second comparison used quantitative page view data to visualize the most frequently traveled paths that different groups used when navigating the site.

Statistical analysis of usage statistics. Since Google Analytics data are anonymized and aggregated, it was not possible to analyze the data with the participants, or even merely their sessions, as the subjects. However, if we consider that the pages themselves are the same across the treatment groups, it is possible to consider those pages effectively as one group of subjects to which all three treatments were applied (as contrasted with the performance data where three groups of subjects received three different treatments). The dependent variables used in this analysis, then, represented how the pages performed under the different treatment methods, thus providing an indirect mechanism for assessing how the student participants interacted with the site at an aggregate level by treatment group.

Page views represented how many times each page was accessed during the study period. In order to compare this across treatment groups containing unequal numbers of participants, however, the number of participants in the treatment groups had to be accounted for in some way. Further, since the website usage data was gathered anonymously, it necessarily included participants who were omitted from the previous analyses, either because they had not consented to have their performance data included or because they were outliers for the purposes of a particular method. To address these issues, the figure used to represent the number of participants in each treatment group was calculated by counting the total number of participants who had submitted the third case assignment. Then, rather than using the value for page views for each page directly, a ratio value calculated from the number of page views divided by the number of participants in the treatment group was substituted as a dependent variable.

Other Google Analytics values were able to be used directly as dependent variables because they represented averages of the aggregated data. Average time per page, average pages per visit, and average visit duration for a treatment group could be compared directly with those averages from other treatment groups because they were not skewed by the number of participants in each group.

While the original intent was to perform a MANOVA on this data using treatment group as an independent variable with page view ratio, average time per page, average pages per visit, and average visit duration as dependent variables, tests of normality showed the dependent variables to be significantly skewed to the left. MANOVA can be considered to be robust to minor violations of the assumption of normality with samples of at greater than 30 subjects, but the violation was judged to be too severe. Consequently, a Kruskal-Wallis H Test for non-parametric distributions was performed. No data screening procedures or post-hoc tests are available for the Kruskal-Wallis test.

Analysis of page navigation. While the statistical analysis of website usage data can give some clues about how much a treatment group might have used the pages in Case 3, it was of limited efficacy in describing whether students navigated those pages multiple times from multiple points in the website as was previously described as the intention of design in a cognitive flexibility hypertext (Spiro et al., 1992). The Google Analytics data does collect data on path navigation, however, which was used to examine the most frequently used paths through a website.

Like the usage data, the path data available from Google Analytics was aggregated. As a result, it was not possible to represent each unique session's full path in an analysis. The aggregated data did allow, however, output of the frequency of every pairing of movements from page to page. In other words, if a participant was on one page and clicked to another, a source and target pairing was recorded. These pairings allowed an aggregated view of the traffic on the site.

The pairings for Case 3 were be partitioned by group to facilitate comparison of differences in navigation patterns among treatment groups. A one-way ANOVA was performed with the unique pairings as rows, the frequencies of those pairings as the dependent variable, and the treatment group as the independent variable. Hochberg's GT2 was selected as a post-hoc test of significance because the sample sizes are different among the groups.

In addition, the Case 3 pairings and their frequencies were used to generate a visualization of each group's navigational use of the various pages provided for the case. Since the website was designed to be non-linear to encourage students to visit a variety of relevant perspectives, concepts, and theories from nearly every other page in the site, each page served as a potential connection to some number of other pages.

According to Newman (2010), the World Wide Web is a network, and the individual sites within it are also networks. Networks are comprised of points, usually called nodes or vertices, that are connected via lines called edges. Websites are directed networks because a particular link

can only take one specific edge to another specific node. Such networks are generally analyzed for centrality based on the importance of the vertices or edges that comprise the network.

Newman's ideas were applied in this study by designating each page as a node in the Case 3 website network. The pairings from the Google Analytics data were used to determine the links, and the frequency with which each pairing was used. The pairings were considered directional so that an edge from the research page for Case 3 to the page for Candidate 1 was distinct from an edge from the page for Candidate 3 to the research page.

To construct the diagram of the Case 3 network for each treatment group, the aggregated data set was sorted by group and divided into separate files, one for each group. Self-referential pairings (i.e. those in which the source and target pages in the pairing were equal) were discarded, as were any data rows that did not contain both a source and a target¹⁰.

These pairings along with the frequency at which each was used by the group was imported into Gephi (Bastian, Heymann, & Jacomy, 2009). The first visualization for each group was created using an unfiltered force-based layout (ARF) with nodes colored by page grouping (all pages for a particular candidate grouped together, main case pages not associated with a candidate grouped together, and pages from teaching cases grouped together) and edges colored black. Metrics were produced using Gephi for degree centrality and eigenvector centrality. The degree centrality metrics calculated average degree and average weighted degree for nodes in the network. Eigenvector centrality also takes into account the relative importance of the other nodes to which a node is connected in order to determine how important a node is (Newman, 2010,

¹⁰ Google Analytics tracks "entrances" and "exits." Since the tracking cookie only takes effect when the browser reaches one of the pages on the site containing the tracking code, it records the first instance of a browser setting the cookie during a session as an "entrance." "Exits" occur when a user departs the site for a page that does not contain the tracking code, such as a page on another site, when the browser has been closed, or when the user has not navigated to another page within the site for 30 minutes.

Chapter 7, Section 7.2, "Eigenvector Centrality"). One-way ANOVAs were performed using SPSS to identify any differences in degree centrality or eigenvector centrality among the groups.

A second visualization was created for each treatment group using a force-based circular layout with an edge-weight filter set between 1.8 and 2.8 (the exact filter setting varied based on the available range of weights). The other settings remained the same as in the ARF visualization.

A third visualization was created for each treatment group using a force-based dual circle layout that forced the main case pages not associated with a particular candidate outside the primary circle. The edge-weight filter settings from the circular visualization were again used to keep the edge lines reasonably distinct within the circle. While the nodes were still organized in the inner circle by their page grouping, they were colored using node degree, defined as the number of edges connected to the node (Newman, 2010, Chapter 1, "Properties of Networks," para. 3).

Limitations of the study

While the research design attempted to reduce the likelihood that statistically differences inherent in the three treatment groups would bias the resulting estimation of treatment effect in the study, selection bias may still have occurred because participants were not randomly assigned to treatment groups. Since different instructors taught the three different sections used to form the three treatment groups, the relative efficacy of the various treatments may be confounded by teaching effects. This specific question is planned to be addressed in future research in which instructors will switch which treatment group they teach from one semester to the next so that the results can be analyzed among both treatment groups and instructors to determine how much of the variance in performance can be attributed to the instructor vs. the treatment.

The Kruskal-Wallis H Test, which is used in this study to assess both granular performance data and website usage data, is widely used and considered valid for non-parametric distributions. However, it should be noted that the tests were performed using SPSS, which

employs asymptotic approximation when determining whether the results are significant rather than exact tests. This is typical of available statistical analysis software and the few exact probability tables that exist to double check the significance findings are limited to samples smaller than those in this study (Meyer & Seaman, in press).

Due to the way that the Google Analytics data are aggregated, it cannot be directly tied to the actions of individual participants. Consequently, detailed statistical analysis about the ways in which website behaviors of participants might affect their performance on the argumentation assignment was not feasible in this study. The results of the analysis of website usage data should be viewed as an exploration of how the case website was used in this study and should not be considered to be generalizable to cognitive flexibility hypertexts in general.

Chapter 4: Results

As noted earlier, the purpose of this study was to examine which of three types of practice activities (summarization of the evidence available for an argument, generation of original components of argument, or generation of components to counter a model argument) result in the highest quality of original argument. The study also sought to validate Jonassen and Cho's (2011) finding that counterargument generation resulted in higher quality arguments and extend it to another field of study.

The instructional methods and data collection strategies used in this study also afforded the opportunity to address questions pertaining to treatment differences in student activities that might impact their argumentation performance. Specifically, the use of SAGrader, which allowed and encouraged student revision to occur within the assignment submission system, afforded the opportunity to explore whether different treatment groups revised more than others. Further, the gathering of website usage statistics allowed some ability to note differences in cognitive flexibility hypertext usage among the treatment groups.

This chapter details the results of the methods employed in the study as described in the preceding chapter. It is organized according to the order of the research questions:

1. Which instructional intervention (summarize, construct, or counter-argue) results in the highest quality of argumentative essays?
 - a. Are there significant differences in performance between genders among the instructional interventions?
 - b. Are there significant differences in performance between first-generation and non-first-generation college students among the instructional interventions?
2. Are there significant differences among the treatment groups in the number of revisions that participants make to their argumentative essays?

- a. If significant differences exist, what proportion of the overall performance variance is represented by revision attempts vs. the treatment itself?
3. Are there significant differences among the treatment groups in the manner in which they used the case website?

RQ1: Learning Outcomes

As described in Chapter 3, two different methods were necessary in order to analyze the two different types of learning outcomes data. This section discusses the results of the analysis of the overall scores followed by that of the component scores.

Overall scores. As described in the previous chapter, a decision was made to use the first revision score from Case 3 rather than the final score for analysis of overall scores. It was hypothesized that final scores would tend to show less variance due to the influence of SAGrader feedback and the ability to revise as often as the student wished in order to achieve his or her desired grade on the assignment. Before proceeding with the analysis, this hypothesis was checked by using the graphing function of SPSS to view the relationship between the first revision score and the final score for all three treatment groups. Figure is a histogram of the first revision scores. With the exception of an outlier at the very low end of the scale, the distribution of the first revision scores is approaching normality.

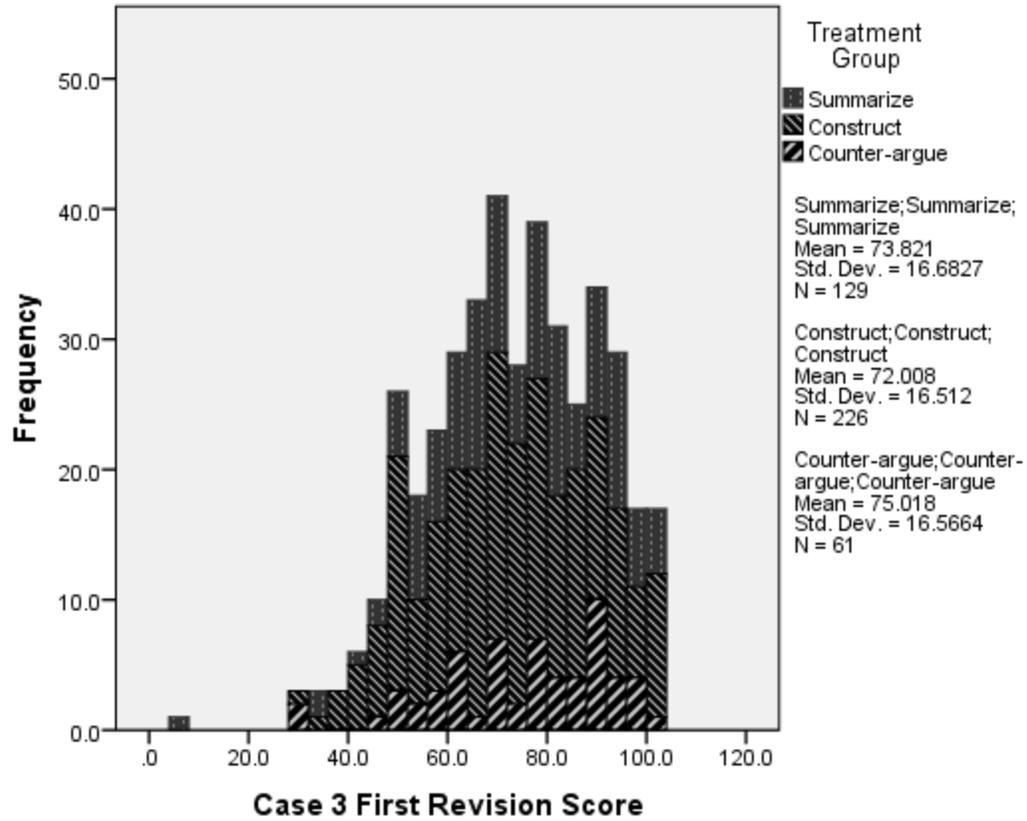


Figure . Distribution of the first revision scores for the Case 3 assignment for all treatment groups.

In contrast, the histogram for the distribution of the final scores shows an obvious negative skew as shown in Figure :

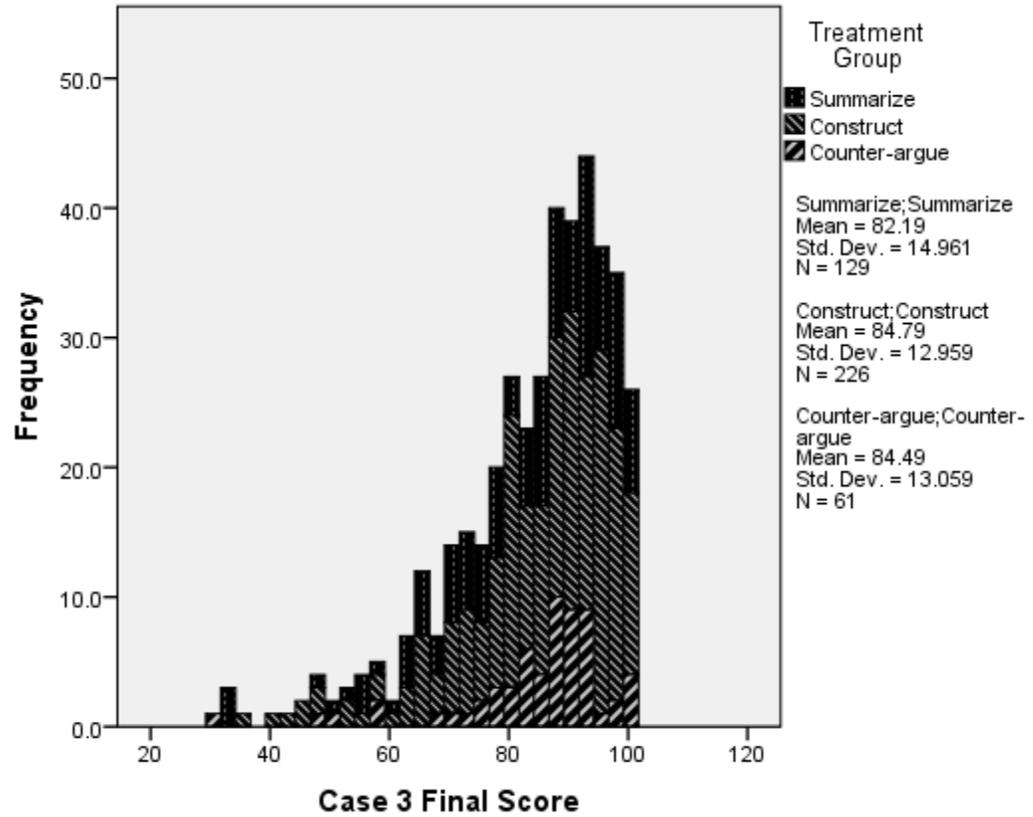


Figure . Distribution of the final submission scores for the Case 3 assignment for all treatment groups.

Figure 5 is a scatterplot illustrating the tendency for the score progression from first revision to final score to converge toward the perfect score mark on the final submissions.

