UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Improving Science Writing in Research Methods Classes Through Computerized Argument Diagramming

Permalink
https://escholarship.org/uc/item/1b28v9cz

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors
Barstow, Brendan J
Schunn, Christian D
Fazio, Lisa K
et al.

Publication Date
2015

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Improving Science Writing in Research Methods Classes Through Computerized Argument Diagramming

Brendan J. Barstow (Brendan.Barstow@Pitt.edu)
Learning Research and Development Center, 3939 O'Hara Street
Pittsburgh, PA 15260 USA

Christian D. Schunn (Schunn@Pitt.edu)
Learning Research and Development Center, 3939 O'Hara Street
Pittsburgh, PA 15260 USA

Lisa K. Fazio (Lisa.Fazio@Vanderbilt.edu)
Department of Psychology & Human Development, Vanderbilt University, 230 Appleton Place #552, Jesop 105
Nashville, TN 37203 USA

Mohammad H. Falakmasir (Falakmasir@cs.pitt.edu)
Learning Research and Development Center, 3939 O'Hara Street
Pittsburgh, PA 15260 USA

Kevin Ashley (Ashley@Pitt.edu)
Learning Research and Development Center, 3939 O'Hara Street
Pittsburgh, PA 15260 USA

Abstract

The purpose of this study was to characterize the ways in which psychologists address research hypothesis risk in academic articles, and to support undergraduates in learning to write about such risk using argument diagramming pre-writing activities. First, 90 articles recently published in top social, developmental, and cognitive psychology journals were examined for their presentation of research hypothesis 'risk' – an element of the intellectual merit of a research study denoting the novelty and importance of the study being conducted. Second, an experimental study was conducted involving 82 students in undergraduate research methods classes. They were assigned to either argument diagram or traditional instruction conditions. Research reports were coded for explicit discussion of risk. Students using argument diagramming were significantly more likely to write about risk when compared to matched classes given no diagramming support.

Keywords: Argument diagram; writing instruction; science instruction; educational intervention; hypothesis risk; philosophy of science

Introduction

American students have a writing problem. Only about one-quarter of 8th and 12th graders are able to write at or above a *proficient* level (National Center for Education Statistics, 2011), with little to no improvement since 1998 (Greenwald, Persky, Campbell, & Mazzeo, 1998; Persky, Daane, & Jin, 2002; Salahu-Din, Persky, & Miller, 2007). This is perhaps unsurprising, as according to a sample of high school English teachers, students may have only one or two opportunities each semester to write longer and/or evidence-based essays (Kiuhara, Graham, & Hawken, 2009). The instruction of writing is especially time-intensive, typically involving multiple drafts requiring careful review and high quality feedback in order to produce student improvements. In lower and mid-tiered colleges, which train the vast majority of our students, writing skill continues to stagnate with negligible improvement over four years – a trend employers are noticing and lamenting (Arum & Roksa, 2011).

Argumentation and argumentative writing are in particular need of attention given that instructional practices tend to emphasize the *presentation* of arguments rather than their *generation*, and the *product* of writing rather than the *process* (Andrews 1995; Andrews & Mitchell, 2001; Chryssafidou & Sharples, 2003; Oostdam, de Glopper, & Eiting, 1994; Oostdam & Emelot, 1991). There is a clear demand, then, for instructional tools that can directly improve argumentation and argumentative writing.

Argument diagramming has a long history of use in philosophy for organizing formal logic (Whately, 1834), but the limitations of pen-and-paper diagrams (e.g. difficulty of revisions, time required) make their use in the classroom less practical. Advances in computing over the last few decades have enabled the development of computerized argument diagramming software, and a growing body of research supports its effectiveness in classrooms (Kozma, 1991; Chryssafidou, 2000; Twardy, 2004; Proske, Narciss, & McNamara, 2010).

