Assessing Teachers' Professional Vision of Oral Scientific Argumentation

April Holton, Arizona State University, april.holton@asu.edu
J. Bryan Henderson, Arizona State University, jbryanh@asu.edu
Nicole Zillmer, Arizona State University, nicole.zillmer@gmail.com
Eric Greenwald, University of California, Berkeley, eric.greenwald@berkeley.edu
Megan Goss, University of California, Berkeley, mgoss@berkeley.edu
Christina Morales, University of California, Berkeley, cmmorales@berkeley.edu

Abstract: This paper reports the initial development of a measure to assess teachers' professional vision of oral scientific argumentation using online video annotation. Teachers watched segments of videos in which students engaged in whole class, small group, and dyad argumentation sessions and annotated the video with what they noticed in these sessions. Based on previous work on professional vision and our larger study on formative assessment of oral argumentation, a coding scheme was developed and tested with a pilot study. Researchers share initial findings of this pilot which showed differences in professional vision existed.

Keywords: professional vision, video annotation, argumentation

To develop student oral scientific argumentation requires a teacher to be able to attend to relevant science content, argument structure, and interpersonal skills within the moments of teaching and learning. Goodwin (1994) coined the term *professional vision* (PV) to describe the ability of experts in a domain to be able to recognize, highlight, and articulate salient components based on accepted knowledge of professionals in the field. Sherin and van Es (2009) investigated how *professional vision* develops in math teachers as they employ knowledge, as *noticings*, in moments of instruction to make sense of events with consideration for response to learning. In order to employ *noticings* during instruction, teachers need to develop an "expert" coding scheme for the context and to engage in knowledge-based reasoning when articulating what they are observing. It is recognized, however, that many teachers do not have such skills in their pedagogical toolbox, and the development of this ability to notice the salient points has professional development implications (Driver, Newton, & Osborne, 2000).

This study of professional vision was situated in a larger design-based research study (DiALoG: Diagnosing Argumentation Levels of Groups) to develop a digital formative assessment system for oral scientific argumentation. The DiALoG system identifies 8 items with teacher prompts to formatively assess: claims, evidence, reasoning, relevance, listening, critique, co-construction, and regulation. In interviews, pilot teachers expressed that holding the digital tool to reference when observing student talk helped them notice things they would have never noticed and they were less likely to fall into the IRE pattern. Furthermore, Sherin et al. (2011) found that teacher noticing led to responsive teaching, learning from teaching, or decomposing practice. This led us to argue that tools to support teachers with highlighting the important components during instruction might contribute to developing teachers' *professional vision* of oral scientific argumentation. Professional vision has been measured quantitatively and qualitatively through video annotation, but previous measures did not meet the needs of our study because they involved face-to-face interviews for verbal annotation (Sherin & van Es, 2009) or followed more general coding schemes for preservice teachers (Stürmer & Seidel, 2015). Thus, we attempted to develop a measure of professional vision, as other studies have, through video annotation, but through the lens of the coding scheme of oral scientific argumentation to answer the research question: *Can we measure teachers' professional vision of oral scientific argumentation through online video annotations?* In this paper, we share the development of a measure and the initial findings of a pilot study to test the instrument.

Methods

The pilot study included nine teachers (5 treatment, 4 control) who taught an argument-based curriculum to middle school students. The treatment group used the DiALoG formative assessment system with two units during student argumentation sessions.

To measure teachers' professional vision, ten minutes of video from the *Argumentation Toolkit* (http://www.argumentationtoolkit.org) showing students engaged in whole group, small group, and dyad oral argumentation sessions were segmented into 2-4 minutes. The segments were limited to avoid teachers becoming overwhelmed with the amount of interactions in the lesson. The video segments were validated by two argumentation experts, and then uploaded into *VideoAnt* (https://ant.umn.edu). Four places were marked for

teachers to annotate the preceding video clip with the simple questions: *What did you notice? What stands out to you?* (Sherin et al., 2009) The questions were broad to avoid prompting specific responses or vocabulary. Furthermore, teachers were invited to add annotations at any point in the video segments. The measure was administered through a shared link. Annotations were downloaded as a transcript and divided into idea units and analyzed by *topic* and *stance* (Sherin & Van Es, 2009).