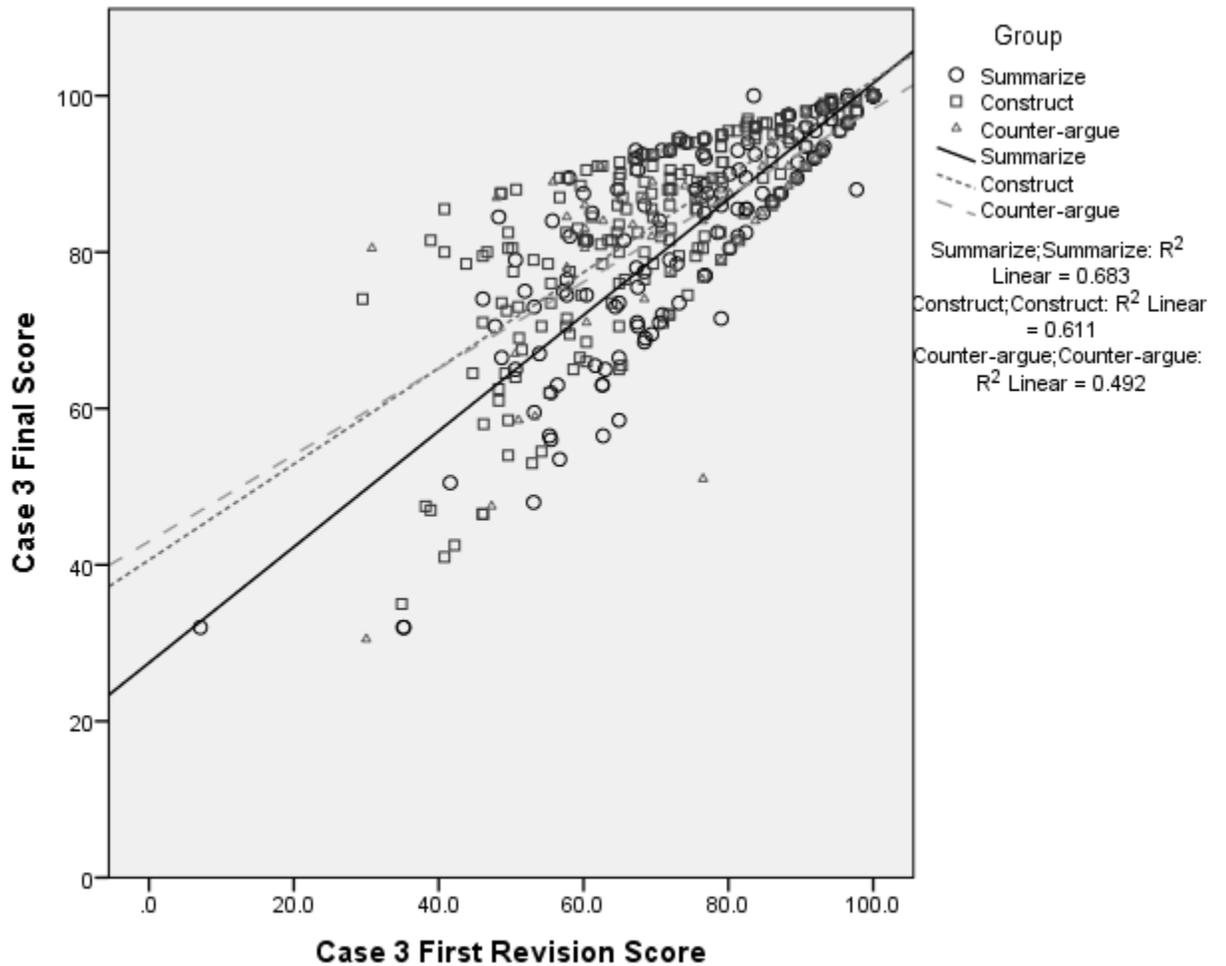


Figure . Simple plot of first revision vs. final scores for all three treatment groups.

Descriptive statistics confirmed the hypothesis. The final scores showed a narrower range than the first revision scores (70 vs. 92.9) as well as less variance (186.128 vs. 274.683), more significant negative skew (-1.358 vs. -.376), and more pronounced kurtosis (2.020 vs. -.249).

A factorial analysis of variance (ANOVA) was performed with the first revision score for case 3 as the dependent variable, and group, gender, and first generation student status as the independent variables.

SPSS ANOVA was used for the analysis. The frequencies report showed no missing values in the data. No univariate outliers were detected, however one participant who identified as transgender was omitted as an outlier. The Shapiro-Wilk test for normality showed significance

less than .05 in all treatment groups for first revision scores, though slight skewness was reported in the descriptives and the slight positive skew was noted in the histograms. Since ANOVA is considered robust to violations of the assumption of normality when sample sizes larger than 30 are used and Q-Q plots showed acceptable normality for all treatment groups, no transformations of the data were performed for normality.

The Shapiro-Wilk test also showed significance less than .05 in both genders, and in both first generation and non-first generation statuses. For these independent variables, the descriptives showed no significant skewness or kurtosis. This was confirmed with histograms and Q-Q plots.

Levene's test for homogeneity of variance was nonsignificant for treatment group and for gender, indicating that the assumption was met for these groupings. The test was significant, however, for first generation vs. non-first generation student status in an untransformed state. Using a square transform of the first revision scores gave a non-significant result on Levene's test for all independent variable groupings. Consequently the scores actually used as the dependent variable for the ANOVA were the actual first revision scores squared.

The results of the ANOVA showed that first generation vs. non-first generation status had a significant main effect on the case 3 first revision score: $F(1)=8.997$, $p < .01$, but the effect was very weak, accounting for only about 2% of the variance in scores with a partial $\eta^2 = .022$. Likewise, there was a weak significant effect in the interaction between group and gender on the case 3 revision score: $F(2)=3.627$, $p < .05$ partial $\eta^2 = .018$.

Table . Between subjects effects on first draft scores for Case 3

Source	df	F	Sig.	Partial Eta Squared	Observed Power ^a
Corrected Model	11	1.949	.032	.051	.895
Intercept	1	717.283	.000	.640	1.000
Group	2	.887	.413	.004	.203
Gender	1	2.368	.125	.006	.336
FirstGen	1	8.997	.003**	.022	.849
Group* Gender	2	3.627	.027*	.018	.669
Group * FirstGen	2	.212	.809	.001	.083
Gender * FirstGen	1	2.930	.088	.007	.401
Group* Gender * FirstGen	2	2.602	.075	.013	.517

a. Computed using alpha = .05

** p < .01; *p < .05

A post-hoc Bonferoni correction procedure was used to obtain pairwise comparisons for the effects of treatment group on the squared first draft scores for case 3. There were no significant differences among the groups found.

Components. In chapter 3, the component data was described as ordinal. Students received a 1 for each component item that they successfully achieved, 0.5 for each component that they partially achieved, and 0 for each component that they missed. As noted, the non-parametric Kruskal-Wallis H Test was used to assess treatment group effects because there were three treatment groups, while the Mann-Whitney U Test was run to determine differences between genders and between first generation and non-first generation students because each of

these variables had only two conditions. Appendix C provides a list and description of all 71 components measured in this analysis.

Table lists only those components with significant results in one or more of the tests. The resulting significance statistic is provided in the column for the test under which it was obtained. In the Overall Measures columns, the Group statistics resulted from the Kruskal-Wallis H Test while the Gender and (First Generation) Status were the result of the Mann-Whitney U Tests. The Gender and Status columns calculated distribution of the component achievement scores among all participants without regard to treatment group, while the Group column shows significant achievement distribution differences among treatment groups without regard to gender or first generation student status.

The remaining columns in Table are the result of Kruskal-Wallis H tests run on sub-samples of the overall dataset. Thus, in the columns for Males and Females, the statistic represents a significant difference in distribution of achievement of specific components for males or females among the treatment groups, but not in comparison to the opposite gender. The First Generation Student Status columns show significant differences in component achievement distribution for first and non-first generation students.

Of 71 components of performance for the Case 3 essay assignment, 41 showed at least one significant distributional difference. Of those, only 13 showed significant differences in the distributions of component achievement at the overall treatment group level. 24 components showed statistically significant differences in achievement distribution between the genders without regard to group. Female participants had the most distributional differences among the different treatment groups at 21. Treatment group resulted in only significant achievement distribution differences among first generation students, while non-first generation students showed significant differences in achievement distribution on 12 of the component items.

Table . Significant component items

Item	By Group Submeasures						
	Overall Measures			Gender		By First Generation Student Status	
	Group	Gender	Status	Males	Females	First Generation Students	Non-First Generation Students
Committee doesn't have authority	0.000	0.001		0.000	0.000	0.001	0.000
False dichotomy	0.000	0.013		0.000	0.013	0.000	0.000
Biases related to method	0.000			0.000	0.000	0.000	0.000
Decision Avoidance	0.000		0.017	0.001		0.027	0.006
Network communication structure	0.000			0.000	0.006		0.000
Hasty Generalization	0.000			0.003	0.001		0.000
Opposing white male candidate	0.000			0.000	0.000		0.000
Support white male candidate	0.000			0.024	0.004		0.001
Must make decision today	0.030		0.048		0.029		
Typology	0.031				0.009		0.017
Learning disabilities example		0.024	0.019	0.042			
Judges evidence based on method	0.010						0.030
Avoids relativism		0.006			0.006		
Raises valued consequences issues		0.001			0.008		
Open to new evidence		0.008			0.011		
Identifies relevant evidence		0.001			0.014		
Anticipates other perspectives		0.015			0.015		
Justifies/defends own position		0.007			0.025		
Explicates logical implications		0.002			0.026		
Acknowledges other perspectives		0.015			0.026		
Correlation is not causation		0.005			0.026		
Examines assumptions		0.007			0.028		
Considers multiple causality		0.009	0.044				
Acknowledges own position		0.003			0.046		
Conclusions qualified by evidence	0.038						
Peter Principle	0.043						
Clearly states their position			0.027				
An applicant							0.004
Promoted and competent				0.045			
Feature of comcon					0.039		
Two applicants							0.004
Two statements			0.011				
Goal displacement			0.023				
Identifies implicit issues		0.003					
Interpret		0.004					
Avoids basic dualism		0.005					
Avoids logical fallacies		0.008					
Raises ethical issues		0.010					
Considers human subjects issues		0.014					
Identifies sub-issues		0.021					
Seeks evidence		0.023					
Feature of star		0.028					

A significant distributional difference in attainment scores, whether as a result of Kruskal-Wallis or Mann-Whitney, does not indicate which group, gender, or first generation student status achieved more items. It simply reflects that the distributions are not proportional at different levels of the independent variable. Figure shows the total proportion of achievement by group for each of the component items that had significant distribution differences among the groups.

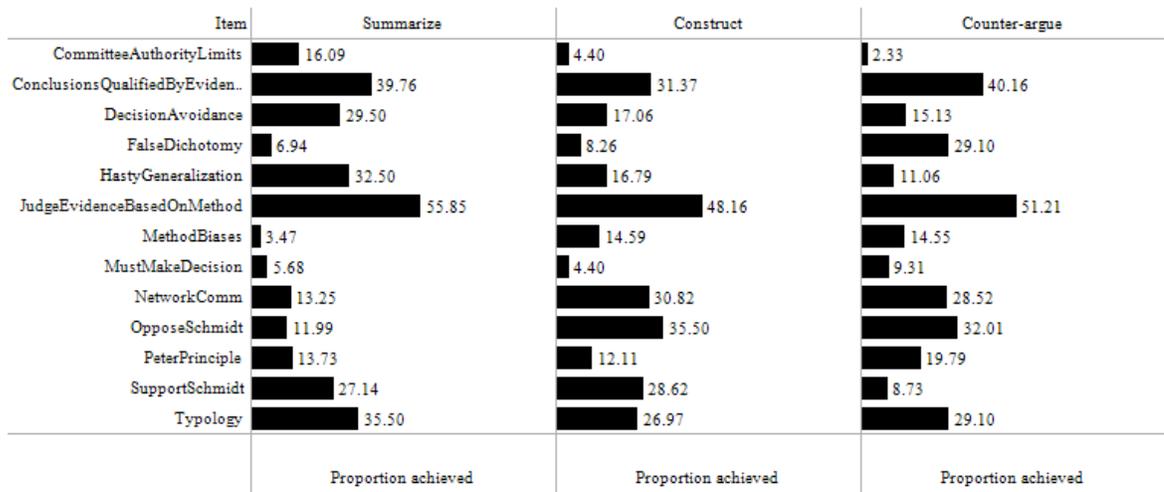


Figure : Proportion of significant items achieved by group.

The proportion in each cell in the figure represents the percentage of full achievement attained by the group. They were derived by dividing the sum of all of the achievement ordinals for the participants in a group by the number of participants in that group. Since the ordinal for full achievement of a component item was 1, the n for the group is equivalent to what the proportion would be if all of the participants obtained full credit.

RQ2: Revisions

A one-way multivariate analysis of variance (MANOVA) was performed on two dependent variables: the final score students achieved on the third case assignment (C3F) and the number of revisions students performed in the third case (C3R). The independent variable was the treatment group to which the students were assigned (group). Prior to running the MANOVA, a

simple linear regression was performed in SPSS to note the amount of variance in C3F that could be predicted by C3R. The adjusted R^2 was .020 with the corresponding ANOVA for the model showing significance of $F(1,413) = 8.520, p = .004$.

SPSS MANOVA was used for the analysis. The frequencies report showed that there were no missing values in the data. No univariate outliers were detected. Multivariate outliers were detected using the criterion of Mahalanobis distance with three variables (group, C3F, and C3R) for which $X^2(3) = 16.266$ at $\alpha = .001$. Using this criterion, three cases were identified as multivariate outliers and discarded from the analysis.

Box's Test of the Equality of Covariance Matrices was significant at $p < .001$ with the full data set. With unequal numbers of participants in the treatment groups, this result suggested that MANOVA results from this data set would be unreliable. Consequently, because Box's Test is only relevant when there are unequal samples, the data set was further screened by selecting 60 cases at random from each of the larger groups (summarize and construct) to match the size of the smaller counter-argue group.

Since the planned MANOVA contained only one independent variable (group number), no screening test was performed for multicollinearity. Using the Shapiro-Wilk test for normality, both dependent variables showed significance less than .05 in all treatment groups. C3F did show a slight negative skew against the hard limit of a perfect score in the histograms, while C3R showed a slight positive skew in the histograms. However, since sample sizes larger than 30 were used, MANOVA and ANOVA tests of significance can be considered to be robust. Further, Q-Q plots showed acceptable normality for all treatment groups.

Based on Wilks' lambda, the results showed that the combined DVs were significantly affected by treatment: $F(4,352) = 4.056, p = .007$. However, treatment accounted for only 4% of the total variability in the combined DVs with partial $\eta^2 = .044$. Univariate ANOVA shows that the effect of treatment on the the number of revisions is significant: $F(2) = 7.344, p = .001$, but

not on the final score on the assignment: $F(2) = 1.033, p = .358$. The Univariate ANOVA also shows that treatment is only minimally associated with the variability in the number of revisions despite the significant result, accounting for only about 8% of the variability in C3R (partial $\eta^2 = .077$).

Table . Tests of between-subjects effects by treatment group on score and number of revisions

Source	Dependent		F	Sig.	Partial Eta	Observed Power ^a
	Variable	df			Squared	
Corrected Model	C3F	2	1.033	.358	.012	.229
	C3R	2	7.344	.001**	.077	.936
Intercept	C3F	1	6828.683	.000**	.975	1.000
	C3R	1	270.977	.000**	.605	1.000
Group	C3F	2	1.033	.358	.012	.229
	C3R	2	7.344	.001**	.077	.936

a. Computed using alpha = .05

** $p < 0.01$; * $p < 0.05$

The post-hoc Games-Howell test confirmed significant pairwise differences between the construct treatment and both other instructional methods (summarize and counter-argue) for revisions, but not between summarize and counter-argue. The construct treatment group generated a mean difference of 2.02 more revisions than to the summarize group with 95% confidence limits from .43 to 3.61, and a mean difference of 1.98 more revisions than the counter-argue group with 95% confidence limits from .44 to 3.52. Both differences were significant at $p < .01$. There were no significant differences observed among the treatment groups on the final assignment score for the case.

