



Quantitatively Investigating Inservice Elementary Teachers' Nature of Science Views

Hallie S. Edgerly¹ · Jerrid W. Kruse¹ · Jesse L. Wilcox¹

Accepted: 14 January 2021 / Published online: 12 March 2021

© The Author(s), under exclusive licence to Springer Nature B.V. 2021, corrected publication 2021

Abstract

Many nature of science (NOS) studies have demonstrated that teachers can improve their understandings of NOS with explicit and reflective instruction; however, the field has relied heavily on qualitative methods. Because qualitative studies can be difficult to compare across studies and contexts, our study sought to determine if a quantitative instrument, Students' Understanding of Science and Scientific Inquiry (SUSSI) (Liang et al., 2008), could detect differences in NOS understandings that might make large studies and comparisons across studies easier. This study investigated changes in 60 elementary teachers' NOS views during a year-long extensive science, technology, engineering, and math (STEM) professional development program designed for elementary teachers. We found four NOS constructs (Social and Cultural, Collaborative, Creative and Imagination, and Scientific Method) out of the eight on the SUSSI had an acceptable Cronbach's alpha values and we found that quantitative assessment could detect interesting differences in participants' NOS views from pre- to post-assessment.

Keywords Inservice elementary teachers · Nature of science · Professional development · Quantitative

Introduction

Nature of science (NOS) is an area of study that is informed by the history, philosophy, and sociology of science (Clough, 2006). NOS includes, but is not limited to, the values and assumptions scientists make as they develop their scientific knowledge (Lederman, 1992), how science works, how scientists collaborate, and the interaction between science and society (Clough, 2006). NOS has been a part of the science education literature for over a century (Lederman, 1992) and continues to be an important part of science education reform efforts (e.g., AAAS, 1989; NRC, 1996; NGSS Lead States, 2013).

✉ Hallie S. Edgerly
hallieedgerly@gmail.com

¹ Drake University, 2507 University Avenue, Des Moines, IA 50311, USA

Despite reform efforts that have included NOS for over thirty years, few K-12 students are explicitly taught NOS (Akerson & Abd-El-Khalick, 2005; Capps & Crawford, 2013). One reason NOS may not be taught is many teachers do not hold adequate views of NOS (Cofré et al., 2019; Lederman, 1992). This issue may be exacerbated for elementary teachers because they often have less preparation in science and science education than their secondary counterparts. A need exists to help elementary teachers better understand NOS ideas.

Although many studies have investigated the NOS views of primarily secondary teachers (e.g., Bell, Mulvey, & Maeng, 2016; Herman & Clough, 2016; Ward & Haigh, 2017), studies investigating elementary teachers' understanding of NOS are less common and are often carried out on preservice elementary teachers. For example, Kaya (2012) notes that even with more science coursework and additional explicit instruction, efforts may not be enough for advancing preservice elementary NOS views. Indeed, Akerson et al. (2006) use the VNOS-B survey and interviews to observe that while preservice elementary teachers' views of NOS ideas did improve after taking one course addressing NOS, their views reverted back to earlier views when surveyed and interviewed again five months after the completion of the course. While the explanation for this reversion is uncertain, the study makes clear how difficult combating years of misconceptions can be. Adding detail to the difficulties preservice teachers encounter, Mesci and Schwarz (2017) found preservice learners may struggle with some concepts more than others.

Considering preservice efforts may not be enough, researchers have also investigated inservice teacher learning of NOS. Unsurprisingly, research with inservice teachers has replicated and reinforced established findings with preservice teachers. For example, Akerson and Hanuscin (2007) found engaging inservice elementary teachers in implicit scientific inquiry activities does *not* improve their NOS understanding. However, elementary teachers do improve their understanding of NOS ideas through the use of the explicit and reflective approaches (Abd-El-Khalick & Akerson, 2004; Akerson et al., 2000).

