ICASE
International Council of
Association for Science Education

Constructing arguments: Investigating pre-service science teachers' argumentation skills in a socio-scientific context

Brooke Robertshaw[*], Todd Campbell[†‡]

ABSTRACT: As western society becomes increasingly reliant on scientific information to make decisions, citizens must be equipped to understand how scientific arguments are constructed. In order to do this, pre-service teachers must be prepared to foster students' abilities and understandings of scientific argumentation in the classroom. This study investigated how instruction in the Toulmin Argumentation Protocol (TAP) impacted pre-service science teachers' ability to write sound and logical scientific arguments. The study occurred in the context of a pre-service methods class on the socio-scientific realm of secondary science education at a university in the USA. Through the use of quantitative methods, investigation findings indicate that there was a positive impact on pre-service science teachers' ability to construct sound scientific arguments through instruction in the TAP within the one semester course where this research took place.

KEY WORDS: Argumentation, Socio-Scientific Issues, Pre-Service Science Teachers

INTRODUCTION

As western society evolves, it has become more and more reliant on scientific concepts to judge how to move forward. Examples of this include looking to science to explain and give guidance about how humans should interact with the earth's ecosystem or understanding what stem cells do and the most ethical uses of these cells. These are just a few examples of hundreds, possibly thousands, of ways that our current western society looks toward scientific work to guide us into the future.

In order to be the best possible consumers of scientific information for making personal, political and ethical decisions, we need to understand how this scientific knowledge came to be or the epistemic origins of this knowledge and the foundations upon which it is being presented. In the U.S., A Conceptual Framework for K-12 Science Education (NRC, 2011) supports this, including science practices as a central strand of science learning. Linked to the practices of science learning, the NRC conceptual framework emphasizes being able to understand and create arguments

---
[*] Utah State University, USA
[†] University of Connecticut, USA
[‡] Corresponding Author: todd.campbell@umassd.edu

about phenomena that use scientific logic. In order to produce citizens that can process and evaluate science information, students must understand how evidence is used in coordination with theory, how to assess the validity and reliability of both data and arguments, and how to engage in the praxis of constructing arguments (Osborne, Erduran, & Simon, 2004; Osborne, 2010; Sampson & Clark, 2008; Venville & Dawson, 2010).

A problem that science educators are trying to overcome is how to encourage instruction on scientific argumentation in the classroom (Erduran, Ardac, & Yakmaci-Guzel, 2008; Osborne, Erduran, Simon, 2004). Two factors that can impede instruction in science argumentation is the need for a great majority of science teachers to change their instructional practices in order to allow their students to effectively learn these skills, and, if students are to practice these skills in the classroom a teacher may have to give up some of the authority in the classroom (Osborne, Erduran, & Simon, 2004). Additionally, work may also be needed to better align science teachers' epistemological commitments so that they are more congruent with those of science (Sandoval & Resier, 2004). Currently there are many studies that investigate ways to aid in-service science teachers in incorporating argumentation in their classrooms (McNeill & Pimentel , 2010; Osborne, Erduran, & Simon, 2004; Osborne, 2010; Sampson & Clark, 2008; Venville & Dawson, 2010), but there are few studies that address instruction in pre-service science education.

This study addresses the dearth of investigations with pre-service science teachers, by focusing at the pre-service level to address recommendations for helping to overcome some of the obstacles faced by in-service teachers (Erduran, Ardac, &Yakmaci-Guzel, 2008). Therefore, the following research question is investigated in this study:

What changes can be found in pre-service students' abilities to write scientific arguments through participation in a one-semester course focused on engaging in and learning about teaching socio-scientific issues?

THEORETICAL FRAMEWORK

Socio-scientific Issues as a Context for Argumentation

Considering argumentation in the context of engaging students in socio-scientific issues has received increased attention in science education and science education literature over the last decade (e.g., Evagoroua, Pilar Jimenez-Aleixandre & Osborne, 2012; Pilar Jimenez-Aleixandre, 2002; Sadler & Donnely, 2006). Part of this increased attention can be attributed to what Roberts (2007) describes as two visions for science education. Vision I is concerned with science education that attends to cultivating understandings about science concepts, laws, theories and processes. These are foci that are well aligned with standards documents from the last 20