Table . Games-Howell pairwise comparisons

Variable	(I) Group	(J) Group	Mean		95% CI	
			Difference (I-J)	Sig.	LL	UL
Case 3 Final Score	Summarize	Counter-argue	-3.34	.395	-9.43	2.75
		Construct	-2.79	.556	-9.20	3.61
	Construct	Summarize	2.79	.556	-3.61	9.20
		Counter-argue	-.55	.966	-5.76	4.66
	Counter-argue	Summarize	3.34	.395	-2.75	9.43
		Construct	.55	.966	-4.66	5.76
Case 3 Revisions	Summarize	Counter-argue	-.03	.997	-1.16	1.09
		Construct	-2.02	.009**	-3.61	-.43
	Construct	Summarize	2.02	.009**	.43	3.61
		Counter-argue	1.98	.008**	.44	3.52
	Counter-argue	Summarize	.03	.997	-1.09	1.16
		Construct	-1.98	.008**	-3.52	-.44

Note. Based on observed means. The error term is Mean Square(Error) = 10.896.

** p < 0.01; * p < 0.05

RQ3: Website Usage

As previously noted, data associated with website usage was gathered anonymously using Google Analytics. Since individual sessions were not associated with individual participants, activity on the site cannot be correlated with performance data. However, because participants in the different treatment groups used identical sets of web pages in different directories, the usage data can be used to compare aggregate differences among the groups.

Two types of comparisons were employed on the website usage data from the third case in order

to examine differences in web usage among the treatment groups. The first comparison examined whether there were statistically significant differences in page views, average time per page, average page depth per visit, and average visit duration among the groups. The second comparison used quantitative page view data to visualize the most frequently traveled paths that different groups used when navigating the site.

Statistical analysis of usage statistics. Since Google Analytics data are anonymized and aggregated, there is no way to associate website usage data with individual participants. While it would be most helpful to know whether certain website usage behaviors were correlated with higher or lower performance on the argumentative essays, it was not possible to analyze the data with the participants, or even merely their sessions, as the subjects using this aggregated data from Google Analytics. However since the pages themselves are the same across the treatment groups, it is possible to consider those pages effectively as one group of subjects to which all three treatments were applied (as contrasted with the performance data where three groups of subjects received three different treatments). The dependent variables used in this analysis, then, represented how the pages performed under the different treatment methods, thus providing an indirect mechanism for assessing how the student participants interacted with the site at an aggregate level by treatment group.

Page views represented how many times each page was accessed during the study period. In order to compare this across treatment groups containing unequal numbers of participants, however, the number of participants in the treatment groups had to be accounted for in some way. Further, since the website usage data was gathered anonymously, it necessarily included participants who were omitted from the previous analyses, either because they had not consented to have their performance data included or because they were outliers for the purposes of a particular method. To address these issues, the figure used to represent the number of participants in each treatment group was calculated by counting the total number of participants who had

submitted the third case assignment. Then, rather than using the value for page views for each page directly, a ratio value calculated from the number of page views divided by the number of participants in the treatment group was substituted as a dependent variable.

Other Google Analytics values were able to be used directly as dependent variables because they represented averages of the aggregated data. Average time per page, average pages per visit, and average visit duration for a treatment group could be compared directly with those averages from other treatment groups because they were not skewed by the number of participants in each group.

While the original intent was to perform a MANOVA on this data using treatment group as an independent variable with page view ratio, average time per page, average pages per visit, and average visit duration as dependent variables, tests of normality showed the dependent variables to be significantly skewed to the left. MANOVA can be considered to be robust to minor violations of the assumption of normality with samples of at greater than 30 subjects, but the violation was judged to be too severe. As shown in Table , both standard tests of normality showed that normality could not be safely assumed in this data.

Table . Tests of normality

	Case	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
PV Ratio	College Applicants	.269	99	.000	.619	99	.000
Avg Pages Per Visit	College Applicants	.231	99	.000	.741	99	.000
Avg. Time on Page	College Applicants	.179	99	.000	.810	99	.000
Avg. Visit Duration	College Applicants	.207	99	.000	.805	99	.000

a. Lilliefors Significance Correction

Consequently, a Kruskal-Wallis H Test for non-parametric distributions was performed. Significant differences among the treatment groups were found for the ratio of pageviews, $\chi^2(2) = 17.757$, $p = .000$; average number of pages per visit, $\chi^2(2) = 9.733$, $p = .008$; and average duration of visit, $\chi^2(2) = 6.631$, $p = .036$.

Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. The pageview ratio was statistically significantly different between the counter-argue and construct treatment groups ($p = .000$) only. Similarly, the counter-argue group was statistically significant from the construct group in average number of pages per visit ($p = .010$), but not from the summarize group. The significance of the difference between the summarize and construct groups was $p = .050$. The pairwise comparisons among the groups for average duration of visit were not statistically significant. The sample average ranks for the treatment groups for each of these usage statistics is shown in Figure .

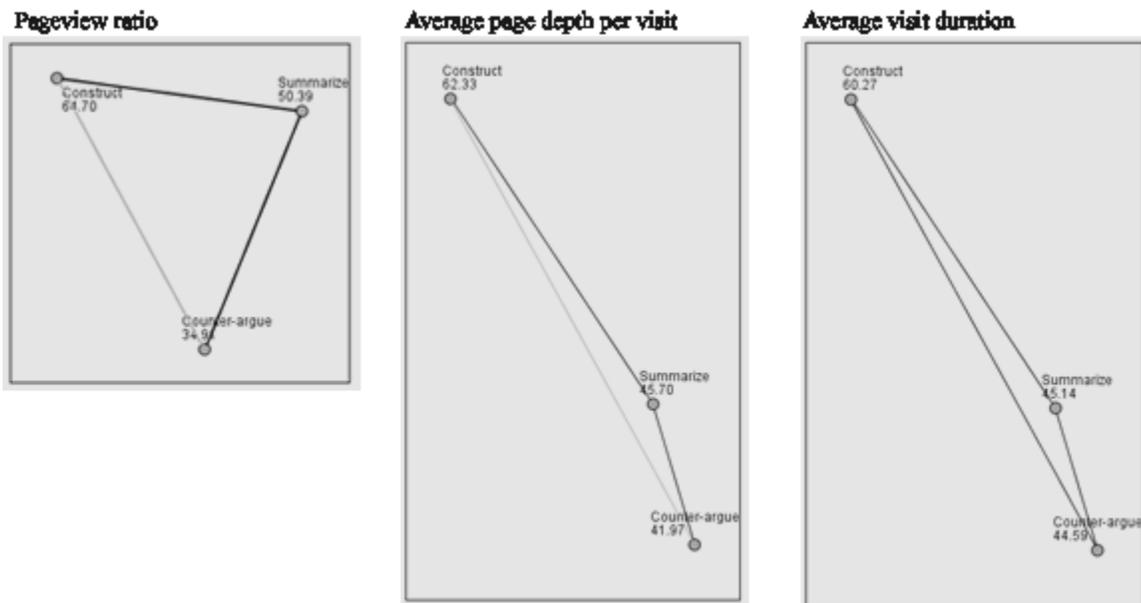


Figure . Average ranks for pairwise comparisons of website usage measures. Lighter gray lines represent statistically significant differences.

Analysis of page navigation. As noted in the methods section, three visualizations of the navigation data were created using Gephi (Bastian et al., 2009). In each of the visualizations, the pages form the nodes and the edges represent connections between the pages. The pages are represented textually in the visualizations by numeric codes. In the first two visualizations, the nodes are colored based on whether they are main case pages not associated with a candidate, case pages for each candidate, and pages from teaching cases to which the participants were permitted to refer when they worked on Case 3. A legend for the codes and colors associated with each page is provided in Figure .

	Code	Page Name
	1.1	case1
	1.3	case1-finish
	2.1	case2
	2.2	case2-research
	0.01	case3-cand1
	0.011	case3-cand1-conflict
	0.012	case3-cand1-social-interaction-theory
	0.013	case3-cand1-race-ethnicity
	0.014	case3-cand1-gender
	0.015	case3-cand1-social-class
	0.016	case3-cand1-letter
	0.017	case3-cand1-pres
	0.018	case3-cand1-politician
	0.02	case3-cand2
	0.021	case3-cand2-conflict
	0.022	case3-cand2-social-interaction-theory
	0.023	case3-cand2-race-ethnicity
	0.024	case3-cand2-gender
	0.025	case3-cand2-social-class
	0.026	case3-cand2-letter
	0.027	case3-cand2-pres
	0.028	case3-cand2-politician
	0.03	case3-cand3
	0.031	case3-cand3-conflict
	0.032	case3-cand3-social-interaction-theory
	0.033	case3-cand3-race-ethnicity
	0.034	case3-cand3-gender
	0.035	case3-cand3-social-class
	0.036	case3-cand3-letter
	0.037	case3-cand3-pres
	0.038	case3-cand3-politician
	0.1	case3
	0.2	case3-research
	0.3	case3-finish
	0.4	case3-meeting
	0.5	case3-characteristics
	0.6	case3-networks

Figure . Legend for network analysis visualizations

The first visualization for each group shows the respective centrality of the nodes for each group. Figures 9, 10, and 11 show the centrality visualization for the summarize, construct, and counter-argue treatment groups, respectively.

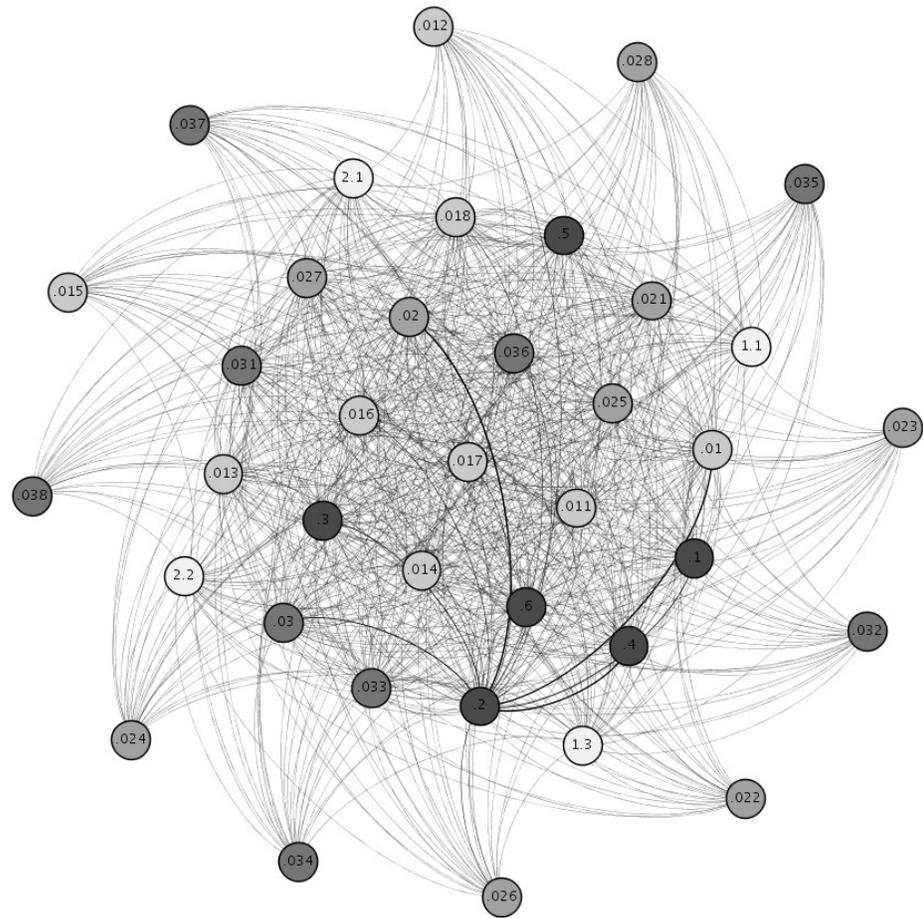


Figure . Centrality visualization for the summarize treatment group

The relative weight of the lines (edges) in the directionality graphs indicates connections with higher numbers of pages (edge weights). Heavier traffic is also indicated by the location of

unweighted degree as well as eigenvector centrality. No significant differences were found among the groups for any of the centrality measures.

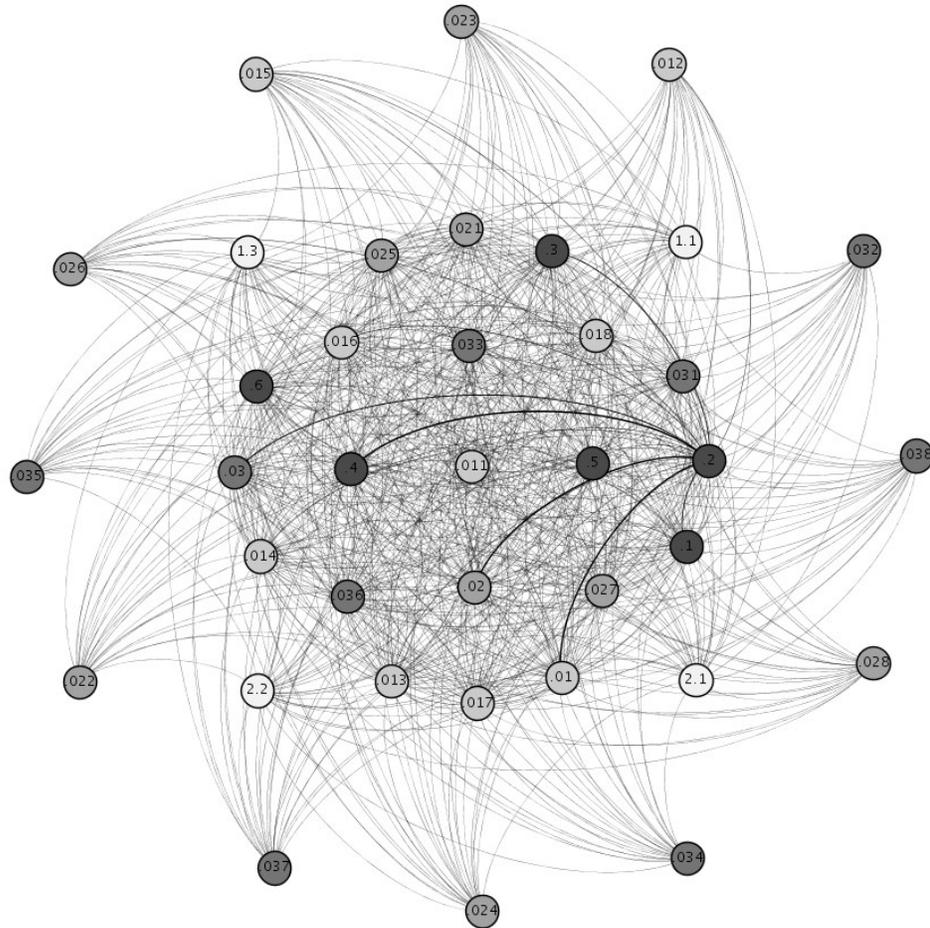


Figure : Centrality visualization for the counter-argue treatment group

A second visualization was created for each treatment group using a force-based circular layout with an edge-weight filter to allow differentiation between the diagrams (unfiltered diagrams are so complex that it becomes impossible to distinguish the flows in a circular layout).

The other settings remained the same as in the ARF visualization. Figures 12, 13, and 14 show the circular layout graphs for the three treatment groups.