The three master *topic* codes, *management*, *argument,* and *argumentation,* were divided into sub-categories. The codes are defined as:

Management

Teacher role: Topics that address the role or responses of teachers during the session.

Student role: Topics that address observations on the role students are playing in the discussion.

Student behavior: Topics that address the actions or attitudes, positive or negative, of students (i.e. misbehavior, distractions, polite).

Argument

Claim: Topics that suggest a tentative answer or position.

Evidence: Topics that address using data to support the claim.

Reasoning: Topics that demonstrate a thinking process that ties evidence(s) to a claim.

Relevance: Topics that suggest the accuracy or appropriateness of evidence to support a claim.

Argumentation

Listening: Topics that suggest that students are hearing and understanding others' contributions (i.e., clarifying, elaborating, body language).

Co-construction: Topics that address students integrating their thinking and ideas with others.

Critique: Topics that address students' attempts to push the thinking of others through critical questioning.

Regulation: Topics that suggest students manage discussion in an equitable way.

Stance codes, which described to what level the teacher engaged in reasoning about what they noticed, were: *describe, evaluate,* and *interpret* (Sherin & van Es, 2009). While coding posttest data, a fourth *stance* code emerged, *reflection.* The *stance* codes are defined as:

Describe: Teacher provides an account focused on observable features of the interaction.

Evaluate: Teacher includes judgments about the quality of the interactions.

Interpret: Teacher includes inferences about what took place in the interaction.

Reflect: Teacher makes connections to their own teaching practice.

First, two teachers' pre-test annotations (Teacher A and B) were analyzed to determine if a difference was detected in the teachers' responses. The two teachers were chosen based on their experience with argumentation. Teacher A worked with our DiALoG research team for two years and showed a developing use of argumentation with her students through observations and post interviews. Teacher B was new to oral scientific argumentation and a control teacher in our study. Based on observations and interview data, we predicted teacher A to have a more developed professional vision than Teacher B on the pretest annotations.

Next, we compared the pretest and posttest annotations by the mean percentage of idea units in topic and stance of the two groups (See Table 2). It is important to note that the pilot of the professional vision measure was not intended to make claims on whether DiALoG affected the professional vision of teachers, but instead, to test the coding schemes and to determine if the measure warrants further development for use in a follow-up random control trial with 100 teachers.

Analysis and findings

The topic and stance data were compared at the teacher level with Teacher A and B (Table 1), as well as at the group level, control vs. treatment (Table 2). The percentage of total idea units for the codes is reported.

Berliner (1994) found that novice teachers notice more surface-level features of classroom interactions while expert teachers are able to discriminate interactions that are more substantive. If a teacher's noticings target *management,* a surface-level feature, then one might consider the professional vision to be less developed when

compared to a teacher who focused less on the *management* and more on the components of argumentation. Teacher A (27%) focused on the management pieces of argumentation nine percent less than Teacher B (36%).

The *argument* and *argumentation* codes were based on the DiALoG progression. *Argument* components move from *claim* to *evidence* to *reasoning* with *relevance* applying to all components. And, the *argumentation* progression built from *listening* to *critique* to *co-construction* with *regulation* supporting all items. Applying these progressions, we proposed that a teacher who noticed components at the higher level of the progression would have a more developed professional vision. This did not hold true in the *argument* components which was not surprising since teachers have been exposed to arguments in other content areas, such as English language arts. However, *argumentation* components confirmed this pattern. Not only did Teacher A notice *argumentation* more than teacher B, 38% to 9% respectively, Teacher A highlighted the higher levels of the progression while Teacher B did not, Teacher A stated, "When you have to articulate what you are thinking, you can develop an idea or disprove your idea" and "They were all working together to come to an understanding of the correct results."