years (Bybee, Fensham, and Laurie 2009; Bybee, McCrae, and Laurie 2009; Osborne 2007). Vision II for science education is concerned with 'situations in which science has a role, such as decision-making about socio-scientific issues' (Roberts 2007, 9). Sadler and Zeidler (2009) describe science instruction emanating from Vision II as 'progressive science education'. This education in science is concerned with citizens' understanding of science (c.f., Fensham, 2004), humanistic science education, context-based science teaching (c.f., Markic & Eilks, 2006), science–technology–society (STS) (c.f., Yager, 2007) and socio-scientific issues (SSI) (c.f., Sadler and Zeidler, 2009). As can be seen, socio-scientific issues are among the platforms for reshaping science education so that Robert's (2007) visions I and II for science education become central for experiences of students in classrooms. Socio-scientific issues become relevant in science education as science literacy is broadly conceptualized to include informed decision making founded on ability with analysis, synthesis, and the evaluation of information, while concurrently weighing moral reasoning and ethical issues and developing an epistemic awareness of the connectedness of issues scientific (Zeidler *et al*., 2005). But, what are Socio-scientific Issues (SSI)? Sadler and Zeidler (2004) provide a response to this question and the fit of these issues in science education:

> Socio-scientific issues describe societal dilemmas with conceptual, procedural, or technological links to science. Many socio-scientific issues stem from dilemmas involving biotechnology, environmental problems, and human genetics. The suggestion that issues such as those related to genetic engineering and environmental challenges can be classified together as "socio-scientific issues" is not meant to imply that science and society represent independent entities. On the contrary, all aspects of science are inseparable from the society from which they arise. However, the topics described by the phrase "socio-scientific issues" display a unique degree of societal interest, effect, and consequent (p. 5).

Because of the complexity of socio-scientific issues, it is not surprising that these issues have been identified for providing a rich context for teaching and learning argumentation (McDonald, 2010; Osborne, Erduran, & Simon, 2004; Sadler & Donnelly, 2006). In fact, much research exists documenting the effectiveness of socio-scientific issues as a platform for developing argumentation skills (Driver, Newton, & Osborne, 2000; Patronis, Potari, & Spiliotopoulou, 1999; Zohar & Nemet, 2002). Socio-scientific issues bring together scientific theories and laws in coordination with evidence, with a social context in which personal, ethical and lawful considerations are aroused or needed. So, the use of socio-scientific issues, along with teaching of argumentation allows students to draw on their own life and relevant community experiences which can lead to understanding and leveraging science concepts, process, laws and theories on a deeper level (McNeill & Pimentel, 2010). Additionally, it is believed

that attention in helping pre-service teachers better understand and envision the role of argumentation is needed, especially if these pre-service teachers are products of the classrooms where argumentation is scarce (Erduran, Ardac, & Yakmaci-Guzel, 2008; Osborne, Erduran, Simon, 2004). So, like the case made for students, socio-scientific issues provide a promising setting for this work. A stronger footing for thinking and understanding argumentation is considered next.

Argumentation and Toulmin's Argumentation Framework

While there are variances in how argumentation is fostered and investigated in science education (Osborne, Eduran, Simon, 2004; Cavegnetto, 2010; Sadler & Zeidler 2005; Sadler & Donnelly, 2006), more generally there is agreement that argumentation is a centrally important practice of science that shapes the work of scientists (Siegel, 1995; Toulmin, 1958) and therefore should also be centrally important in science classrooms (Erduran, Simon, & Osborne, 2004; NRC, 2011; Sadler & Zeidler, 2005).

A framework that has been used extensively to help students, and teachers, learn how to construct sound scientific arguments is the Toulmin Argumentation Protocol (TAP) (McNeill & Pimentel, 2010; Novak, McNeill & Krajcik, 2009; Osborne, Erduran, & Simon, 2004; Sadler & Zeidler, 2005; Venville & Dawson, 2010). The TAP defines seven different structural components that make up an argument: claim, data, warrant, backing, qualifier and rebuttal (Toulmin 1958, 1988). See Table 1 for an explanation of each of these components.

A properly formed argument has these components being inter-dependent and building on each other (Toulmin 1958, 1988). Figures 1 and 2 visually illustrate a properly formed argument using the TAP.