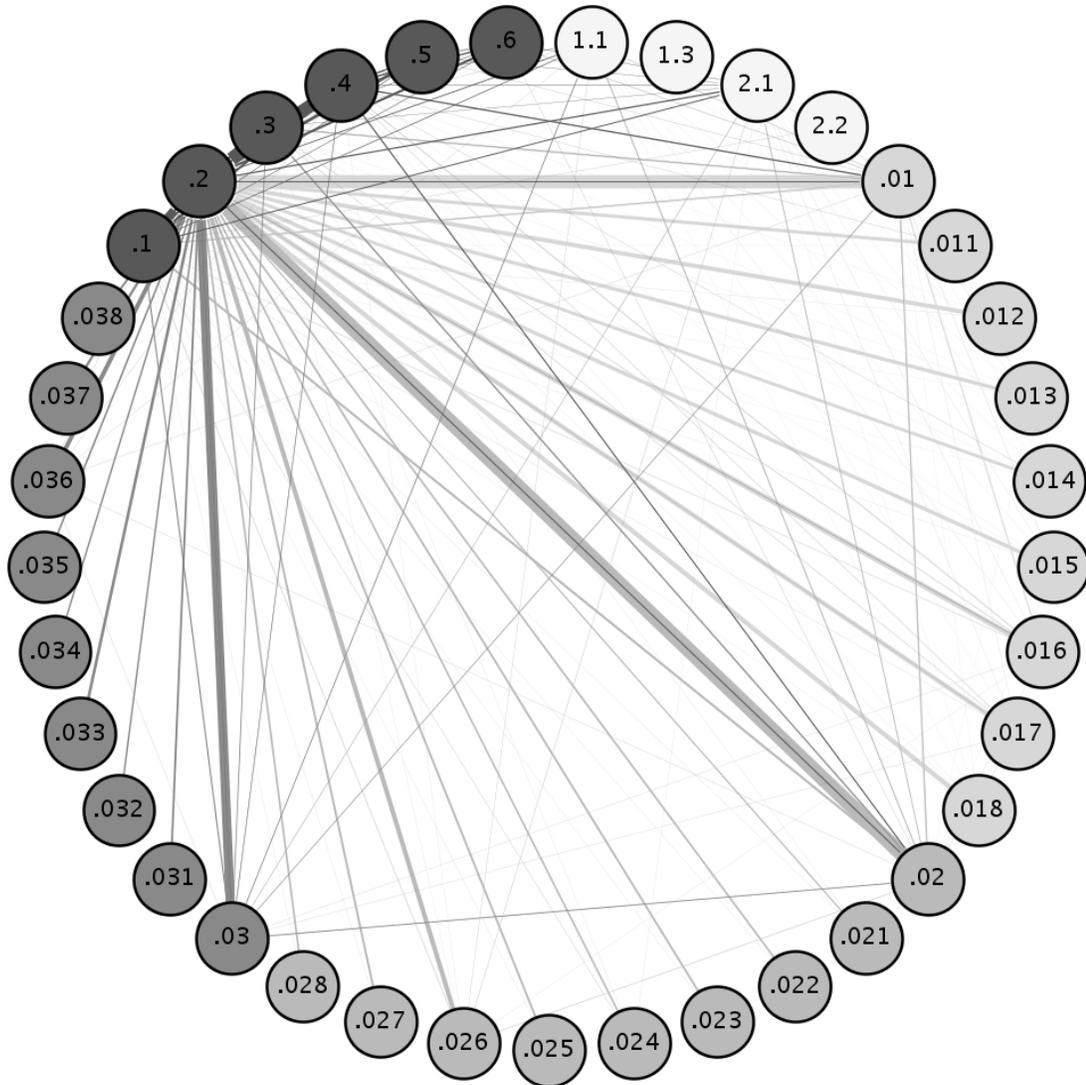


Figure : Circular layout for the summarize treatment group

This layout represents filtered data to give a more clear view. Thus, the lines that show in each circular graph do not reflect all of the connections from node to node; merely the strongest ones.

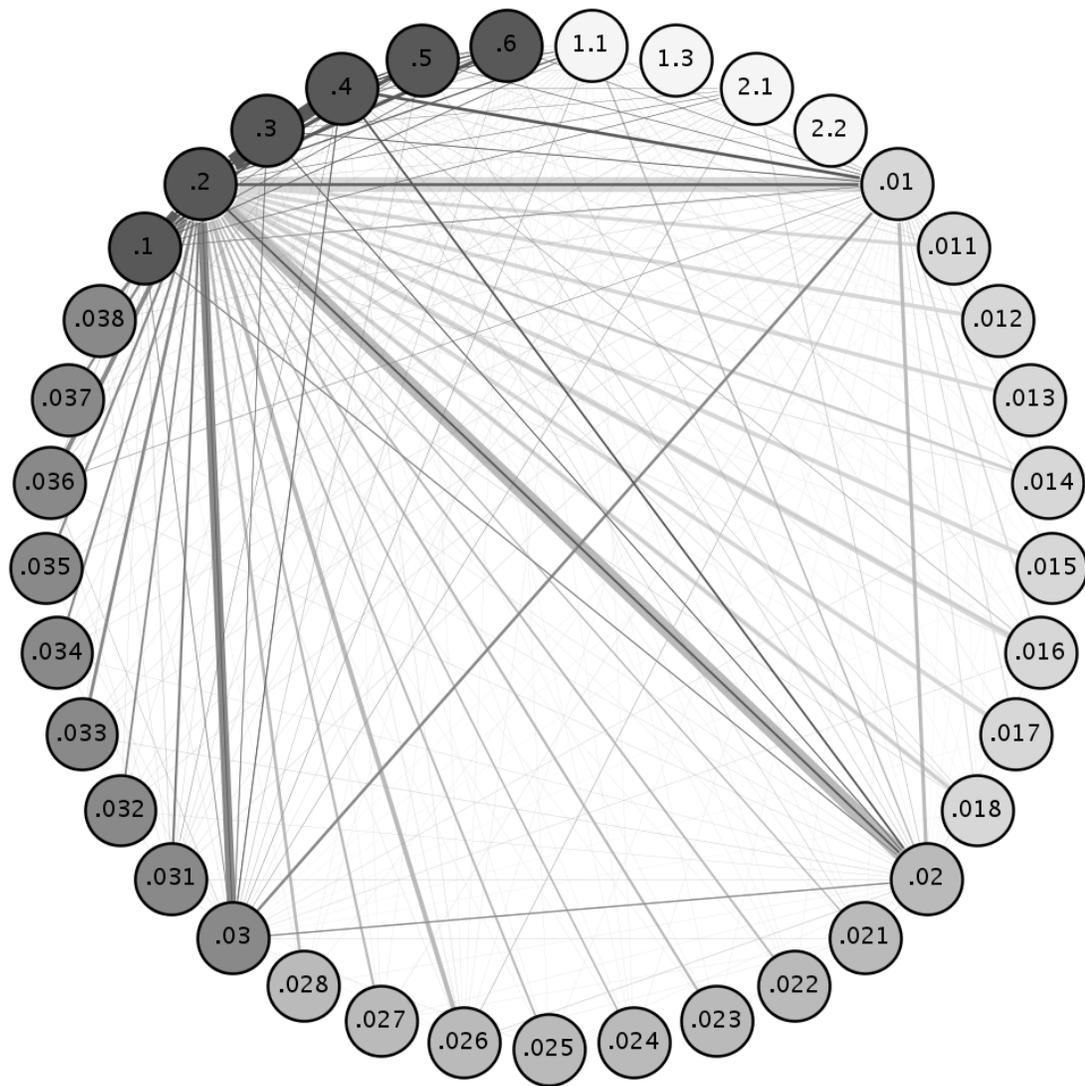


Figure : Circular layout graph for the construct treatment group

This visualization gives a bit clearer picture of the idea of navigating the problem landscape by treating all pages as equal rather than organizing them according to centrality. The visualization demonstrates the heavy influence of the “Research the Case” page (.2) to guide learners to the three candidate pages, and the ways that learners navigate from nearly all of the various pages for a particular candidate to the home page of another candidate.

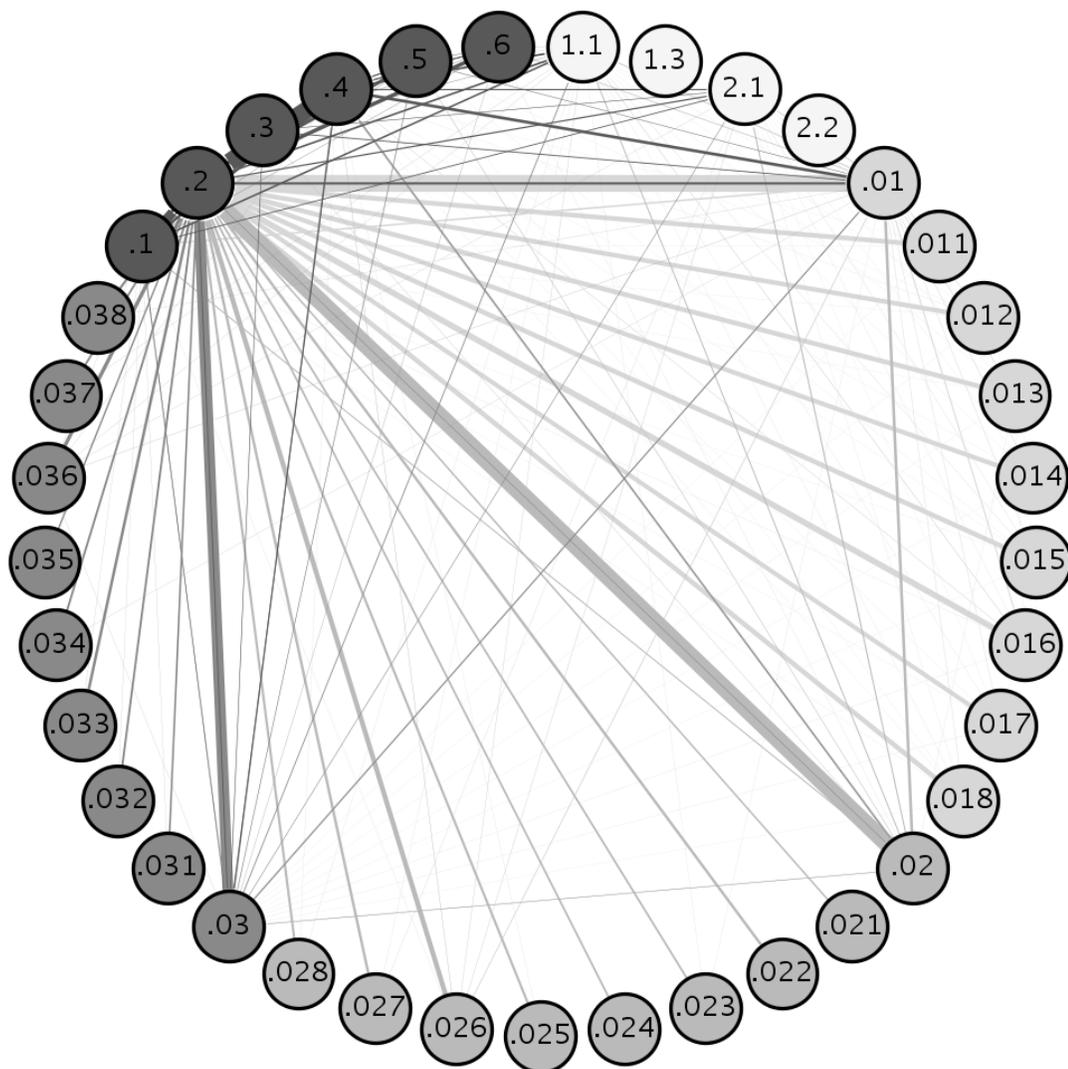


Figure : Circular layout graph for the counter-argue group

The final visualization created for each treatment group used a force-based dual circle layout that forced the main case pages not associated with a particular candidate outside the primary circle. The edge-weight filter settings from the circular visualization were again used to keep the edge lines reasonably distinct within the circle. This allows a view of the navigation that shows how the research page (.2) served as the main hub for navigating the site. By pulling the

main case pages out of the inner circle, it is easier to view the connections among the candidate pages.

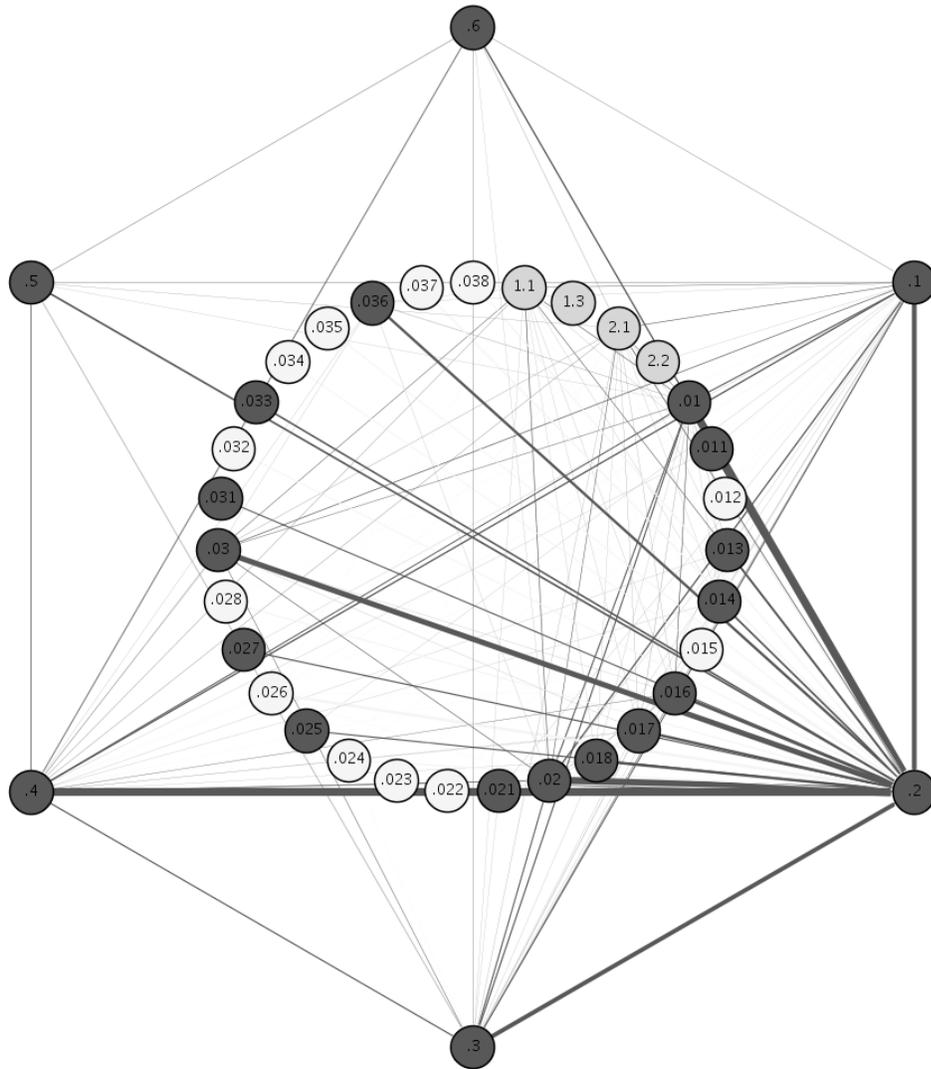


Figure : Dual-circular layout graph for the summarize treatment group

In this set of graphs, the node color is determined by the number of edges that connects to the node, also known as the node's degree. The degree in the graphs is without regard to source or target status, so a node with many outbound edges to other pages but few inbound edges could have a similar degree to a node with many inbound edges and few outbound edges.