Argument diagramming has been shown to facilitate better argumentation for a variety of component elements. Training in the construction of argument diagrams and their subsequent use has been tied to improved critical thinking ability (Harrell, 2011; 2012; Twardy, 2004), a skill important both for the analysis and generation of arguments. Diagrams can also aid students in argument-counterargument integration, or the development of an informed conclusion (Nussbaum & Schraw, 2007). We see these benefits synthesized in a study by Harrell (2013), where students trained in diagramming produced more elaborate, coherent arguments following the intervention than students taught traditional argument analysis methods.

The link between the quality of students' diagrams and their subsequent writing has been established through recent modeling work (Lynch, Ashley, & Chi, 2014; Lynch, 2014), suggesting that the complexity and coherence of a diagram can be used to predict the grade earned by the resulting essay. Despite these promising results, the impact of computerized argument diagramming has yet to be tested in the domain of science writing.

Writing in science poses unique challenges for both the student and the instructor. For the student, the construction of a literature review often requires the synthesis of a large number of sources, or claims. This synthesis requires not only breadth but depth as well, for the author must be able to extract the core finding of each study cited, evaluate the validity or scientific strength of these findings, and then compare the relevance of each study to the others and to the current hypothesis being explored.

For the instructor, the breadth and depth required in science writing makes its review time-consuming and challenging. The writer often knows the background research much better than the instructor, making it difficult for instructors to level informed criticisms. This difficulty is compounded in the case of student peer review – one approach to improving writing instruction - where student peers may not have an instructor's general field knowledge to fall back on in the absence of specific content familiarity.

One approach to help mitigate these issues, spearheaded by Hand and Keys, seeks to provide students with more opportunities for informal writing in science, with an emphasis on writing to learn (Keys, Hand, Prain, & Collins, 1999). Their instructional template, the Science Writing Heuristic, provides guides for instructors and students to engage in meaningful writing, reflection, and discussion about concepts and experiences in science. Evidence suggests that these informal writing experiences help students create a richer representation of scientific understanding, and enable them to respond more deeply to related test questions (Keys, Hand, Prain, & Collins, 1999; Hand, Prain, & Wallace, 2002; Hand, Wallace, & Yang, 2004). While their work focused on informal writing for understanding, the current study sought instead to develop and test an intervention for students' *formal* writing in science, with an emphasis on writing rather than learning outcomes.

For both students and experts in science, the generation of a suitable hypothesis poses an additional challenge in that the purpose of a scientific introduction is to pose a problem to be explored, unlike in other forms of argumentation where the goal may be a conclusive statement. The introduction should serve to justify the hypothesis(es) being explored, and should present an element of appropriate risk.

The eminent philosopher of science Karl Popper asserted the following regarding this idea of risk: "Confirmations should count only if they are the result of risky predictions; that is to say, if unenlightened by the theory in question, we should have expected an event which was incompatible with the theory – an event which would have refuted the theory" (1963).

We see hypothesis risk as both an essential element in the scientific method and one that should be communicated explicitly in scientific writing. Risk is defined here as the strength of belief that a given hypothesis will not be supported based on existing research. We consider this idea of risk to be one component of the intellectual merit of a given study, which, when combined with its practical merit, forms the whole of a convincing scientific rationale. Too much risk would indicate an insufficient basis for the proposed hypothesis in the literature and might be a poor use of resources, while not enough risk would indicate that the proposed hypothesis is not scientifically novel - that the question has already been sufficiently explored. A proper hypothesis represents a balance between these two extremes.

This specific framing of scientific writing is novel, and therefore the first stage of our work was to verify that research writing in psychology does explicitly frame hypothesis risk. For instance, it is possible that psychologists as writers, reviewers, and readers understand the implicit risks in a given research area and do not need to explicitly present the risks in their papers. By studying the way that risk is presented in published articles, we learned how commonly this component of science writing is included and established a standard for what students should aim to achieve. Additionally, by developing a system to categorize risk in published articles we were able to use this system to understand the nature of risk in student writing and the effectiveness of an argument diagramming intervention.