Limitations of Toulmin's Argumentation Framework

Common criticisms of the TAP are that (a) it only focuses on the *structure* of an argument, while not addressing the *quality* of the argument being examined (Abi-El-Mona & Abd-El-Khalick, 2011), (b) the lines between its different structural components can be uncertain at times (Sampson & Clark, 2008), and (c) the dissection of structural components of an argument may leave the dialectical features responsible for driving arguments under-examined or under-emphasized (Nielsen, in press).

Table 1. **Criteria for judging the quality of structural components of argumentation**

Structural Facet	Level	Descriptor
Claim	High	A claim without an opinion that includes back-

		ground information
	Medium	Stating an opinion with background information or stating a stance on an issue, that isn't stated as an opinion, but without background information
	Low	Simply stating an opinion
Data	High	Empirical: The use of specific data to back up the claim. The use of specific data to back up the claim. This evidence can include conceptual information as well. This is connected with evidence and data to the claim
	Medium	Conceptual: The use of conceptual information to back up a claim. This level *may* also include a personal opinion in linking the conceptual information to the claim. It does not rely on specific data to back up the claim, but includes more than a personal opinion
	Low	Opinion: The use of a personal opinion to back up a claim
Warrants, Backings, Qualifiers	High	Scientific: Data and reasoning that scientists use to investigate the phenomenon being argued, such as glaciers melting, sea levels, air temperature, water temperature, or species disturbance (McNeill & Pimentel, 2010) Data and theoretical groundings are connected in logical ways similar to ways in which scientists do this as well; Coordination of theory and evidence in the same ways that scientists use to connect data to hypotheses.
	Medium	Rational: Logical, attempts to use scientific understanding and language, is expressed through discussions of general scientific principles, possibly connected to personal experiences. (Dawson & Venville, 2010)
	Low	Personal: This is reasoning that relies on ideas from students' everyday lives, including, but not limited to, a student's opinion, personal feelings about the phenomena being studied, or expression of a student's expertise in an area to justify their claim. (McNeill & Pimentel, 2010)
Rebuttal	High	A counter-statement to the claim that uses empirical or conceptual evidence as well as using scientific reasoning. A high level rebuttal also refutes the counter-claim using scientific reasoning and empirical or conceptual evidence. This level of rebuttal is almost a complete argument within itself.
	Medium	A counter-statement to the claim that uses conceptual evidence, with a personal opinion possibly connected to refute the claim. The reasoning uses rational logic that makes an attempt to use scien-

	tific understanding and language. A medium level rebuttal also refutes the counter-statement using personal opinion and/or conceptual evidence.
Low	A counter-statement to the claim that uses a personal opinion to refute the claim. It may or may not also refute the counter-statement to the claim and if it does it relies solely on a personal opinion. Conceptual information may be included in the personal opinion, but the overall effect of the statement is an opinion. The conceptual information is not the central focus of the statement.