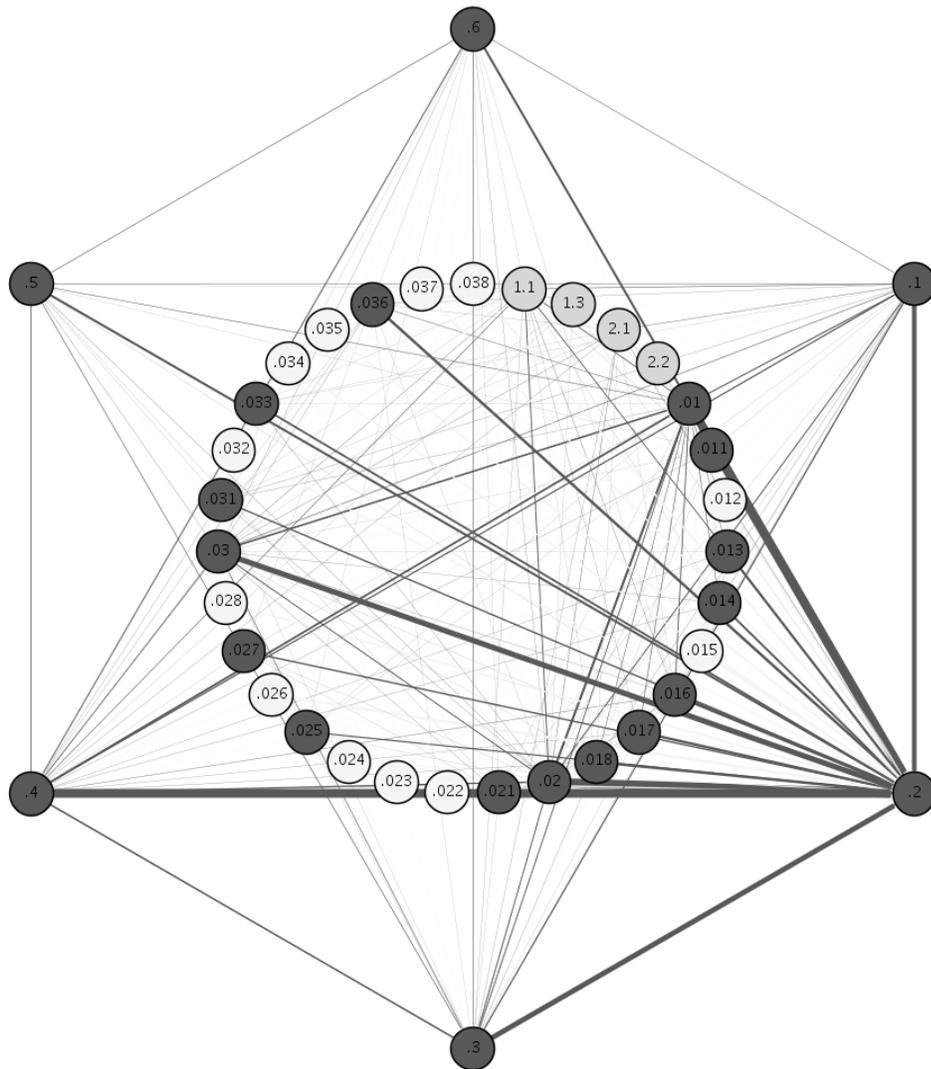


Figure : Dual-circular layout for the construct treatment group

The lines in the dual-circular layout graphs are colored according to the source node to give some directionality to this view of the connections. The weight of the edges represents the relative frequency of the connections they represent.

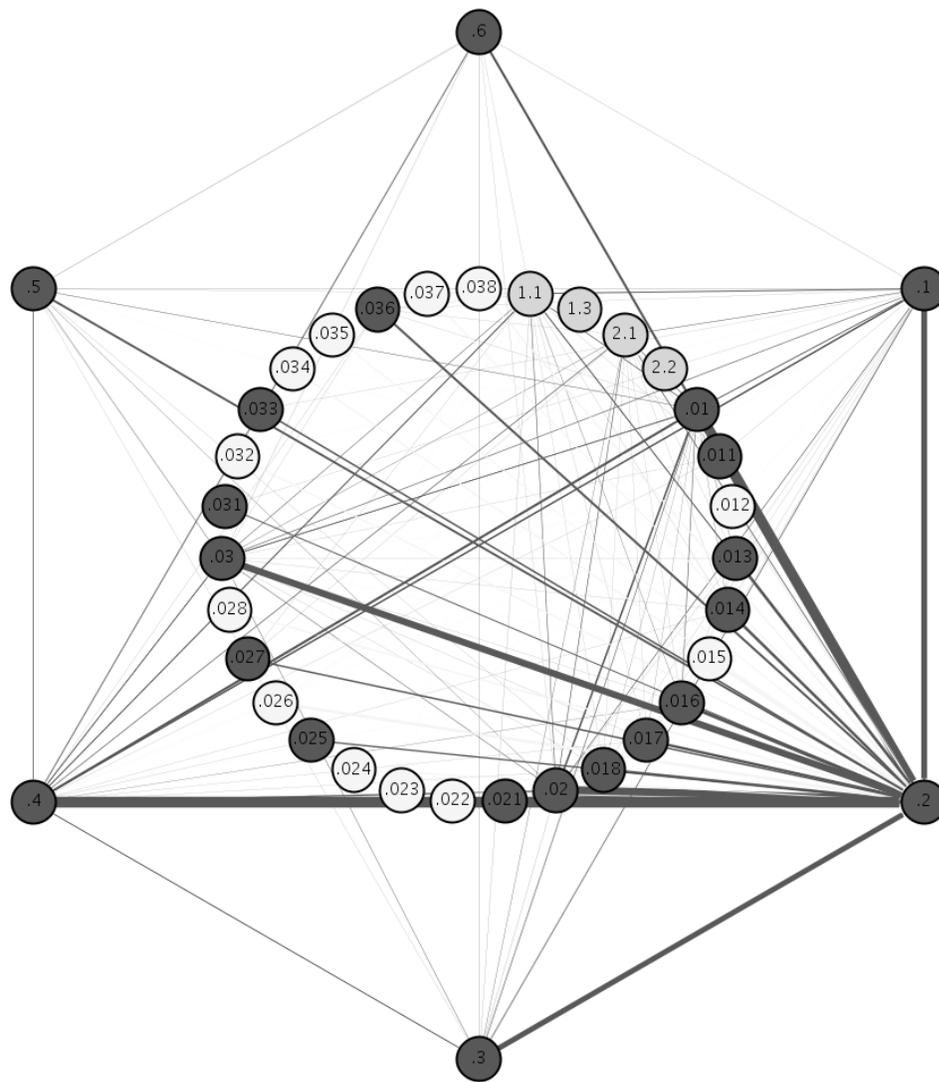


Figure : Dual-circular layout for the counter-argue treatment group

Chapter 5: Discussion

This chapter explores the results detailed in the preceding chapter in the context of the original research questions, the theoretical frameworks under which those questions were raised, and the related research that informed the design of this study. The chapter is organized in three sections that examine the efficacy of the instructional interventions employed in the study based on themes raised by the research questions: effects on argument quality, effects on the writing process, and effects on cognitive flexibility hypertext usage.

Effects on Argument Quality

When designing the treatments, it was hypothesized that learners who were given teaching cases designed to scaffold the construction of arguments or the generation of counterarguments would outperform learners who were given teaching cases with instructions to summarize the evidence provided in the case. The analysis of overall scores showed that there was no statistically significant effect on the quality of arguments produced by the three different treatment groups. This finding was contrary to that of Jonassen and Cho (2011) who did find significant differences in holistic essay scores for their counter-argue treatment condition when compared to their other treatment groups.

Similarly, in analyzing the related questions of impact of the treatments for males and females, no overall significant difference was detected for gender on performance, contradicting the National Report Card assessment (National Assessment of Educational Progress, 2008c) finding that females outperformed males, though a weak interaction effect was found between gender and treatment group. Subsequent post-hoc pairwise comparisons showed the effect to be insignificant in explaining the variances among student scores. The study found a significant though weak overall influence of first generation status on the participants' argument scores, but no interaction effect with group, and no significant pairwise difference in the post-hoc tests. The

result that first generation student status had some impact on student learning outcomes was consistent with the National Report Card assessment (National Assessment of Educational Progress, 2008c).

Given that the National Report Card assessment (National Assessment of Educational Progress, 2008c) found that females outperformed males at a statistically significant level and that the significant difference between first generation and non-first generation students showed a much higher power in the effect than the current study, it is possible that all of the forms of the instructional intervention in the introductory sociology course had somewhat of a mitigating effect on the anticipated demographic differences. It could be possible to study that possibility in a future research design that incorporated a pre-test to determine whether the first generation and male students in the study population demonstrated a lower initial ability to write an argumentative essay than their peers.

There were significant differences among the groups in the distributions of their component score achievement. Referring back to Figure in Chapter 4, the summarize treatment group included the following items more frequently than the construct and counter-argue groups: an item that questioned whether the committee had the authority to make the decision, an example of decision avoidance, an example of a hasty generalization, an example of judging evidence based on method, and an item related to topology (organizational structure). The construct group included the following items more frequently than the summarize and counter-argue groups: an example of method biases, an item related to communication in networks, and examples of statements supporting and opposing the white male candidate. Compared to the summarize and construct groups, the counter-argue group had a higher proportion of the following items: conclusions qualified by evidence and context, an example of false dichotomy, a statement about the need to make the decision quickly, and an example of the Peter Principle.

In general, the differences among the groups with regard to the types of evidence they tended to highlight in their essays are not definitive in exposing differences related to the treatment. It is interesting that the counter-argue group was more likely than the other two groups to use examples of false dichotomy and conclusions qualified by evidence and context. These could suggest that the method of constructing a counterargument to an existing model trained the counter-argue group to be more critical of the evidence that others might use to make their argument. However, the summarize and construct groups seemed to pay more attention than the counter-argue group to questions of structural influence on the arguments that the characters in the case were making as exemplified particularly in the topology and network communication items.

It is also curious to note that the construct group had higher proportions of examples of statements both supporting and opposing the white male candidate than the other groups. From the performance data set, it is not possible to determine which candidate a student elected to support, and so it could be that more of the construct group selected the white male candidate and thus were more likely to include arguments in support of that character and rebut arguments against him. Future research into the data that came out of this study could examine differences among students based on which candidate they chose.

Effects on the Writing Process

Based upon results from data associated with The Nation's Report Card (National Assessment of Educational Progress, 2008c) retrieved from the NAEP Data Explorer (National Assessment of Educational Progress, 2008a), it was hypothesized that higher revision attempts would result in higher scores. Thus, one area of interest in the current study was the impact of the three instructional treatments on the number of revision attempts as well as on the associated final scores for the assignment.

Simple regression suggested that, without regard to treatment group, the number of revisions can account for 20% of the variation in final scores for the Case 3 assignment. This is consistent with the expectation of the influence of revisions on performance in writing assignments noted above.

When comparing the treatment groups, instructional treatment was shown to have a weak effect on the number of revisions. The construct treatment group generated a mean difference of 2.02 more revisions than the summarize group and 1.98 more than the counter-argue group. The means for the summarize and counter-argue groups were nearly identical.

While the results did not indicate a statistically significant effect on final scores in Case 3, it is interesting to note that the mean score for the counter-argue group was slightly higher than that of the other groups given that the mean number of revisions for that group was lower than the construct group at a statistically significant level and roughly equivalent to that of the summarize group. This cannot be considered to contradict the suggested direction of the regression line, which showed a positive relationship between revisions and scores, or the NAEP findings discussed above since there is no statistically significant difference among the scores.

The lack of significant differences among the scores by treatment group combined with the significant differences among the number of revisions is, in and of itself, a significant finding because the regression model showed that, when treatment was not taken into account, scores rose along with revisions as expected. Consequently, whatever aspects of the instructional treatments contributed to the higher number of revisions for the construct group did not translate into the expected higher scores, suggesting that the assignments for the teaching cases for that group may have been less efficient in promoting the desired learning outcomes than those provided to the counter-argue group.

When considering the significant difference in number of revisions for the construct group in comparison to the counter-argue group in particular, one possible explanation for the

fact that the construct group did not achieve higher scores to correspond with the higher number of revisions could be that the simple act of providing a model in the teaching cases for the counter-argue group gave them a better understanding from the outset of Case 3 about what a fully constructed argument should look like. If so, the students in the counter-argue condition may have required fewer revisions to attain the final score that they desired. Referring back to Figure from the results of the analysis on overall scores from Chapter 4, the mean for the first revision score was in fact somewhat higher for the counter-argue group than for the other treatment groups, however as the results of that analysis showed, the difference was not a statistically significant one.

This conundrum of more revisions not equating to higher scores suggests the need for further research, particularly between groups trained with models of what will be expected of them in their essays and groups that are not. Jonassen (2011a) notes that using cases as worked examples are well-researched and shown to demonstrate efficacy in situations where the problems are well-structured, but questions whether it is possible to effectively model and ill-structured problem-solving process because there is no single, clear path to a solution and asserts that any attempt to do so would increase cognitive load for learners to an unacceptable level.

Still, he notes the gap in empirical research in using models when training students to solve ill-structured problems. It could further research in problem-solving in ill-structured domains to explore the circumstances under which expending the effort of further revisions not result in the reward of higher achievement as one opportunity to begin to address this gap. A qualitative analysis of the differences from revision to revision among participant essays in the construct and counter-argue treatment groups, particularly with a representative sample of high revision counts at different levels of final score achievement, might provide a way to explore this question in further detail.

Effects on Cognitive Flexibility Hypertext Usage

The three treatment groups were provided with mirror copies of the same website for each of the three cases used in the study. Only the instructions for the assignments differed among treatment groups, and only then for the two teaching cases. In Case 3, the transfer case, all groups were given the same assignment with the same instructions along with the same website from which to draw the evidence needed to construct an essay arguing for their candidate of choice. Thus, examining the differences in website usage in Case 3 provides some insight into the ways in which the differences in teaching case assignments trained them to use cognitive flexibility hypertexts to explore the problem space for Case 3.

The primary goal for design of a cognitive flexibility hypertext is to encourage learners to explore both the depth and breadth of the problem space while gaining an appreciation for the ways in which multiple perspectives, concepts, and overarching theories are interrelated. In this study, the differences between groups with regard to average number of pages per visit (page depth) and total number of pages viewed (pageviews) were examined to give insight into the questions of depth and breadth while network analysis techniques were employed to better understand how students navigated the site as an indicator of their exploration of the interrelatedness of the case material.

As was noted in Chapter 4 that the participants in the counter-argue group viewed fewer of the Case 3 web pages overall and by visit than the construct treatment group at statistically significant levels. This finding again suggests that the construct group expended additional effort to achieve effectively the same results as the counter-argue group. This also lends support to the idea that the models provided to the counter-argue group, or some other aspect of the instructional intervention provided to them, allowed them to achieve equivalent results with less effort than their counterparts as noted in the preceding section.

Three different sets of visualizations provided insight into the navigation habits of each of the three groups within the case website. Of particular interest in examining the centrality graphs are the similarities and differences among the group for the pages with high centrality and those with low centrality.

In the third case, the pages that discussed sociological concepts of structural inequality (race/ethnicity, gender, and social class) as well as the pages that covered the two theories which were to inform the students' arguments (conflict theory and social interaction theory) provided guidance in the form of question prompts at a high conceptual level rather than the more case-specific guidance in Case 2 and the very candidate-specific guidance in Case 1. In Case 3, these pages were the same for each of the three candidates. Likewise, the perspectives of the president of the University and the politician were general perspectives that applied to all three candidates and were not unique to each.