Following this analysis, we then developed and tested the effectiveness of a computerized argument diagramming intervention for improving undergraduates' APA-style paper introductions. The diagramming ontology used in the experiment prompted students to describe the relevance and validity of cited studies (one logical basis of hypothesis risk) and note the relationship of cited studies to their hypotheses as either supporting or opposing (another logical basis for hypothesis risk). We expect that these two components will enable students to write more about risk through uncertainty (a hypothesis with novel contents) and risk through opposition (a hypothesis with mixed prior support), respectively.

We hypothesized that undergraduate university students in psychological research methods classes undergoing an argument diagramming intervention would produce higher quality first draft introductions as measured by the degree to which they explicitly addressed hypothesis risk, compared to a control group of students undergoing no intervention.

Methods

Review of Risk in Published Articles

Selection

We selected nine journals spread across three major psychological disciplines – cognitive, social, and developmental. In each discipline, we selected the top three empirical research journals based on impact factor rankings. From each of these nine journals, we then examined the ten most recently published original empirical articles as of December 2014 for a total of 90 articles. We focused only on empirical articles (no short reports, reviews, or corrections), as they are the closest to what the student participants in our diagramming intervention would be writing, and empirical articles would be expected to include the element of hypothesis risk that interested us. The set of journals examined included the following: *Journal of Experimental Psychology: Learning, Memory, and Cognition*; *Cognitive Psychology*; *Cognition*; *Developmental Science*; *Developmental Psychology*; *Child Development*; *Journal of Personality and Social Psychology*; *Journal of Experimental Social Psychology*; and *Personality and Social Psychology Bulletin*.

Procedures

We iteratively developed a coding scheme to use in systematically categorizing the risk-related language found in psychology research. This process resulted in three ways authors typically address risk in psychology.

Risk through uncertainty (RU) is addressed when the author claims that there is insufficient or problematic evidence for his/her hypothesis(es) in the literature. E.g., "First, although the effect of fluency on a variety of judgments has been well documented, it is unknown whether fluency can influence two different attributes at once" (Westerman, Lanska, & Olds, 2014).

Risk through opposition (RO) is addressed when the author claims that there is both supporting and opposing evidence for his/her hypothesis(es) in the literature. E.g., "While there is strong evidence for such a process of combination, there has been some debate as to when metric and categorical cues are combined…" (Holden, Newcombe, & Shipley, 2014).

Risk through difficulty (RD) is addressed when the author claims that his/her hypothesis(es) is particularly difficult to test. E.g., "However, formally specifying a prior distribution that captures the expectations of humans (or even just a select group of a more homogenous population) is particularly difficult, and will be the main concern of the present paper" (Yeung & Griffiths, 2014).

Using these definitions, we coded the introductions of each article by annotating individual sentences that addressed risk, and coding each article based on the types of risk addressed, regardless of the number of instances above one (1). For example, if an article had four sentences tagged as RU and one as RO, that article would be tagged as [RU, RO]. In other words, one instance of a risk type was sufficient to be considered. Agreement between two expert coders was calculated in a 3x3 matrix (RU, RO, Combination) on a 30 article, three journal subset of the data (JEP: LMC, PSPB, & Dev. Psych.), *kappa* = .93. RD was omitted from this calculation given its low occurrence.

Diagramming Intervention

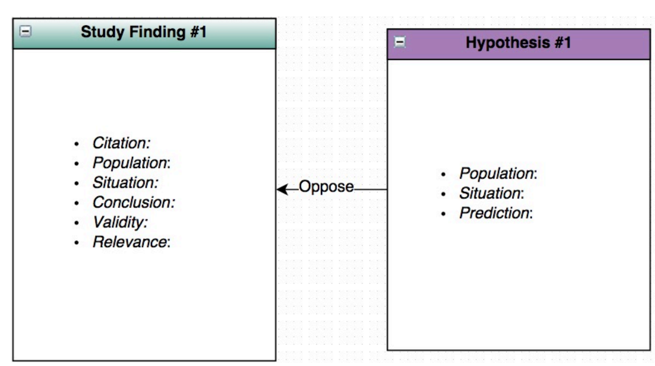
Participants

82 participants were enrolled in Research Methods classes at a large public university, including 2.5 lecture hours and 3 lab hours per week. The intervention was limited to the lab sections only. Two pairs of classes were matched for day of week, time of day, and instructor experience. These four lab sections (2 experimental, 2 control) were taught by three instructors – one of the instructors taught both an experimental and a control section.