These limitations are well founded, and studies, including this one, have worked to overcome the first two of these limitations (i.e., uncertainty and quality). To address argument quality this study, and others, created explanations to consider how well an argument, using the TAP, is written (McNeill & Pimentel, 2010; Venville & Dawson, 2010). While understanding the uncertainties between the different TAP components continues, this current research has achieved some level of satisfaction in disaggregating the components of the TAP through the creation and use of a rubric.
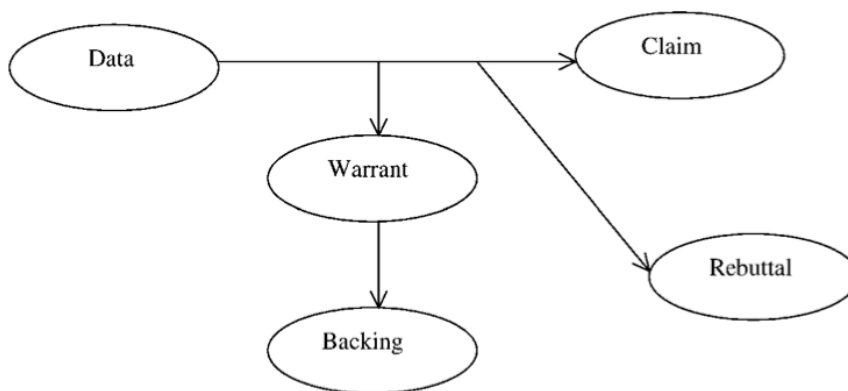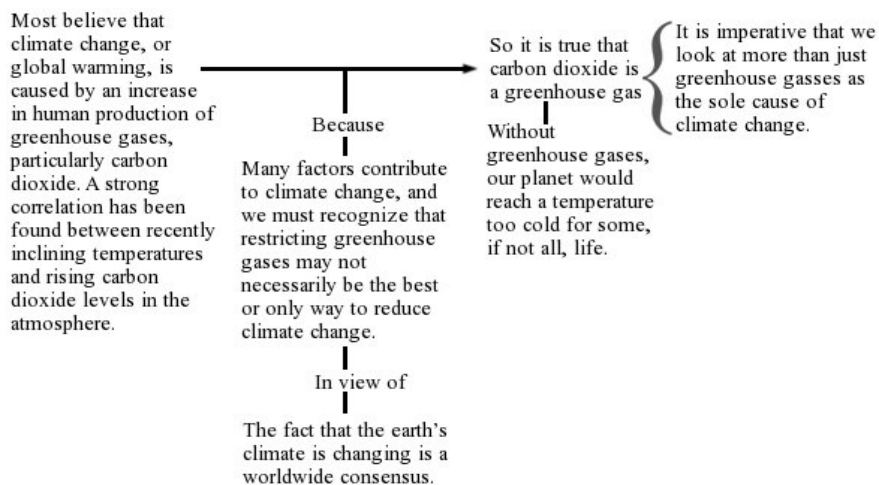
Data → Claim
Data → Warrant
Data → Rebuttal
Warrant → Backing

Figure 1.   A diagram of the Toulmin's (1958) framework for argumentation.

Our research acknowledges these two commonly levied issues that are critical for proper use of the TAP (i.e. structure and quality) and has taken measures to address these so that the TAP can be used to aid students in developing improved scientific literacy connected to understanding and creating sound scientific arguments (Erduran, Ardac, & Yakmaci-Guzel, 2008, McNeill & Pimentel, 2010; Novak, McNeill & Krajcik, 2009; Os-

borne, Erduran, & Simon, 2004; Sadler & Zeidler, 2005; Venville & Dawson, 2010).

Most believe that climate change, or global warming, is caused by an increase in human production of greenhouse gases, particularly carbon dioxide. A strong correlation has been found between recently inclining temperatures and rising carbon dioxide levels in the atmosphere.

Because

Many factors contribute to climate change, and we must recognize that restricting greenhouse gases may not necessarily be the best or only way to reduce climate change.

So it is true that carbon dioxide is a greenhouse gas

Without greenhouse gases, our planet would reach a temperature too cold for some, if not all, life.

It is imperative that we look at more than just greenhouse gasses as the sole cause of climate change.

In view of

The fact that the earth's climate is changing is a worldwide consensus.

Adapted from Toulmin (1988)

Figure 2.   The TAP with example components.

Finally, Neilsen's (in press) criticism of the inability of TAP to deal with the dialectical nature of arguments is acknowledged as an ongoing concern. Nielsen (in press) suggests that "[t]here will always be a certain trade-off between having measurable constructs in the form of (informal) logical relations between core elements, on the one hand, and taking account of the dialectical context in which they originated, on the other" (p. 19). In this present research and work with pre-service teachers, we willingly 'trade off' a focus on the dialectical features, to ensure that these participants begin to consider 'core elements' of an argument, both in their own arguments and as a precursor to focusing on their future students' arguments. We see the TAP as a tractable framework for this and intend to continue to consider and look for ways to examine the dialectical features of arguments, especially ways that will lend themselves as tractable means for teachers to introduce and look for the same dialectical features with their students.

RESEARCH PURPOSE

In order to address the lack of argumentation instruction in the classroom, this study set out to instruct pre-service teachers in argumentation, using the Toulmin Argumentation Protocol (TAP), in a pre-service science methods course, set in a socio-scientific context. The study set out to investigate whether there was a change in secondary science education pre-

service science teachers' ability to write scientific arguments, using the TAP, from the beginning to the end of the course. Focusing on pre-service science teachers' abilities to craft scientific arguments in writing is seen as a precursor or foundation for future important work of supporting and fostering pre-service teachers' ability to engage students in meaningful instruction targeting increased student understanding and abilities with scientific argumentation.