Consequently, it was hypothesized that students would refer more frequently to the pages for a candidate that provided information specific to that candidate; namely, their "home page" (.01, .02, and .03 for candidates 1, 2, and 3 respectively) and their letters of recommendation (.016, .026, and .036 for candidates 1, 2, and 3 respectively). Graphs where these pages are not relatively central for all three candidates may suggest a preference for certain candidates over others. For example, the letter for candidate 2 shows as not central for all three groups. Pages that provided information that was not candidate-specific were expected to be less central. Thus, in graphs where particular general pages showed high centrality, such as the president's letter for the discuss and construct groups or the race and ethnicity page for the counter-argue group, these may reflect either a high number of participants following similar paths through the material, confusion about the concepts on those pages, or stronger inclination to consider those concepts in constructing their own essays. It is possible that qualitative analysis of essays from the groups might provide additional data to explore those theories.

The circular and dual-circular graphs for each group give an idea of how thoroughly the treatment groups might have traversed the landscape of the problem space, which is a primary objective of the design of the website as a cognitive flexibility hypertext. The graphs show that the groups tended to visit all of the pages in the case, and that the traffic did seem to represent the desired non-linear traversal of the site that might typify “criss-crossing the landscape” of the problem space (Spiro et al., 1992).

The study was not designed to evaluate the efficacy of cognitive flexibility hypertexts to support problem-solving. To explore that question would necessitate comparing a cognitive flexibility hypertext with some alternative form of support and assess the impact on learning outcomes. Instead, it was hypothesized that if there were differences among the groups in their use of the Case 3 website, those differences could explain differences in scores and could point to the tendency of one treatment over another to promote deeper exploration of the problem space among the learners. Instead, the results showed that a higher level of effort in using the website (as demonstrated by a higher mean number of pageviews and a higher average number of pages per visit) did not translate into higher scores in neither the first nor the final draft of the assignment.

A subsequent study designed in a similar way could consider making use of a web usage tracking system that provided data on a session basis rather than aggregated as is the case with Google Analytics. Ideally, such a system could require a user login so that individual sessions could be tied to individual assignment scores which would allow more rigorous analysis of the relationship between website usage and performance. Participant essays for qualitative analysis of the relationship between website usage and performance could then be selected based upon criteria resulting from the analysis of website usage data.

Chapter 6: Summary

The purpose of the instructional intervention examined in this study was to foster argumentation in an introductory sociology course in order to situate student learning of structural inequality in the process of solving ill-structured problems typical of those that form the heart of the discipline. The intervention was designed around theoretical frameworks associated with problem-solving, argumentation, and using Cognitive Flexibility Hypertexts to represent the multiple perspectives and considerations typical of ill-structured problem domains.

This study focused on which, if any, of the three treatment versions of the instructional intervention enhanced student learning about how to argue, resulted in making more use of the revision process to hone student writing, and promoted more exploration of the facets of the problem. This chapter summarizes the results and discusses their implication for practice and for future research.

Summary of Results

There were significant differences among the three treatment groups in terms of what could be characterized as the effort that they expended, but not in their outcomes. The construct treatment group, which was instructed in the teaching cases to choose a candidate and build an argument to support that candidate step by step, spent more time on the transfer case website and visited more pages within it than the counter-argue group. The construct group also generated significantly more revisions from first to final draft of their case 3 essays than either the summarize or counter-argue treatment groups.

Despite this additional effort, though, the construct group did not outperform the other treatment groups. Teaching students to build an argument step-by-step could be said to be the least efficient of the three instructional methods studied in this case in terms of fostering argumentation. This would tend to support, albeit weakly, the significant results from Jonassen

and Cho (2011). This conclusion must be qualified with the understanding that the effects uncovered here were not particularly strong. It is likely that there were other influences on the differences in revision attempts and on website usage that were not controlled for in this study.

The design of the website under the framework of cognitive flexibility theory appeared to be successful in that all three groups of students were shown to have “criss-crossed the landscape” (Spiro et al., 1992) of the problem in navigating the case. The navigation visualizations demonstrated that, at an aggregate level, the various pages representing the problem were all traversed and that no strong linear path was developed by any of the groups.

Implications for Practice

Practitioners should examine the results of this study in light of their own learning objectives. On the surface, these results suggest that training students to develop their skills in writing argumentative essays is most efficiently accomplished by giving them models to argue against. The learners in this study who were trained with this instructional method expended less effort than their peers trained using the alternative methods for effectively the same results.

Modeling in general is a longstanding instructional method. Giving students a model of what the instructor expected them to produce may have given students in the counter-argue treatment group an advantage in understanding the structural considerations of what a successful argumentative essay looks like, such as expected length; number of warrants, counterarguments, rebuttals, etc.; and the order of elements within an essay. However, given Jonassen’s (2011a) concerns about models of problem-solving in ill-structured domains, a note of caution is warranted because the current study did not explore whether the model gave students an artificially constrained understanding of valid solutions given that the model represented only one possible solution path.

Since the models also highlighted certain evidence from within the teaching cases for which they were written, they might also have served to suggest what types of evidence were

most important in crafting a successful argument. It is important to remember that none of the groups had a model essay written about the specific evidence in the transfer case. However, by using specific evidence from the teaching case for which it was written, each model effectively gave one example of how to select evidence from a case about choosing among candidates to make an argument for one's candidate of choice.

On the other hand, it is important to consider whether argumentation is the only goal even if it is the primary one. The group trained to construct arguments step-by-step in the teaching cases likely spent more time on the tasks associated with the transfer case assignment than the group trained with models. This additional time on task could represent additional benefits in terms of longer-term retention of concepts and theories relevant to the subject matter that were not tested in this study. For many instructors situated within disciplines, the primary objective is not merely to improve argumentative essay skills, but to use the argumentative essay as a tool for demonstrating an understanding of how to apply subject matter knowledge in a particular context.

In addition, the fact that the study investigates only short term transfer effects doesn't illuminate which instructional method might have had longer-term transfer on the ability to structure an effective argument in essay form. One of the measures in which the construct group differed significantly from the counter-argue group was the mean number of revisions from first to final drafts of their essays. The additional practice provided by those additional revisions could have had effects on how well learners internalized structural aspects of the argumentative essay since the counter-argue group could have been operating from memory of the teaching cases that might fade over time.

In light of these considerations, it is apparent that the results of this study do not imply that one instructional method is superior to another. The fact that each treatment group in this study showed some differences in the evidence that they offered in their essays, all of which could be used in an effective argument, could reflect the possibility that each of the methods has

its own strengths. In fact, since all three methods used in this study are based on a variety of best practices in teaching, the best result may be achieved by using a combination of them over time in a larger set of teaching cases.

The study did not test the efficacy of cognitive flexibility hypertexts for promoting argumentation since no comparison was made between groups who used the case website and groups who used some other presentation of the case. Thus, it cannot be inferred from these results that cognitive flexibility hypertext is an effective means of representing the types of ill-structured problems typical of sociology. The use of cognitive flexibility theory to inform the design of the case websites was based upon earlier studies discussed in Chapter 2 (cf. Cernusca, 2007; Dabbagh, 2002; Jacobson & Spiro, 1995) that did make such comparisons. The analysis of the navigational patterns of the participants in the study showed that participants demonstrated behavior that could be characterized as “criss-crossing the landscape” (Spiro et al., 1992) of the problem; that is, that they used the design as intended.

For practitioners, this suggests that non-linear navigation design encourages the type of learner behavior sought after in cognitive flexibility theory. The case website’s non-linear design encouraged learners to follow a variety of paths in exploring the problem space, which, according to the studies that informed that design, should have helped them to understand the ways that the multiple concepts, theories, and perspectives in the problem were interconnected. It also gave them some amount of locus of control which allowed them to choose their next move based on what they individually felt was most relevant to their needs at the time.

As noted in the literature discussion, this type of problem space navigation has been considered most appropriate to mastery-level learners rather than novices who would be expected to be enrolled in an introductory course such as the one in this study. However, the participants had completed foundational learning activities associated with learning the concepts of structural inequality prior to the activities described in this study. In addition, as Reigeluth has argued

(1999), care must be exercised when simplifying problems for novice learners to ensure that the simplification is not giving a false understanding about the scope of problems in the domain. For instructional material in which exposing learners to some amount of the complexity inherent in the domain is desirable, the design incorporated in this study can be used to inform how to approach non-linear presentation of that material.

Implications for Future Research

Several areas of opportunity for future research have been suggested in the discussion of the results of this study. First and foremost, the study leaves open the question of the efficacy of the construct versus counter-argue instructional methods initially raised in Jonassen and Cho (Jonassen & Cho, 2011). While this study provided some support for their conclusion that the counter-argue treatment better supported argumentation quality, it did so indirectly and thus weakly. Here, the counter-argue treatment was a more efficient instructional method because participants expended less effort than their counterparts, but it did not result in better outcomes.

Where there were significant differences among the groups, such as in revision attempts, the effect shown tended to explain little of the variance in the participants' scores on the assignment. The design of this study could be improved by implementing a pre-test assignment that would expose how well students are able to achieve the assignment goals prior to the implementation of the instructional methods used in the study as prior knowledge could account for a significant portion of the variance in the assignment scores.

It would be most interesting to delve into the different levels of effort among the groups in this treatment to find out why increased effort did not result in better outcomes. In terms of revisions, as noted in Chapter 5, a qualitative assessment of a purposive sample of high revision count essays with differing final scores could illuminate why higher revision attempts did not generally result in higher scores. It might also be beneficial to examine other demographic differences among the participants that could have affected either the number of revisions or

outcomes. Do students with particular majors revise more than others, for example? Does socio-economic status affect outcomes? Other than gender and first generation student status, individual differences among the students were not studied for their possible relationship to performance data. Another interesting variation might be to study whether students revised more when they had more or less time remaining until the deadline. For example, do procrastinators who start much closer to the deadline complete fewer revisions than do those students who allow themselves more time to work on the assignment?

Likewise, a subsequent study that refined the web usage data gathering techniques so that individual sessions could be associated with individual participants could give more insight into the relationship between what a learner does on the website and his/her performance on the transfer case assignment. This would allow not only a deeper examination of the relationship between effort in exploring the problem space and performance, but also allow for more precise comparison between the treatment groups.

Qualitative examination of essays might also afford a more detailed understanding of why significant differences existed among groups in distribution of achievement on component items. As previously mentioned, some differences in achievement of these items could result from the choice of candidate which was not captured in the quantitative data used in this study.

A key direction for future research is to examine the longer-term effects of instruction designed to foster argumentation. A future implementation of this or a similar study could focus on the relationship between performance in the activities associated with the instructional intervention and other, subsequent activities in the course, such as a final exam, which could highlight content knowledge effects, and/or a similar essay assignment later in the semester, which could highlight argumentation skill effects. Such research could address questions such as:

1. In ill-structured domains such as sociology, does learning to apply concepts and theories in this manner affect how well students retain their knowledge of the subject matter content?
2. Does their performance on subsequent argumentation assignments in the discipline, perhaps in subsequent courses, benefit from early application of the foundational material in this way?

An overarching theme that remains unanswered after this study is the efficacy of instruction in argumentation skill across a variety of ill-structured domains. While the current study was situated in sociology and the Jonassen and Cho (2011) study was situated in engineering, neither study compared differences in learning outcomes between students who engaged in the argumentative assignments and students who didn't. Further, these represent only two of a myriad of possible disciplines typified by ill-structured problems with more than one possible solution, which Jonassen (2011a) has asserted are often "solved" through argumentation. While it would likely be impossible to control for the many factors that might influence learner performance in different domains, it could be possible to target introductory courses where most of the students are freshmen to minimize some of the differences. Ideally, such a study should be designed to examine performance between a group in each discipline that received the argumentation instruction with one that did not on a final exam that included at least one argumentative essay so that variance due to the treatment would more likely be exposed.

Conclusion

This study examined three different treatment methods for fostering argumentation in an introductory sociology course. Overall learning outcomes did not show any significant differences among the methods, though there were differences in participant level of effort in completing the assignments and in the inclusion of particular elements of argument. While these results afford some insight into the merits of particular instructional methods for fostering

argumentation, more research is needed to determine why the differences occurred and whether there are other significant differences, such as in long-term knowledge and skill retention, that were not exposed in the current study.

References

- Abelev, M., Vincent, M. B., & Haney, T. J. (2008). The Bottom Line: An Exercise to Help Students Understand How Social Inequality Is Actively Constructed. *Teaching Sociology, 36*(2), 150-160.
- Atkinson, M. P., & Hunt, A. N. (2008). Inquiry-Guided Learning in Sociology. *Teaching Sociology, 36*, 1-7.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: An Open Source Software for Exploring and Manipulating Networks*. Paper presented at the International AAAI Conference on Weblogs and Social Media, San Jose, CA. Retrieved from: <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Bixler, B. A., & Land, S. M. (2010). Supporting College Students' Ill-Structured Problem Solving in A Web-Based Learning Environment. *Journal of Educational Technology Systems, 39*(1), 3-15. doi: 10.2190/ET.39.1.b
- Bloom, B. S. (1956). *Taxonomy of educational objectives; the classification of educational goals* (1st ed.). New York: Longmans, Green.
- Bowen, W. G., Kurzweil, M. A., Tobin, E. M., & Pichler, S. C. (2005). *Equity and excellence in American higher education* [Google Books Edition]. Retrieved from https://play.google.com/store/books/details?id=Ku60XoLE_ngC
- Brem, S. K., Russell, J., & Weems, L. (2001). Science on the Web: Student Evaluations of Scientific Arguments. *Discourse Processes, 32*(2), 191 - 213.
- Brent, E., Atkisson, C., & Green, N. (2010). Time-Shifted Online Collaboration: Creating Teachable Moments Through Automated Grading. In A. A. Juan, T. Daradoumis, F. Xhafa, S. Caballé & J. Faulin (Eds.), *Monitoring and Assessment in Online Collaborative Environments: Emergent Computational Technologies for E-Learning Support* (pp. 55-73). Hershey, PA: Information Science Reference.
- Brent, E., Carnahan, T., & McCully, J. (2006a). SAGrader™ Helps Disadvantaged Students Overcome Initial Deficits. Retrieved from https://www.sagrader.com/static/content/whitepapers/disadvantaged_students.pdf
- Brent, E., Carnahan, T., & McCully, J. (2006b). Students Improve Learning by 20 Percentage Points with Essay Grading Program, SAGrader™. Retrieved from https://www.sagrader.com/static/content/whitepapers/improves_learning.pdf
- Bridwell, L. S. (1980). Revising Strategies in Twelfth Grade Students' Transactional Writing. *Research in the Teaching of English, 14*(3), 197-222. doi: 10.2307/40170856

- Brislen, W., & Clayton, D. P. (2005). Using a Hypothetical Distribution of Grades to Introduce Social Stratification. *Teaching Sociology*, 33(1), 74-80.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1), 32-42.
- Cala Carrillo, M. J., & De La Mata Benítez Maria, M. L. (2004). Educational Background, Modes of Discourse and Argumentation: Comparing Women and Men. *Argumentation*, 18(4), 403-426. doi: 10.1007/s10503-004-4906-1
- Carter, C. S., & Robinson, R. (2002, April 1-5). "Can We Send Some of the Money Back Home to Our Families?" *Tensions of Transition in an Early Intervention Program for Rural Appalachian Students*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Center for the Study of Problem Solving. (2009a). Sociology 1000: Case 1, from <http://csps.missouri.edu/soc1000/SP09/case1.html>
- Center for the Study of Problem Solving. (2009b). Sociology 1000: Case 2, from <http://csps.missouri.edu/soc1000/SP09/case2.html>
- Center for the Study of Problem Solving. (2009c). Sociology 1000: Case 3, from <http://csps.missouri.edu/soc1000/SP09/case3.html>
- Cernusca, D. (2007). *A design-based research approach to the implementation and examination of a cognitive flexibility hypertext in a large undergraduate course*. Ph.D. Dissertation, University of Missouri - Columbia, United States -- Missouri. Retrieved from [https://mospace.umsystem.edu.proxy.mul.missouri.edu/xmlui/handle/10355/4810\(3351627\)](https://mospace.umsystem.edu.proxy.mul.missouri.edu/xmlui/handle/10355/4810(3351627))
- Chan, Y., & Walmsley, R. P. (1997). Learning and understanding the Kruskal-Wallis one-way analysis-of-variance-by-ranks test for differences among three or more independent groups. *Phys Ther*, 77(12), 1755-1762.
- Chiu, M. M. (2008). Effects of argumentation on group micro-creativity: Statistical discourse analyses of algebra students' collaborative problem solving. *Contemporary Educational Psychology*, 33(3), 382-402.
- Cho, K.-L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, 50(3), 5-22.
- Choy, S. P. (2001). *Students whose parents did not go to college: Postsecondary access, persistence and attainment*. (NCES 2001-126). Washington, D.C.: U.S. Department of Education Retrieved from <http://nces.ed.gov/pubs2001/2001126.pdf>.

- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). London ; New York: Routledge.
- Collier, P. J., & Morgan, D. L. (2008). "Is that paper really due today?": differences in first-generation and traditional college students' understandings of faculty expectations. *Higher Education*, 55(4), 425-446. doi: 10.1007/s10734-007-9065-5
- Creswell, J. W. (2008). *Educational research : planning, conducting, and evaluating quantitative and qualitative research* (3rd ed. ed.). Upper Saddle River, N.J.: Pearson Prentice Hall.
- Dabbagh, N. H. (2002). Assessing Complex Problem-Solving Skills and Knowledge Assembly Using Web-Based Hypermedia Design. *Journal of Educational Multimedia and Hypermedia*, 11(4), 291-322.
- Demetriadis, S., & Pombortsis, A. (1999). Novice student learning in case based hypermedia environment: a quantitative study. *Journal of Educational Multimedia and Hypermedia*, 8(2), 241-269.
- Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, 6(3), 241-252.
- Ewert, S. (2012). *What it's Worth: Field of Training and Economic Status in 2009*. Washington, DC: U.S. Census Bureau Retrieved from <http://www.census.gov/prod/2012pubs/p70-129.pdf>.
- Fisher, E. M. (2008). USA Stratified Monopoly: A Simulation Game About Social Class Stratification. *Teaching Sociology*, 36, 272-282.
- Fitzgerald, J. (1987). Research on Revision in Writing. *Review of Educational Research*, 57(4), 481-506. doi: 10.2307/1170433
- Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, 32(4), 365-387. doi: 10.2307/356600
- Ge, X. (2001). *Scaffolding students' problem-solving processes on an ill-structured task using question prompts and peer interactions*. Ph.D Dissertation, The Pennsylvania State University, United States -- Pennsylvania.
- Ge, X., & Land, S. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development*, 51(1), 21-38.
- Guerra-López, I. (2008). *Performance Evaluation: Proven Approaches for Improving Program and Organizational Performance*. San Francisco, CA: Jossey-Bass.

- Harlow, R. (2009). Innovations in Teaching Race and Class Inequality: "Bittersweet Candy" and "The Vanishing Dollar". *Teaching Sociology*, 37(2), 194-204.
- Hattery, A. J. (2003). Sleeping in the Box, Thinking outside the Box: Student Reflections on Innovative Pedagogical Tools for Teaching about and Promoting a Greater Understanding of Social Class Inequality among Undergraduates. *Teaching Sociology*, 31(4), 412-427.
- Hernandez-Serrano, J., & Jonassen, D. H. (2003). The effects of case libraries on problem solving. *Journal of Computer Assisted Learning*, 19(1), 103-114. doi: 10.1046/j.0266-4909.2002.00010.x
- Howe, C., & McWilliam, D. (2001). Peer Argument in Educational Settings. *Journal of Language and Social Psychology*, 20(1-2), 61-80. doi: 10.1177/0261927x01020001003
- Ideaworks Inc. (2008). SAGrader, from <http://www.sagrader.com/>
- Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *journal of educational computing research*, 12(4), 301.
- Jeong, A., & Davidson-Shivers, G. (2006). The Effects of Gender Interaction Patterns on Student Participation in Computer-Supported Collaborative Argumentation. *Educational Technology Research and Development*, 54(6), 543-568. doi: 10.1007/s11423-006-0636-4
- Jeong, A. C. (2007). The Effects of Intellectual Openness and Gender on Critical Thinking Processes in Computer-Supported Collaborative Argumentation. *Journal of Distance Education*, 22(1), 1-18.
- Jessup, M. M. (2001). Sociopoly: Life on the Boardwalk. *Teaching Sociology*, 29(1), 102-109.
- Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63-85.
- Jonassen, D. H. (2011a). *Learning to solve problems : a handbook for designing problem-solving learning environments*. New York: Routledge.
- Jonassen, D. H. (2011b). Supporting problem solving in PBL. *Interdisciplinary Journal of Problem-based Learning*, 5(2).
- Jonassen, D. H., & Cho, Y. H. (2011). Fostering argumentation while solving engineering ethics problems. *Journal of Engineering Education*, 100(4), 1-23.

- Jonassen, D. H., Cho, Y. H., Easter, M., Henry, H. R., & Kwon, K. (2010). *Eliciting counterarguments in ethics problems*. Paper presented at the American Educational Research Association, Denver, CO.
- Jonassen, D. H., Cho, Y. H., Kwon, K., Henry, H. R., & Easter, M. (2009). *Solving Ethical Dilemmas Through Argumentation*. Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Louisville, KY.
- Jonassen, D. H., Cho, Y. H., Kwon, K., Henry, H. R., Shen, D., Easter, M., & Lo, J. (2009). *Evaluating Versus Constructing Arguments*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Jonassen, D. H., Shen, D., Marra, R., Cho, Y. H., Lo, J., & Lohani, V. (2009). Engaging and Supporting Problem Solving in Engineering Ethics. *Journal of Engineering Education*, 98(3), 235-254.
- Kerbo, H. R. (2006). Social Stratification. In C. D. Bryant & D. L. Peck (Eds.), *21st Century Sociology: A Reference Handbook* (Vol. 1, pp. 228-236). Thousand Oaks, CA: Sage Publications Inc.
- Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). *What Matters to Student Success: A Review of the Literature*. Washington, DC: National Postsecondary Education Cooperative, Retrieved from http://nces.ed.gov/npec/pdf/Kuh_Team_Report.pdf.
- Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking & Reasoning*, 13(2), 90 - 104.
- Leech, N. L., Barrett, K. C., Morgan, G. A., Clay, J. N., & Quick, D. (2005). *SPSS for Intermediate Statistics: Use and Interpretation* (Second ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lund, A., & Lund, M. (2012). Kruskal-Wallis Test in SPSS. *Laerd Statistics* Retrieved September 30, 2012, from <https://statistics.laerd.com/premium/kruskal-wallis-test-in-spss.php>
- MacArthur, C. A. (2007). Best Practices in Teaching Evaluation and Revision. In S. Graham, C. A. MacArthur & J. Fitzgerald (Eds.), *Best practices in writing instruction: Solving problems in the teaching of literacy* [Kindle Edition]. New York: Guilford Press. Retrieved from http://www.amazon.com/gp/product/B005DIASAQ/ref=docs-os-doi_0
- McCammom, L. (1999). Introducing Social Stratification and Inequality: An Active Learning Technique. *Teaching Sociology*, 27(1), 44-54.

- Meyer, J. P., & Seaman, M. A. (in press). A comparison of the Exact Kruskal-Wallis Distribution to Asymptotic Approximations for All Sample Sizes Up to 105. *Journal of Experimental Education*. Retrieved from <http://faculty.virginia.edu/kruskal-wallis/paper/A%20comparison%20of%20the%20Exact%20Kruskal-v4.pdf>
- National Assessment of Educational Progress. (2007). *Writing: Student Background Questionnaire, Grade 12*. Washington, DC: U.S. Department of Education Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/bgq/student/BQ07-NAEP-student-writing-g12.pdf>.
- National Assessment of Educational Progress. (2008a). NAEP Data Explorer. In National Center for Education Statistics (Ed.). Washington, D.C.
- National Assessment of Educational Progress. (2008b). NAEP Questions Tool. In National Center for Education Statistics (Ed.). Washington, D.C.
- National Assessment of Educational Progress. (2008c). *The Nation's Report Card: Writing 2007*. Washington, DC: US Department of Education Retrieved from <http://nces.ed.gov/nationsreportcard/pdf/main2007/2008468.pdf>.
- Newman, M. E. J. (2010). *Networks : an Introduction* [Kindle version]. Retrieved from Amazon.com
- Nichols, L., Berry, J., & Kalogrides, D. (2004). Hop on the Bus: Driving Stratification Concepts Home. *Teaching Sociology*, 32(2), 213-221.
- Nunez, A.-M., & Cuccaro-Alamin, S. (1998). *First-Generation Students: Undergraduates Whose Parents Never Enrolled in Postsecondary Education*. (NCES 98-082). Washington, DC: U.S. Department of Education Retrieved from <http://nces.ed.gov/pubs98/98082.pdf>.
- Nussbaum, E. M. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology*, 30(3), 286-313.
- Nussbaum, E. M., & Bendixen, L. D. (2003). Approaching and avoiding arguments: The role of epistemological beliefs, need for cognition, and extraverted personality traits. *Contemporary Educational Psychology*, 28(4), 573-595.
- Nussbaum, E. M., & Kardash, C. M. (2005). The Effects of Goal Instructions and Text on the Generation of Counterarguments During Writing. *Journal of Educational Psychology*, 97(2), 157-169. doi: 10.1037/0022-0663.97.2.157
- Nussbaum, E. M., & Schraw, G. (2007). Promoting Argument-Counterargument Integration in Students' Writing. *Journal of Experimental Education*, 76(1), 59-92. doi: 10.3200/jexe.76.1.59-92

- Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28*(3), 384-395.
- Office of Postsecondary Education. (2011, 9/9/2011). History of the Federal TRIO Programs Retrieved 3/24/2012, 2012, from <http://www2.ed.gov/about/offices/list/ope/trio/triohistory.html>
- Osana, H. P., Tucker, B. J., & Bennett, T. (2003). Exploring adolescent decision making about equity: Ill-structured problem solving in social studies. *Contemporary Educational Psychology, 28*(3), 357-383.
- Pascarella, E. T., Pierson, C. T., & Wolniak, G. C. (2004). First-Generation College Students: Additional Evidence on College Experiences and Outcomes. *Journal of Higher Education, 75*(3), 249-284. doi: 10.1353/jhe.2004.0016
- Pascarella, E. T., Whitt, E. J., Nora, A., Edison, M., Hagedorn, L. S., & Terenzini, P. T. (1995). *What Have We Learned from the First Year of the National Study of Student Learning?* Washington, DC: Office of Educational Research and Improvement Retrieved from <http://www.eric.ed.gov/PDFS/ED381054.pdf>.
- Pence, D. J., & Fields, J. A. (1999). Teaching about Race and Ethnicity: Trying to Uncover White Privilege for a White Audience. *Teaching Sociology, 27*(2), 150-158.
- Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday Reasoning and the Roots of Intelligence. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education*. Hillsdale, N.J.: L. Erlbaum Associates. Retrieved from <http://www.questia.com/read/99104792/informal-reasoning-and-education>.
- Persell, C. H., Pfeiffer, K. M., & Syed, A. (2007). What Should Students Understand After Taking Introduction to Sociology? *Teaching Sociology, 35*, 300-314.
- Pike, G. R., & Kuh, G. D. (2005). First- and Second-Generation College Students: A Comparison of Their Engagement and Intellectual Development. [Article]. *Journal of Higher Education, 76*(3), 276-300.
- Reigeluth, C. (1999). The elaboration theory: Guidance for scope and sequence decisions. *Instructional design theories and models: A new paradigm of instructional theory, 2*, 425-453.
- Renzulli, L. A., Aldrich, H. E., & Reynolds, J. (2003). It's up in the Air, or Is It? *Teaching Sociology, 31*(1), 49-59.
- Rijlaarsdam, G., & van den Bergh, H. (2006). Writing Process Theory: A Functional Dynamic Approach. In C. A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research* [Kindle Edition]. New York: Guilford Press.

Retrieved from http://www.amazon.com/gp/product/B001QCYOAI/ref=docs-os-doi_0.

- Ross, S. M., & Hurlbert, J. M. (2004). Problem-Based Learning: An Exercise on Vermonts Legalization of Civil Unions. *Teaching Sociology*, 32, 79-93.
- Sernau, S. (1995). Using a Collaborative Problem-Solving Approach in Teaching Social Stratification. *Teaching Sociology*, 23(4), 364-373.
- Shen, D. (2008). *A comparison of embedded links and question links in cognitive flexibility hypertext (CFH) learning environments for problem solving in engineering ethics* Ph.D. Dissertation, University of Missouri - Columbia, United States -- Missouri. Retrieved from <https://mospace.umsystem.edu/xmlui/handle/10355/6689>
- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). Cognitive Flexibility Theory: Advanced Knowledge Acquisition in Ill-Structured Domains *Tenth Annual Conference of the Cognitive Science Society* (pp. 375-383). Hillsdale, NJ: Erlbaum.
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill-Structured Domains. In D. H. Jonassen & T. M. Duffy (Eds.), *Constructivism and the technology of instruction : a conversation* (pp. 121-128). Hillsdale, N.J.: Lawrence Erlbaum Associates Publishers.
- Toulmin, S. (1958). *The uses of argument*. Cambridge England: University Press.
- Toulmin, S. (2003). *The uses of argument* (Updated ed.). Cambridge, U.K. ; New York: Cambridge University Press.
- van den Bergh, H., Rijlaarsdam, G., & Breetvelt, I. (1993). Revision process and text quality: An empirical study. In G. Eigler & T. Jechle (Eds.), *Writing: Current trends in European research* (pp. 133-147). Freiburg: Hochschul Verlag.
- Wetcher-Hendricks, D., & Luquet, W. (2003). Teaching Stratification with Crayons. *Teaching Sociology*, 31(3), 345-351.
- Wills, J. B., Brewster, Z. W., & Fulkerson, G. M. (2005). The Stratification Puzzle: An Active-Learning Exercise in Hard Work and Success. *Teaching Sociology*, 33(4), 389-395.
- Wright, W., & Ransom, E. (2005). Stratification on the Menu: Using Restaurant Menus to Examine Social Class. *Teaching Sociology*, 33(3), 310-316.

Appendix A: Holistic Argument Scoring Rubric

The following rubric from (Jonassen & Cho, 2011) was used to determine how to program SA Grader to assess the quality of argument in student essays in this study:

Level	Description
0. Response to topic	Essay does not provide a clear claim.
1. Undeveloped opinion	Essay states a clear claim but no reason is given to support the claim or the reason given is unrelated to the claim.
2. Minimally developed	Essay provides a clear claim and reasons supporting the claim, but the reasons are not well explained or elaborated.
3. Partially developed	Essay provides a clear claim and substantial reasons that are well explained and elaborated, but no counterclaim is addressed.
4. Well developed	Essay provides a clear claim and counterclaim and rebuts the counterclaim, but some reasons supporting the claim or rebutting the counterclaim are not well explained or elaborated.
5. Fully developed	Essay provides a clear claim and a plausible counterclaim, and they are supported by substantial reasons that are well explained and elaborated. The essay effectively rebuts the counterclaim with substantial reasons and/or proposes a valid alternative solution that addresses counterclaim concerns.

Appendix B: Assignment Instructions

Copyright for all assignment text in this appendix is ©2009 by IdeaWorks, Inc.