Materials

Students created argument diagrams using a free, web-based, open-ended environment called Draw.IO. Students were instructed to construct diagrams containing two hypotheses, a rationale for each hypothesis (why should your study be done?), supporting and opposing evidence, and counterarguments for any opposing evidence. Students were provided with a framework for these instructions in the form of a diagram template (Figure 1). In the absence of directly opposing evidence, students were asked to demonstrate appropriate risk in other ways (i.e., through the absence of data, methodologically weak previous studies). For each piece of evidence cited, students were asked to include the APA-style citation, the population tested, the situation (what were they testing?), the conclusion (what did they find?), the validity (e.g. correlational), and the relevance of the evidence to the student's hypothesis(es). Relevance was classified as slightly, partially, or highly relevant, and students were encouraged to explain their classification by comparing the variables and context between the cited study and their hypothesis.

Figure 1: Partial example of argument diagram template



Procedures

In the experimental lab sections, students were first given a brief presentation explaining the idea of hypothesis risk. This included the issues associated with insufficient risk (i.e., study is redundant) and excessive risk (i.e., study is a poor use of resources), and conveyed the importance of relevant and valid research in locating strong support for one's hypothesis. Afterward, these students completed an activity where they viewed three argument diagrams and were tasked with choosing which one was too risky (insufficient supporting evidence), which was too 'safe' (only supporting evidence), and which demonstrated an appropriate level of risk (some supporting and some opposing evidence).

In a later class, students were introduced to the diagramming software through a practice activity where they worked in pairs to diagram a short scientific paper given to them by their instructor. Finally, students were instructed to construct an argument diagram for an observational experiment they would later conduct, and to use the diagram when writing the associated paper.

The paper assignment for both conditions included an abstract, introduction, methods, results, and discussion. Students were asked to include at least five peer-reviewed references in their introduction (two of which were provided by the instructor) and two hypotheses, and to discuss and explain at least one study or theoretical position that conflicted with their hypotheses.

Using the coding scheme developed for the review stage detailed previously, all 82 introductions of students' first draft papers were coded for risk (RU, RO, RD) in an identical process.

Results

Review of Risk in Published Articles

93% of the articles examined addressed at least one (1) type of risk, while 21% of the articles addressed at least two (2) types of risk. The majority of articles that did not discuss risk were concentrated in the social psychology journals (83%), with the only other no-risk article in a cognitive psychology journal. Aside from this, distributions of the types of risk addressed were relatively consistent across both disciplines and journals. Only 2 (<3%) of the 90 articles demonstrated risk through difficulty.

Table 1: Risk distribution by discipline

	Social	Dev.	Cog.
RU	25	28	26
RO	3	10	11
RD	0	1	1
No Risk	5	0	1
Total Articles	30	30	30
>1 types of risk	25	30	29
>2 types of risk	3	9	7

Diagramming Intervention

In the control sections, 71.8% of students addressed risk in at least one form compared to 93.0% of students in the diagramming condition. For all of the following analyses, we used α=.05. A χ^2 test of independence applied to all four lab sections revealed that students in the diagramming condition wrote about risk significantly more than those in the control condition, $\chi^2(1, n=82)=6.5$, $p=.02$. This difference represents a moderate effect (d=0.58). A χ^2 analysis looking at RU versus no RU and RO versus no RO across all four lab sections indicated more writing about RU in the diagramming condition, $\chi^2(1, n=82) = 11.7$, $p<.001$, a large effect (d=0.81), and combinations of risk types $\chi^2(1, n=82)=4.7$, $p<.001$, a moderate effect (d=0.49) but no difference in writing about RO, $\chi^2(1, n=82)=0.5$, $p=.46$. Most of these results held when examined for only the two within-instructor sections, showing more writing about any risk, χ^2 (1, n=42)=8.8, $p<.001$, (d=1.02) and more writing about RU, χ^2 (1, n=42)=4.8, $p=.03$, (d=0.71) in the diagramming condition, but no significant difference across conditions in writing about RO, $\chi^2(1, n=42)=0.8$, $p=.37$ or combinations of risk types $\chi^2(1, n=42)=0.5$, $p=.49$, (d=0.21).