Although in line with other research, studies on elementary teachers tend to be done with small numbers of participants. Typically, these studies have used qualitative methods (e.g., Abd-El-Khalick & Akerson, 2004; Akerson & Hansucin, 2007). These qualitative studies are very detailed, but their findings can be difficult to transfer or generalize. While the quotes used to illustrate findings in qualitative studies provide nuanced views, the quotes are necessarily attributed to individual participants. Despite the value of qualitative research, alternative avenues in which larger sample sizes can be used warrant exploration.

Purpose of Study

This study quantitatively investigates changes in the NOS views of elementary teachers participating in an intensive STEM professional development (PD) program. While many studies have demonstrated that teachers can improve their understanding of NOS with explicit and reflective instruction (e.g., Abd-El-Khalick & Akerson, 2004; Akerson & Hanuscin, 2007), the field has relied heavily on qualitative methods. Even though we do not doubt the rigor of these methods and hold with high regard the nuance gleaned from qualitative methods, like Shim, Young, & Paolucci (2010), we wondered to what extent a more quantitative approach might be useful to shed light on larger sample sizes in addition to also providing for detailed comparisons across participants. Although representative quotes often provided in qualitative studies are interesting, they perhaps may not transfer across studies or even across participants within a study. As the field progresses, ways to compare strategies, interventions,

or treatment conditions across studies may provide for more generalizable conclusions about how to best enact explicit and reflective NOS instruction. Therefore, this study sought to explore what possible nuances might come from quantitative approaches so that larger studies can be done. To explore the utility of quantitative approaches in detecting nuances in the changes of inservice teacher NOS views, this study investigates the following research questions.

Research Questions

1. Is there a statistically significant difference between pre- and post-measures of NOS concepts for elementary teachers who completed a year-long PD program in which NOS concepts were explicitly and reflectively taught?
2. What is the nature of the changes, if any, observed in the elementary teachers' NOS views?

Research Methods

Context of Study

Description of the PD and Participants The 60 participants of this PD program were inservice elementary teachers of a large urban school district in the Midwestern United States teaching a variety of grade levels from kindergarten through sixth grade. Like many elementary teachers, few had previous experience with NOS. The PD program was funded through the No Child Left Behind Title IIA program and a collaboration of the aforementioned urban school district and a mid-sized Midwestern university.

The PD program was an intensive year-long program starting in the summer and continuing during the academic year. Participants took four courses: mathematical practice, physical science, earth science, and life science. The three science courses focused on science content knowledge, science pedagogy, NOS, and the nature of technology and engineering (NOTE). NOS was consistently included throughout the science courses occurring weekly for three hours during an entire school year (30 total PD sessions). Therefore, because one semester of NOS instruction is not enough (Akerson et al., 2006), we studied the changes in inservice teachers' views of NOS over a whole year of instruction.

Holliday, Lederman, and Lederman (2014) make clear the qualifications of the program instructors matter in the success of a PD program in NOS. The qualifications of the two PD leaders of this study include successfully teaching NOS in K-12 settings, publishing on NOS instruction and research, as well as teaching NOS to inservice and preservice teachers for many years.

Description of NOS Teaching in the PD Program The activities and discussions to address NOS within the PD program were based on the explicit and reflective NOS instructional framework (Abd-El-Khalick & Lederman, 2000) and the conceptual change framework (Clough, 2006; Mesci & Schwartz, 2017). While our participants were engaged in using, doing, and reading about science, they were also asked general NOS questions (e.g., How is what you have done like what scientists do?) as well as more specific questions (e.g., How does your experience illustrate that science can change?) to guide their reflection toward particular aspects of NOS (Kruse et al., 2020).

An example of explicit and reflective NOS instruction from the PD is an investigation into falling objects. The instructor had participants predict the motion of a bean bag dropped from a plane, then asked further questions about when the participants believe they should release the bean bag to fall on a specific target, how they might explain trajectory, and predicting possible trajectories. Participants tried their ideas and made observations of bean bags falling with a horizontal motion. After this concrete experience, the instructor led a NOS discussion in which the following questions were asked: “How is what we’ve done like what scientists do? Why might scientists make predictions? Why do you think scientists might use models like we did? How might working together help scientists in their work?”