The *stance* codes represented the level of knowledge-based reasoning. Seidel and Sturmer (2015) argued that tasks requiring evaluation and/or inference/prediction are more difficult than tasks requiring descriptions of classroom situations. Teacher A had 51% *describe*, 20% *evaluate*, and 29% *interpret*, while teacher B was 73% *describe*, 18% *evaluate*, and 9% *interpret*. Based on the percentages of each stance, we might say that teacher A has a more developed PV because of the lower percentage of *describe* codes than teacher B, 51% to 73% respectively. In addition, Teacher A has more than three times the percentage of *interpret* codes than teacher B, 29% to 9% respectively. As teachers' PV develops, we expect the *describe* codes to shift to *evaluate* and *interpret*.

Table 1: Teacher A and B Pre-test Annotations

	Teacher A (n=34)	Teacher B (n=11)
	Percentage	Percentage
Topic		
Management		
Teacher role	.09	.09
Student(s) role(s)	.09	.09
Student behavior	.09	.18
Argument		
Claims	.06	.09
Evidence	.09	.18
Reasoning	.14	.27
Relevant	.09	0
Argumentation		
Listening	.03	.09
Critique	.14	0
Co-construction	.11	0
Regulation	.09	0
Stance		
Describe	.51	.73
Evaluate	.20	.18
Interpret	.29	.09

*n refers to the number of idea units

When comparing the pretest and posttest data of the DiALoG and control groups, we applied the framework from the two-teacher analysis to see if the same patterns appeared in the mean percentages. An additional code, *reflect*, emerged with the posttest analysis (See table 2).

The *topic* data showed the same pattern of decreased focus on *management* from pretest to posttest (DiALoG from 24% to 10% and control from 27% to 10%) which indicated that teachers attended to other components of the argumentation session, rather than the surface management, which suggested growth in PV. In addition, teachers showed a decrease in the percentage of *describe* stance (DiALoG decreased by 14% and control by 28%) and *evaluate* and *interpret* increased.

Table 2: Percentage of Idea Units for Topic and Stance for DiALoG and Control Groups

	DiALoG Group (n=5)		Control Group (n=4)	
	Pretest	Posttest	Pretest	Posttest

Topic				
Management				
Teacher role	.08	.03	.08	.07
Student(s) role(s)	.15	.07	.08	.00
Student behavior	.02	0	.11	.03
Argument				
Claims	.10	.02	.11	.03
Evidence	.24	.34	.20	.25
Reasoning	.08	.11	.24	.17
Relevance	.15	.10	0	.06
Argumentation				
Listening	.07	.10	.04	0
Critique	.04	.14	0	.10
Co-construction	.10	.13	0	.09
Regulation	.02	.09	.14	.16
Stance				
Describe	.64	.50	.76	.48
Evaluate	.15	.20	.18	.34
Interpret	.21	.27	.04	.17
Reflect	0	.02	.02	.02

*n represents the number of teachers

Discussion

The initial analysis provided evidence of differences to support that PV can be measured using remote video annotation. While the progressions of the argument and argumentation topics do not follow a consistent pattern, we will continue to explore this within the larger data set with the 100 teachers in a random control trial. The stance codes showed the most consistent trend to analyze teachers' knowledge-based reasoning. As teachers engage in deeper connections and reasoning of the topics they are highlighting, this signifies a growing professional vision. We have determined that the measure warrants continued development based on our findings. Next, multiple raters will use the coding scheme with pilot study data to determine rater reliability. With a larger sample size in the next phase, we will determine whether the differences are significant between the two groups after accounting for other factors such as years of teaching experience and hours of professional development.

References

Berliner, D. C. (1994). Expertise: The wonders of exemplary performance. *Creating powerful thinking in teachers and students*, 161-186.

Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287–312.

Goodwin, C. (1994). Professional vision. American anthropologist, 96(3), 606-633.

Sherin, M.G. & van Es, E. A. (2009). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education*, *60*(1), 20–37. https://doi.org/10.1177/0022487108328155

Sherin, M., Jacobs, V. & Philipp, R. (2011). *Mathematics Teacher Noticing: Seeing Through Teachers' Eyes*. Routledge.

Stürmer, K. & Seidel, T. (2015). Assessing Professional Vision in Teacher Candidates. *Zeitschrift Für Psychologie*, *223*(1), 54–63. https://doi.org/10.1027/2151-2604/a000200