Figure 2: Proportion of student papers addressing risk in any form with SE bars.

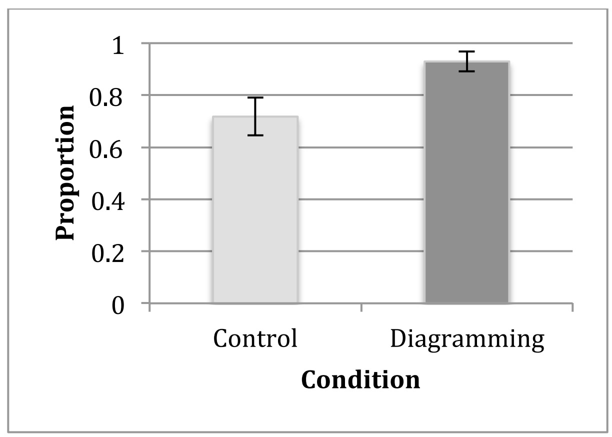
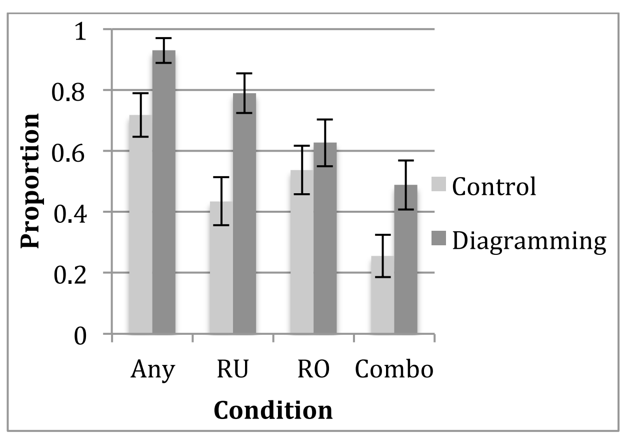


Figure 3: Proportion of student papers addressing any risk (Any), risk through uncertainty (RU), risk through opposition (RO), and a combination of risk types (Combo)



Discussion

The results of our review of science writing in the field matched our expectations – that hypothesis risk is a common element of scientific argumentation and that it is explicitly addressed in the vast majority of recently published papers in psychology. It is often addressed multiple times in one article, and occasionally authors will even address risk through multiple forms. Although we examined only a small sample of articles (n=90), it is interesting that 5 of the 6 articles that did *not* address risk were published in social psychology journals. It is possible that there are subdiscipline-level differences in this aspect of science and science writing, like those found by Okada and Shimokido (2001) indicating a major difference in writing style between papers in physical science journals as compared to psychology journals. Their investigation found that psychology papers were much more likely to follow a hypothesis-testing style of writing and that most physical science introductions included no hypotheses at all. If similar discipline or subdiscipline level differences exist for the presentation of hypothesis risk, they may be uncovered from the examination of a larger dataset.

The distribution of risk types appeared to be rather consistent across journals and disciplines except with respect to no-risk articles. Authors most commonly demonstrated hypothesis risk through uncertainty, secondly through opposition, and only two articles addressed risk through difficulty. Although these cases of risk through difficulty were rare in our examination, we do not believe that this invalidates risk through difficulty as an acceptable method for demonstrating hypothesis risk given its independence from the other categories, but instead that its rate of use may be very low. It might be *difficult* to address risk through difficulty, and this approach may only be appropriate for certain types of studies which deal with issues like obscure populations, complex modeling, etc.