METHODS

This study used a quantitative approach, whereby inferential statistics and effect sizes were used to examine the change in the pre-service teachers' abilities to write a sound scientific argument based on the TAP. To this end, a rubric was created by the researchers to examine the quality of the individual facets of the TAP and the overall quality of each written argument. Finally the results of the quantitative analyses were compared to the self-report of participant changes in their ability to write arguments collected at the end of the semester.

CONTEXT

This study was completed within the context of a required secondary science education methods course at a university in the Rocky Mountain Region of the USA. The course focused on understanding how science and society interact, and how pre-service science teachers could discuss these issues with their students. A professor and a doctoral student teaching assistant, who were also the researchers, taught the course. This was the sixth time the course had been taught by the professor, and the first time the TA had participated in the course. The course used a mix of pedagogies, but relied most on self-reflection within student-centered pedagogies. Students used blog posts to reflect and construct arguments, and analyzing self-chosen writings by others for facets of sound scientific arguments. The last third of the course focused on the group selecting a socio-scientific issue to examine more deeply before creating resources to educate others about the issue. This project included the students writing arguments for candidate topics, and then, as a group selecting and writing a TAP based argument for the topic selected as the basis of the class action.

PARTICIPANTS

There were nine pre-service teachers in the course, seven of whom participated in the research. All pre-service teachers were enrolled in the sec-

ondary education program and all had finished their science disciplinary courses, but varied in whether they had taken secondary science methods coursework, with some students having completed none, one, or two semesters of the secondary science education courses. One participant was working on her master's degree in secondary science education and had previous experience teaching in a secondary school. She had not taken a course focusing on the socio-scientific realm of science instruction prior to this course, nor had she experienced a course focusing on argumentation.

PROCEDURE

Two assignments given during the course were used to answer the research questions. In the first assignment, which was designated as the pretest, students wrote a scientific argument during week 3 of the course, before they were exposed to the TAP. At the end of the course, week 15, the students re-wrote the arguments from week 3 with instructions to consider what they had learned about TAP as they re-wrote their assignments. They were also instructed to reflect on how their arguments had changed from week 3 to the end of the course. This second assignment was designated as the post-test. Between the two assignments students received direct and experiential instruction on the TAP.

ANALYSIS

There were four steps involved in the study. First, each researcher using the TAP components as pre-defined categories, analyzed each argument independently. Agreement on coding of the arguments was reached through discussion. Stemming from the discussion of the components of the TAP, the second step was undertaken – a rubric was created to establish the quality of each component as well as the overall argument. As stated in the literature review, one of the drawbacks of using the Toulmin argumentation protocol is that it does not give an indication of how to judge the *quality* of an argument or facets of an argument. We worked to address this issue through following similar ideas by McNeill and Pimentel (2010), Dawson and Venville (2009), and Venville and Dawson (2010). Due to the complexity of Toulmin's scheme, and the differences in the facets of the TAP we felt we could not judge each facet with the same logic. Included in this rubric are four different labeling systems – 1 for the claim, 1 for the data, 1 for warrants, backings and qualifiers, and 1 for rebuttals. Each instance of a component of the TAP was given a score of 1-3, indicating low-high quality of each facet, and then summed to give a score for each individual pre- and post-argument. The logic of how each facet of the TAP was given a high, medium or low rating was as shown in

table 2. As part of the development of the rubric, another science education researcher with research interests and publications focused on socio-scientific issues and argumentation reviewed and offered suggestions for strengthening the rubric. Based on this feedback and subsequent revisions, the rubric was finalized. Finally, using the rubric, the arguments were scored separately by the TA and professor. Using a weighted kappa, agreement between the two raters indicated absolute agreement, Kappa=.97 (p<.0001). Student reflections were also analyzed to see if students' self-reports of change in ability to write a scientific argument matched the findings by the professor and TA.

Table 2.   Toulmin Argumentation Protocol components and definitions

Facet	Definition
Claim	What is to be established in the argument
Data	Explicit facts that give foundation to the claim
Warrant	Justification for the data, rules and principles to strengthen the argument
Backing	Why the warrant should be accepted, without it the warrant has no weight
Qualifier	Gives the degree to which the data strengthens the claim in light of the warrant
Rebuttal	A circumstance when the warrant should be put aside.