Case 1 Instructions

The following are the instructions and specific assignment questions for the first case for each of the three treatment groups. Case 1A represents the instructions for the construct treatment group, Case 1C for the counter-argue treatment group, and Case 1D for the summarize treatment group.

Case 1A Argumentation: Make a Choice and Argue for Your Case

Description

In this exercise you are to learn about Mrs. Chou who is looking for a renter for the upstairs apartment in her house, and two potential renters, Ashley and Delon. You will first need to research the case, looking at information about each renter and their interview. Then you will be asked to examine the conversation between Mrs. Chou and her daughter. Based on this information your paper should identify various argumentation strategies used by Mrs. Chou and her daughter. Then you should indicate to whom you would rent, provide reasons that support that choice, and provide reasons against renting to the other person.

Questions

1. Identify at least two fallacies made by Mrs. Chou or her daughter in their arguments that challenge your choice, including the comment in which the fallacy occurred and the type of fallacy.
2. Identify your final claim. To whom would you rent?
3. Provide at least two reasons that support your choice, and indicate whether each is evidence, opinion, or something else.

4. Provide at least two reasons that challenge the other choice, and indicate whether each is evidence, opinion, or something else.
5. Identify at least one rebuttal made by Mrs. Chou or her daughter that supports your argument, and identify what it rebuts.

Case 1C Argumentation: Counter an Argument

Description

In this exercise you are to learn about Mrs. Chou who is looking for a renter for the upstairs apartment in her house, and two potential renters, Ashley and Delon. You will first need to research the case, looking at information about each renter and their interview. Then you will be asked to examine the conversation between Mrs. Chou and her daughter. Finally, you will read a paper arguing Mrs. Chou should rent to Ashley. Based on this information your paper should identify various argumentation strategies used by Mrs. Chou and her daughter. Then you should make an argument that supports renting to Delon by providing reasons in favor of him and reasons against renting to Ashley.

Questions

1. Identify at least two fallacies made by Mrs. Chou or her daughter in their arguments that challenge your choice of Delon, including the comment in which the fallacy occurred and the type of fallacy.
2. Provide at least two reasons supporting your choice to rent to Delon and indicate what type of support they provide.
3. Provide at least two reasons against renting to Ashley and indicate whether they are evidence or opinion.
4. Identify at least one rebuttal made by Mrs. Chou or her daughter that supports renting to Delon, and identify what it rebuts.

Case 1D Argumentation: Discuss the pro's and con's of each potential renter

Description

In this exercise you are to learn about Mrs. Chou who is looking for a renter for the upstairs apartment in her house, and two potential renters, Ashley and Delon. You will first need to research the case, looking at information about each renter and their interview. Then you will be asked to examine the conversation between Mrs. Chou and her daughter. Based on this information your paper should identify various argumentation strategies used by Mrs. Chou and her daughter, and discuss the reasons supporting and challenging the choice of each of the potential renters.

Questions

1. Identify at least three kinds of fallacies in reasoning in Mrs. Chou's or her daughter's comments and cite the comment that illustrates each of those fallacies.
2. Argumentation Case 1 - Identify the final claim made by Mrs. Chou. (To whom will she rent?)
3. Provide at least two supporting arguments for each potential renter and two arguments against each, and indicate whether each argument is evidence, opinion, or something else.
4. Argumentation Case 1 - Both Mrs. Chou and her daughter issued at least one rebuttal in their conversation. Name at least one of those rebuttals and identify what it rebutted.

Case 2 Instructions

The following are the instructions and specific assignment questions for the first case for each of the three treatment groups. Case 2A represents the instructions for the construct treatment group, Case 2C for the counter-argue treatment group, and Case 2D for the summarize treatment group.

Case 2A: Choose and argue

Description

You are the CEO of a company hiring a national sales director. You are meeting with your Vice Presidents to discuss the three finalists for the job. Because the VPs do not agree on which candidate to hire, you must make the final decision. In order to explain to the VPs why you have selected the candidate you choose, use what you know from sociology to help you make a fair decision and to justify your choice to the VPs. Be sure to use a number of relevant sociological concepts and facts. You should pay attention not only to why you believe a particular candidate should get the job, but also to ways in which the people in this case reflect common sociological concepts, recurring themes, and sociological theories.

Hints: You should consider the different letters and resumes as well as the arguments made by the VPs about the candidates and consider how they reflect concepts or issues relevant to sociology.

Questions

1. Identify who you would hire and point out at least two reasons supporting your choice.
2. For the candidate you would hire, identify at least one argument against them that is based primarily on opinion with little or no factual evidence.
3. For each of the candidates you would not hire identify at least one argument against them that is based on evidence and not just opinion.
4. Acknowledge at least one point supporting each of the other two candidates. Be sure to indicate the candidate supported by the point.
5. Identify at least one argument against a candidate and a rebuttal expressed in the discussion of the vice presidents.

6. Select one of the vice presidents and identify statements by him or her that illustrate at least three different kinds of critical thinking. Be sure to specify the kind of critical thinking illustrated by each statement.
7. Identify at least one fallacy for each vice president.
8. Who had the fewest fallacies in their arguments?
9. Did anyone make an assumption that was challenged by someone else? If so, who made the assumption, what was the assumption, and who challenged the assumption?
10. Which speakers show some evidence of being reflexive?

Case 2C: Counter an argument

Directions

You are the CEO of a company hiring a national sales director. You are meeting with your Vice Presidents to discuss the three finalists for the job. Because the VPs do not agree on which candidate to hire, you must make the final decision. After you hear the VPs discussing the candidates you read the argument in favor of hiring John Meyer. Your task now is to make a counterargument, arguing in favor of either of the other two candidates and against hiring John Meyer. In order to explain to the VPs why you have selected the candidate you choose, use what you know from sociology to help you make a fair decision and to justify your choice to the VPs. Be sure to use a number of relevant sociological concepts and facts. You should pay attention not only to why you believe a particular candidate should get the job, but also to ways in which the people in this case reflect common sociological concepts, recurring themes, and sociological theories.

Hints: You should consider the different letters and resumes as well as the arguments made by the VPs about the candidates and consider how they reflect concepts or issues relevant to sociology. Don't forget your task is to make a counterargument, arguing in favor of either of the other two candidates and against hiring John Meyer.

Questions

1. Identify which of the other two candidates (Otis or Jessica) you would hire and point out at least two reasons supporting your choice.
2. For the candidate you would hire, identify at least one argument against them that is based primarily on opinion with little or no factual evidence.
3. For each of the candidates you would not hire identify at least one argument against them that is based on evidence and not just opinion.
4. Acknowledge at least one point supporting each of the other two candidates. Be sure to indicate the candidate supported by the point.
5. Identify at least one argument against a candidate and a rebuttal expressed in the discussion of the vice presidents.
6. Select one of the vice presidents and identify statements by him or her that illustrate at least three different kinds of critical thinking. Be sure to specify the kind of critical thinking illustrated by each statement.
7. Identify at least one fallacy for each vice president.
8. Who had the fewest fallacies in their arguments?
9. Did anyone make an assumption that was challenged by someone else? If so, who made the assumption, what was the assumption, and who challenged the assumption?
10. Which speakers show some evidence of being reflexive?

Case 2D: Discuss

Description

You are the CEO of a company hiring a national sales director. You are meeting with your Vice Presidents to discuss the three finalists for the job. Because the VPs do not agree on which candidate to hire, you must make the final decision. Before making that decision, however, your task here is to discuss the relative merits and demerits of each candidate. Use

what you know from sociology to help you highlight reasons supporting and challenging each candidate. Be sure to use a number of relevant sociological concepts and facts. You should pay attention not only to why a particular candidate should or should not get the job, but also to ways in which the people in this case reflect common sociological concepts, recurring themes, and sociological theories.

Hints: You should consider the different letters and resumes as well as the arguments made by the VPs about the candidates and consider how they reflect concepts or issues relevant to sociology.

Questions

1. For each candidate, point out at least two reasons supporting that candidate.
2. For each of the candidates, identify at least one argument against them that is based primarily on opinion with little or no factual evidence.
3. For each of the candidates, identify at least one argument against them that is based on evidence and not just opinion.
4. Acknowledge at least one point supporting each of the other two candidates. Be sure to indicate the candidate supported by the point.
5. Identify at least one argument against a candidate and a rebuttal of that argument expressed in the discussion of the vice presidents.
6. Select one of the vice presidents and identify statements by him or her that illustrate at least three different kinds of critical thinking. Be sure to specify the kind of critical thinking illustrated by each statement.
7. Identify at least one fallacy for each vice president.
8. Who had the fewest fallacies in their arguments?
9. Did anyone make an assumption that was challenged by someone else? If so, who made the assumption, what was the assumption, and who challenged the assumption?

10. Which speakers show some evidence of being reflexive?

Case 3 Instructions

The instructions for all three treatment groups for the third case were the same. In this case, unlike the previous cases, the participants' first submissions were counted for a portion of their grade in order to determine any effect of the treatments on their performance prior to revision.

Description

You are the head of the admissions committee. You are a new member and this is your chance to make a good impression on the President of the University, who is sitting on the discussion. First, listen to what the other members have to say and then make your own speech. In this speech you should support one of the candidates and explain why. Part of your explanation should anticipate additional arguments the other members may make when they respond. You should also rebut points they have already made that are faulty. You may want to do things like identify any logical fallacies or points that are not supported by the data or are only supported by opinion or pseudo-evidence. Another way you can strengthen your argument is to place it in context of broad social issues, important events or problems facing the university.

Since you are newly appointed as head of the committee, President Smith also wants you to discuss any problems you see in how the committee functions or generic problems such as those that typically plague bureaucracies and large organizations where communication is always an issue.

FINALLY, THERE IS A VERY IMPORTANT DIFFERENCE BETWEEN THIS ASSIGNMENT AND ALL OF THE EARLIER SAGRADER ASSIGNMENTS. THE PRESIDENT WANTS TO SEE HOW WELL YOU CAN WORK ON YOUR OWN WITHOUT HELP SO IN THIS ASSIGNMENT YOUR FIRST RESPONSE WILL COUNT FOR 25% OF

YOUR GRADE AND YOUR LAST RESPONSE WILL COUNT FOR THE OTHER 75%. SO
BE SURE TO DO THE BEST JOB YOU CAN IN YOUR FIRST SUBMISSION.

Appendix C: Component Items

Item Name	Item Description
SeekEvidence	An item related to seeks evidence
Theme	Two items related to identifies question or theme
AvoidFallacies	An item related to avoids logical fallacies
JeffHighRep	An item that is an example of Jefferson High School has reputation for tough grades
Interpret	An item related to interpret
MethodBiases	An item related to biases related to method
CommitteeAuthorityLimits	An item that includes as an example “I don’t really think this committee has the authority to make that decision. That is an issue that the Head of Admissions should resolve.”
Applicant	three applicants
RespectOtherPerspectives	Two items related to respects other perspectives
AdHominemArgument	Two statements that are ad hominem argument
PeterPrinciple	An item that has the negative consequence Peter Principle
RelevantEvidence	An item related to identifies relevant evidence
ImplicitIssues	An item related to identifies implicit issues
GarciaDishonest	An item that is an example of Jennifer Garcia is dishonest
AvoidNegativeConsequences	An item that is network communication structure
NetworkComm	An item that has the negative consequence decision avoidance

AckOtherPerspectives	An item related to acknowledge other perspectives
Fallacy	Three examples of fallacy
MustMakeDecision	An item that includes as an example, “But we have to make this decision today. The guidelines in the faculty bylaws are clear.”
JustifiesOrDefendsPosition	Two items related to justifies or defends their position (argues for it)
NetworkComm2	Network communication structure
ApplicantA	three applicants
Typology	Two items related to typology
NotOurProblem	An item that includes as an example, “That’s not our problem. If they can’t give us dependable information, how do they expect us to make such important decisions?”
ConsidersMethod	An item related to considers method
PutsInContext	An item related to puts in context
IfThenReasoning	An item related to if-then reasoning
SupportDiaz	Four reasons that support Ramona Diaz
FeatureOfChain	An item that is a feature of chain (communication structure)
SubIssues	An item related to identifies sub-issues
FeatureOfStar	An item that is a feature of star (communication structure)
HastyGeneralization	A reason that argues against Matt Schmidt
OpposeSchmidt	A statement that is a hasty generalization
AckOwnPosition	Two items related to acknowledges own position

ExamineAssumptions	Two items related to examines assumptions
SupportWashington	Two reasons that support Scott Washington
ComparePositionToOtherViews	Four items related to compare their position to other views
JudgeEvidenceBasedOnMethod	Three items related to judges evidence based on method
RaiseValuedConsequenceIssues	An item related to raises valued consequences issues
FalseDichotomy	A statement that is a false dichotomy
SupportSchmidt	Three reasons that support Matt Schmidt
ExplicateImplications	Three items related to explicates logical implications
ClearClaim	Three items related to clearly states their position
NotGoingToMakeItWithProblem	An item that is an example of not going to make it with problems reading and writing
AvoidRelativism	An item related to avoids relativism
Generalization	An item related to generalization
AvoidBasicDualism	An item related to avoids basic dualism
AppealToAuthority	A statement that is an appeal to authority
RaiseEthicalIssues	An item related to raises ethical issues
ExampleOfLearningDisabilities	An item that is an example of “students with learning disabilities almost never make it here”
OpposeWashington	A reason that argues against Scott Washington
CorrelationNotCausation	An item related to does not confuse correlation with causation
HardWorker	An item that is an example of hard worker
ConclusionsQualifiedByEvidenceAndContext	An item related to conclusions qualified by evidence and context

MultipleCausality	An item related to considers multiple causality
OpenToNewEvidence	An item related to open to new evidence
AboutWashington	An item that can be Scott Washington
Reasons	Another reason
ExampleOfGoodRoleModel	An item that is an example of good role model
FeatureOfComCon	an item that is a feature of comcon (communication structure)
Statements	Three statements
CloneOfFalseDichotomy	A statement that is ad populum
AdPopulum	A statement that is a Clone of false dichotomy
AnticipatesOthersArguments	An item related to anticipates arguments of other perspectives
PromotedExperience	An item that includes as an example “promoted experience and was competent at that job. But no one wants to demote him since that would be so embarrassing.”
EverSincePromoted	An item that includes as an example “Ever since he was promoted to head admissions a few years ago, things have been a mess. We should have kept him as the head interviewer instead”
AboutDiaz	An item that can be Ramona Diaz
OpposeDiaz	A reason that argues against Ramona Diaz
AboutSchmidt	An item that can be Matt Schmidt

GoalDisplacement	An item that has the negative consequence goal displacement
HumanSubjectIssues	An item related to considers human subjects issues

VITA

Originally from Cincinnati, Ohio, Holly R. Henry first studied psychology as an undergraduate at the former Mount Union College (now the University of Mount Union) in Alliance, Ohio. After leaving Mount Union, Dr. Henry worked for several years in the information technology field during which time she also completed a B.S. in Information Technology at the University of Cincinnati.

During her work in this field, Dr. Henry often designed and delivered training to end-users and fellow technology professionals. Her desire to improve her teaching skills provided the impetus for her further study in education. She studied Instructional Technology and Media at Teachers College, Columbia University where she received an MA in 2005.

While she continued her work in the information technology field in private industry, after completing the MA, Dr. Henry also consulted with the School of Continuing and Professional Studies at New York University to provide faculty development for online instructors and served as an adjunct lecturer in the school's MS program in Instructional Design.

Dr. Henry decided to pursue her Ph.D. in the School of Information Science and Learning Technologies in the College of Education at the University of Missouri starting in Fall semester, 2007 to further deepen her understanding of teaching and learning. Her research interests include supporting novice-to-expert learning, problem-based learning, STEM education, distance education, workplace training, vocational education, adult education, and the design of learning systems, environments, and activities.