The results of our argument diagramming intervention supported our hypothesis, indicating that it is at least moderately effective for improving this particular element of science writing.

The mechanism behind this improvement, however, is less clear. The diagram template given to students (Figure 1) encouraged the inclusion of opposing findings, and so we anticipated that this would encourage students to find opposing research in the first place, and that having it in their diagram would make it more accessible when it came time to write their paper. Additionally, previous research (Nussbaum & Schraw, 2007) indicates argument diagramming can improve argument-counterargument integration for students, and one could argue that the parallel of this in our study would be an increase in the discussion of risk through opposition. However, a chi-square analysis focusing only on risk through opposition revealed no apparent differences between the experimental and control conditions on this type alone at both the aggregate level (all four classes) and the within-instructor level (two classes). We believe this may be due to the nature of scientific literature reviews, in that it is often difficult to find directly opposing evidence to a given hypothesis and that it is easier to find gaps in scientific knowledge to explore.

Nussbaum and Schraw (2007) controlled for this by providing students with a sample text including arguments for two sides of the argument prompt. The diagramming activity may have prompted students to consider opposing evidence more seriously, but if they could not find any opposing evidence to use then this additional consideration would not be reflected in their written introduction. Interestingly, a similar result was found by Toth, Suthers, and Lesgold (2002), where students *considered* more opposing evidence when constructing argument diagrams, but this consideration had no effect on their final prose.

Instead, the bulk of the difference between the two conditions in our study is captured by risk through uncertainty - the mechanism behind which may be subtler. In their diagrams, students were asked to describe the relevance and validity of studies, and it is possible that this task enabled students to see gaps or issues in the existing research literature more clearly. We plan to shed light on this possibility by examining the ways in which the argument diagramming task impacted students' writing of relevance and validity..

These results add to a growing body of research supporting the effectiveness of computerized argument diagramming for improving students' argumentative writing across diverse academic disciplines (Chryssafidou, 2000; Proske, Narciss, & McNamara, 2012). The present study provides evidence that at least some of the positive outcomes associated with diagramming in philosophy education (Harrell, 2011; 2012; 2013) may be found for science education as well.

This study focused on one small but important element of science writing. Additionally, students' baseline conceptual knowledge about psychology and science writing was not measured, nor any changes thereafter, and therefore we are unable to determine the interactive role that conceptual knowledge may have played in the effect of the diagramming intervention. Finally, it is possible that some of the differences seen across the two conditions could be explained by the difference in time spent on the assignment, although this difference is only slight. Further research is needed to determine how the use of argument diagrams may affect other components of science writing and learning.

Acknowledgments

This research was supported by the National Science Foundation Grant IIS-1122504. Any opinions expressed in this work reflect those of the authors and do not necessarily reflect those of the National Science Foundation or the University of Pittsburgh.

We would like to thank Diane Litman, Huy Nyugen, and Caitlin Rice for their feedback regarding this project.

References

Andrews, R. (1995). Teaching and learning argument. London: Cassell.

Andrews, R., & Mitchell, S. (2001). *Essays in argument*. Middlesex University Press.

Arum, R. & Roska, J. (2011). *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press.

Chryssafidou, E. (2000). DIALECTIC: Enhancing essay writing skills with computer-supported formulation of argumentation. In C. Stephanidis (Ed.), *Proceedings of the ERCIMWG UI4ALL one-day joint workshop with i3 Spring Days 2000 on "Interactive Learning Environments for Children"*, (pp. 14 pages). Athens, Greece.

Chryssafidou, E., & Sharples, M. (2002). Computer-Supported Planning of Essay Argument Structure. *Proceedings of the 5th International Conference of Argumentation*. Amsterdam: Sic Sat.