FINDINGS

To determine whether there was a difference in students' ability to write a scientific argument following the TAP guidelines, the Friedman Rank Sum test was used. This test was used because of the low number of participants, the violation of assumptions of normality and homogeneity of variance, and its robustness with ordinal data and for testing repeated measures (Cohen, 2001; Sheskin, 2004). After converting the data to rank scores, the results showed that there was no difference in students' ability to write a TAP argument from the beginning to the end of the course ($\chi^2(1)=1.29$, p=.26). To further illustrate the findings, Table 3 shows the median for each facet, pre- and post-, as well as the minimum and maximum scores for each.

Table 3. **Median, Minimum and Maximum score for each argument and the total argument score for pre and post arguments**

	Claim		Data		War-rant		Backing		Qualifier		Rebuttal		Total	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Median	1	2	2	5	1	2	1	1	1	0	1	2	9	13
Min	1	1	1	1	1	0	0	0	0	0	0	0	4	6
Max	3	3	5	7	3	6	3	5	3	4	3	4	15	23

Note: There is no minimum or maximum possible score for the argumentation scores.

While the Friedman's test showed no difference in the argumentation scores, effect sizes were computed for the total scores for each component of the TAP as well as the total score. Often a small sample size will not show significance when there may be, plus this enables us to examine possible gains among the separate facets of the students' argumentation and nature of science skills. Vargha-Delaney's A (2000) was used to compute the effect size. This statistic is more robust to ordinal data (Leech & Onwuegbuzie, 2002; Vargha & Delaney, 2000) and Leech and Onwuegbuzie (2002) found the A statistic, being one of two effect size measures, to be most robust to violations of normality assumptions. As can be seen in Table 4 the effect size calculation was able to detect large gains for argumentation skills in understanding the use of data, rebuttals, and overall creation of arguments, and showed a medium gain for students' ability to write claims. It also detected that there was a medium drop in students' use of qualifiers.

It is also important to note that those facets with the lowest effect sizes were warrants and backings (A=.57 for each). One reason for this could be because of the difficulty in discerning between backings and warrants; a limitation of the Toulmin Argumentation protocol that has been discussed in the literature (Abi-El-Mona & Abd-El-Khalick, 2011; Sampson & Clark, 2008). This limitation was also a point of discussion between the two researchers throughout the study, as well as an issue that was brought up on more than one occasion by the participants in the study during class sessions. Another possible explanation for the lower effect sizes of the warrants and backings might also be explained by the difficulties others researchers have identified, in U.S. schools in particular. In this research, students in U.S. classrooms are found making claims that rely on data or are aptly described as correlational conclusions, but are not found moving

beyond these conclusions to coordinate theoretical ideas with data to develop explanations (Kang, Orgill, & Crippen, 2008). In some ways, this resembled the findings in this current study as arguments improved with respect to increased focus on claims, data, and rebuttals, but did not improve with respect to the backings and warrants or the reasons that connect (warrants) and ground (backings) data and claims.

Table 4. **Effect sizes for each TAP component and overall score. (Vargha and Delaney (2000) set values to be 0.56=small, 0.64=medium and 0.71=large)**

Component	Value
Backing	A= 0.57
Claim	A= 0.64
Data	A= 0.71
Qualifier	A= -0.64
Rebuttal	A= 0.79
Warrant	A= 0.57
Argumentation Total Score	A= 0.74

DISCUSSION

Due to past concerns for the existence of unclear boundaries between warrants and backings when applying the descriptions to written socio-scientific arguments, the researchers worked to address this limitation during analysis through diagramming the pieces of the arguments and considering them within the context in which they occurred (e.g., Figure 2). Based on the level of rater agreement that was found and reported in this research, this strategy proved helpful.

As can be seen in the effect size calculations, there was an overall positive impact on pre-service teachers' abilities to write scientific arguments as laid out by Toulmin, through participation in the course. This can also be seen in Table 3, which shows a general trend of improvement in scores from pre- to post-argumentation instruction. These findings were further reinforced by the students in their reflections about how their arguments had changed from the beginning to the end of the course. Most students commented that the first version of their argument was disorganized, jumbled, and one student even found fallacies in the version written during week three of the course. Additionally, all students felt that the TAP helped them to become more organized, and more directed in their argument. They were able to see how the different components fit together, and how that made for a stronger, more scientific argument. Examples

of reflections offered by the students that help reify these findings are as follows:

> As I started the revision process, I realized how disorganized my thoughts were . . . they were all over the place. There were claims with no data, backing with no qualifiers, and I didn't dare add a rebuttal for fear that it might look like I was weakening my argument.

> My revised societal issue paper had significantly more direction . . . the first time I looked up some information to support my thoughts, but it was primarily just my opinions. This time I applied more actual examples or 'data' to support my claim.

So, while there is still room for improvements in certain structural components of argumentation after this one semester course, promise is found in the impact of concurrently engaging students in socio-scientific issues while also instructionally attending to the structural components of argumentation.

Finally, the one area where the effect size indicated a negative trend was the use of qualifiers by students. As we consider what might have contributed to this finding, we are left to wonder whether or not enough attention was focused on the use of qualifiers. This finding warrants additional examination into the future. As an example, it may be possible that as the pre-service teachers recognized that their arguments were stronger because they had better supported them with data, as one example, they felt less compelled to qualify the claims made in the arguments. This is an untested explanation, but one that merits further exploration in future research. In summary, while the instructors of the course tried to help students understand the TAP, this finding (i.e., the decreased quality of rebuttals) indicated an area for focus into the future.

CONCLUSION

As society continues moving in a direction of putting citizens in the position to make personal, ethical, and lawful decisions based on science, these citizens must be prepared to understand how these arguments are formed. They must even be able to create their own scientific arguments. This study resulted from a science in society course where pre-service science teachers were instructed in scientific argumentation in order to prepare them to do the same in their own future classrooms. The instruction given during the course was direct, as well as experiential. While this study only investigated seven students, it did show that instruction in the TAP can improve argumentation skills within a one semester course, a finding that is significant as university science teacher educators work to balance the time focused on one important aspect of science teaching with

time spent on the multitude of other important foundational science teaching foci that support pre-service teachers as they enter classrooms.

LIMITATIONS AND FUTURE RESEARCH

A major limitation to this study was the number of participants. We expect to continue this research by investigating the impact of TAP in informing and improving pre-service teachers' argumentation skills among larger populations of participants in the future, but feel that the use of non-parametric statistics, at this point, at least help us ensure that reliable methods of detection have been employed appropriately within the population studied to date.

Additionally, while the instructor has taught this particular course six times previously, he had not used the TAP as explicitly in the past. It is expected that increased experience with this framework into the future, informed by the findings of this research, will continue to improve the promising outcomes already realized, while also focusing more intently on those outcomes that are not as promising to date (e.g., looking for more improvement over time with respect to the quality of claims, warrants, and backings). This study indicated, as did other research using the TAP (Abi-El-Mona & Abd-El-Khalick, 2011; McNeill & Pimentel, 2010; Venville & Dawson, 2010), that a common measure of the quality of TAP based arguments is needed. Future work by the researchers of this study would be to repeat this study, as well as to explore partnerships with other researchers and instructors who were using the TAP to investigate the impact of their practices on pre-service science teachers argumentation skills.

ACKNOWLEDGEMENTS

REFERENCES

Abi-El-Mona & Abd-El-Khalick, (2011). Perceptions of the nature and 'goodness' of argument among college students, science teachers, and scientists. *International Journal of Science Education*, 33(4), 573–605.

Bybee, R., Fensham, P., & Laurie. R. (2009). Scientific literacy and contexts in PISA 2006 science. *Journal of Research in Science Teaching,* 46(8), 862–864.

Bybee, R., McCrae. B and R. Laurie. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching,* 46(8), 865–883.

Cavagnetto, A. R. (2010). Argument to Foster Scientific Literacy: A Review of Argument Interventions in K–12 Science Contexts. *Review of Educational Research*, 80(3), 336–371

Cohen, B. (2001). *Explaining psychological statistics* (2nd ed.). Hoboken, NJ: Wiley.

Dawson, V. & Venville, G. (2009). High-school students' informal reasoning and argumentation about biotechnology: an indicator of scientific literacy? *International Journal of Science Education*, 31(11), 1421-1445.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84, 287–312.