Greenwald, E. A., Persky, H. R., Campbell, R. R., & Mazzeo, J. (1998). *The NAEP 1998 Writing Report Card for the Nation and the States,* (NCES 1999-462). National Center for Education Statistics, Office of Educational Research and Improvement. Washington, DC: 1999.

Hand, B., Prain, V., & Wallace, C. (2002). Influences of Writing Tasks on Students' Answers to Recall and Higher-Level Test Questions. *Research in Science Education, 32*, 19-34.

Hand, B., Wallace, C. W., & Yang, E. (2004). Using a Science Writing Heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: quantitative and qualitative aspects. *International Journal of Science Education, 26(2)*, 131-149.

Harrell, M. (2011). Argument Diagramming and Critical Thinking in Introductory Philosophy. *Higher Education Research & Development, 30(3)*, 371-385.

Harrell, M. (2012). Assessing the Efficacy of Argument Diagramming to Teach Critical Thinking Skills in Introduction to Philosophy. *Inquiry*, *27(2)*, 31-38.

Harrell, M. (2013). Improving First-Year Writing Using Argument Diagramming. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*.

Keys, C. W., Hand, B., Prain, V., & Collins, S. (1999). Using the Science Writing Heuristic as a Tool for Learning from Laboratory Investigations in Secondary Science. *Journal of Research in Science Teaching, 36(10)*, 1065-1084.

Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, *101*, 136-160.

Kozma, R. (1991). Learning with media. *Review of Educational Research*, 61(2), 179-212.

Loll, F. & Pinkwart, N. (2013). LASAD: Flexible Representations for Computer-Based Collaborative Argumentation. *International Journal of Human-Computer Studies*, *71(1)*, 91-109.

Lynch, C. (2014). *The Diagnosticity of Argument Diagrams*. Doctoral Dissertation, University of Pittsburgh.

Lynch, C., Ashley, K. D., & Chi, M. (2014). Can Diagrams Predict Essay Grades? *Lecture Notes in Computer Science*, *8474*, 260-265.

National Center for Education Statistics (2011). *The Nation's Report Card: Writing 2011* (NCES 2012–457). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Nussbaum, E. M., Schraw, G. (2007). Promoting Argument-Counterargument Integration in Students' Writing. *The Journal of Experimental Education, 76(1)*, 59-92.

Okada, T. & Shimokido, T. (2001). The role of hypothesis formation in a community of psychology. In K. Crowley, C.D. Schunn, & T. Okada (Eds.), Designing for Science: Implications from everyday, classroom, and professional settings (pp. 445-464). Mahwah, NJ: Erlbaum.

Oostdam, R. J., & Emmelot, Y. W. (1991). Education in argumentation skills at Dutch secondary schools. *Proceedings of the Second International Conference on Argumentation*. Amsterdam: Sic Sat.

Oostdam, R., Glopper, K. D., & Eiting, M. H. (1994). Argumentation in written discourse: secondary school students' writing problems. *Studies in Pragma-dialectcs*. Amsterdam: Sic Sat.

Persky, H. R., Daane, M. C., & Jin, Y. (2002). *The Nation's Report Card: Writing 2002*, (NCES 2003-529). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Popper, K. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge* (5th ed.). London: Routledge.

Proske, A., Narciss, S., & McNamara, D. S. (2012). Computer‐based scaffolding to facilitate students' development of expertise in academic writing *Journal of Research in Reading*, *35*(2), 136-152.

Salahu-Din, D., Persky, H., and Miller, J. (2008). *The Nation's Report Card: Writing 2007* (NCES 2008–468). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Toth, E. E., Suthers, D. D., & Lesgold, A. M. (2002). "Mapping to know": The effects of representational guidance and reflective assessment on scientific inquiry. *Science Education, 86(2)*, 264-286.

Twardy, C. (2004). Argument maps improve critical thinking. *Teaching Philosophy*, *27*(2), 95-116.

Whately, R. (1834). *Elements of logic: comprising the substance of the article in the Encyclopaedia metropolitan: with additions, &c.* (5th ed.). London: B. Fellowes.