Erduran, S., Ardac, D., & Yakmaci-Guzel, B. (2006). Learning to teach argumentation: Case studies of pre-service secondary science teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 2*(2), 1-14.

Evagoroua, M., Pilar Jimenez-Aleixandre, M., & Osborne, J. (2012). 'Should We Kill the Grey Squirrels?' A Study Exploring Students' Justifications and Decision-Making. *International Journal of Science Education*, 34(3), 401-428.

Fensham, P. (2004). Increasing the relevance of science and technology education for all students in the 21st century. *Science Education International*, 15(1), 7-26.

Kang, N., Orgill, M., & Crippen, K. J. (2008). Understanding Teachers' Conceptions of Classroom Inquiry With a Teaching Scenario Survey Instrument. *Journal of Science Teacher Education*, 19(4), 337–354

Leech, N. L. & Onwuegbuzie, A. J. (2002). *A Call for Greater Use of Nonparametric Statistics*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Chattanooga, TN.

Markic, S. & Eilks, I. (2006). Cooperative and context-based learning on electrochemical cells in lower secondary chemistry: A project of participatory action research. *Science Education International*, 17(4), 253-273.

McDonald, C. V. (2010). The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science. *Journal of Research in Science Teaching, 47*(9), 1137-1164.

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229.

National Research Council (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Nielsen, J. A. (in press). Dialectical Features of Students' Argumentation: A Critical Review of Argumentation Studies. *Research in Science Education*, Online first DOI 10.1007/s11165-011-9266-x

Novak, A., McNeill, K., Krajcik, J. (2009). Helping students write scientific explanations. *Science Scope, 33*(1), 54-56.

Osborne, J. (2007). Science education for the twenty first century. *Eurasia Journal of Mathematics, Science and Technology Education,* 3(3) 173–184.

Osborne, J. (2010). An Argument for Arguments in Science Classes. *Phi Delta Kappan, 91*(4), 62.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020.

Patronis, T., Potari, D., & Spiliotopoulou, V. (1999). Students' argumentation in decision-making on a socio-scientific issue: Implications for teaching. *International Journal of Science Education*, 21, 745–754.

Pilar Jimenez-Aleixandre, M. (2002). Knowledge producers or knowledge consumers? Argumentation and decision making about environmental management. *International Journal of Science Education*, 24(11), 1171-1190.

Roberts, D. (2007). Scientific literacy/science literacy. In *Handbook of research on science education*, S.K. Abell, and N.G. Lederman, 729–80. Mahwah, NJ: Lawrence Erlbaum Associates.

Sadler, T. D. & Donnelly, L. A. (2006). Socio-scientific Argumentation: The effects of content knowledge and morality. *International Journal of Science Education,* 28, 1463-1488.

Sadler, T.D. & Zeidler D. L. (2009). Scientific literacy, PISA, and socio-scientific discourse: Assessment for progressive aims of science education. *Journal of Research in Science Teaching,* 46(8), 909–921.

Sadler, T. D. & Zeidler, D. L. (2004). The morality of socio-scientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education*, 88, 4–27.

Sadler, T. D. & Zeidler, D. L. (2005). The significance of content knowledge for informal reasoning regarding socio-scientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education*, 89(1), 71-93.

Sampson, V. & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and

recommendations for future directions. *Science Education, 92*(3), 447-472.

Sandoval, W.A. and Reiser, B.J. (2004).Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88, 345-372.

Siegel, H. (1995). Why should educators care about argumentation? *Informal Logic*, 17(2), 159–176.

Sheskin, D. (2004). *Handbook of Parametric and Nonparametric Statistical Procedures*, Boca Raton, Fl; CRC Press LLC.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Toulmin, S. (1988). *The uses of argument* (9th ed.). Cambridge, UK: Cambridge University Press.

Vargha, A. & Delaney, H. D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics, 25*(2), 101.

Venville, G. J. & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, 47(8), 952-977.

Yager, R. E. (2007). The six "C" pyramid for realizing success with STS instruction. *Science Education International*, 18(2), 85-91.

Zeidler, D.L., Sadler, T.D., Simmons, M.L. & Howes, E.V. (2005). Beyond STS: A research- based framework for socio-scientific issues education. *Science Education, 89*(3), 357-377.

Zohar, A. & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35–62.