Data Collection

Many NOS frameworks and lists have been put forth to describe NOS constructs (e.g., McComas, 2004; Lederman et al., 2002; Clough, 2007). Researchers have begun to call into question the efficacy of reducing NOS ideas into a list of tenets, which can eschew the complexities and contextual NOS (Allchin, 2011; Clough, 2007). While we support a highly nuanced conception of NOS such as the Family Resemblance Approach (Izrik & Nola, 2011, 2014; Dagher & Erduran, 2016), we sought quantitative assessment of NOS views. At the time of this study (2015–2018), the Student Understanding of Science and Scientific Inquiry (SUSSI) (Liang et al., 2008), which was further modified by Herman et al. (2013), presented a promising opportunity and the Reconceptualized Family Resemblance Approach for NOS (RFN) questionnaire was not yet developed (Kaya, Erduran, Aksoz, & Akgun, 2019). We chose to use the original wording of all six of the Liang et al.’s (2008) constructs as well as the original wording of two of the constructs added by Herman et al. (2013). Therefore, our investigation was limited to NOS ideas included on the SUSSI instrument including: subjectivity, tentative, laws vs. theories, social and cultural influences, collaboration, scientific methods, creativity, and methodological naturalism. This list is not exhaustive, but rather focuses on well-established NOS ideas that have a fairly broad level of consensus across the history, philosophy, and sociology of science.

Participants’ views of NOS were gathered through responses to Likert scale questions and open-ended prompts of the SUSSI instrument (Herman et al., 2013; Liang et al., 2008) in a pre-test/post-test design before and after the PD program. Participants completed the SUSSI via an online survey tool and were given the assignment to complete the SUSSI outside of PD class time. The SUSSI includes four Likert items per NOS construct for a total of 32 Likert items on the pre- and post-assessments. A valuable aspect of the SUSSI instrument is the inclusion of open-ended prompts for each NOS construct. For each of the eight constructs evaluated using the SUSSI, the participants responded to four Likert items regarding the NOS construct. Following the four Likert items for each NOS construct, participants responded to an open-ended prompt to explain their thinking further. While the writing prompts were not the main source of the data analysis for this study, they did provide for greater confidence in data interpretation by comparing Likert responses with open-ended responses. Together, the two types of responses (Likert and open-ended) can provide for increased construct validity. Furthermore, because the instrument was developed by two different sets of NOS experts, (Herman et al., 2013; Liang et al., 2008), we have high confidence in the instrument’s content validity.

Liang et al. (2008) used calculations of the overall Cronbach's Alpha for the instrument to assess reliability of the SUSI instrument. Their findings indicated consistent alpha values of three sample groups tested. Therefore, the researchers concluded that the SUSI instrument can be used as a reliable test of participants' understanding of NOS with quantitative aspects allowing for inferential statistics.

Data Analysis

Using responses to the SUSI instrument, this study used a quantitative approach to assess participants' NOS views prior to and after a year-long PD. Likert response data was the main data used. However, constructed response data was used to validate data interpretation.

Transforming Selected (Likert) Response Data Participants responded to the Likert items related to each NOS construct evaluated. The Likert scale had five choices (i.e., strongly agree, agree, undecided, disagree, strongly disagree). These categories can be converted to numerical values (e.g., 1–5) with higher scores reflecting more accurate understanding. In Herman and Clough (2016), the authors used the Likert items of the SUSI to make initial determinations about participants' NOS understanding. In that study, the authors considered either a four or a five to indicate an "informed" view, while scores of three or lower were deemed to be "naive." Similarly, for our study, participant responses for each Likert item were scored as either aligned or not aligned based on the side of the scale most aligned with recent NOS literature. For example, in the item "Scientific theories are subject to on-going testing and revision," the aligned response was to agree or strongly agree, participants received a 1 if they responded with either category indicating an aligned, or informed view. If participants responded with strongly disagree, disagree, or undecided, they received a 0 indicating and not aligned, or naive view. This dichotomous approach for each item prevented having to assume that participants viewed the Likert scales as interval. Furthermore, experts consulted in the field of NOS noted that a response of "strongly agree" is not necessarily better than "agree" as counter examples may cause an individual with a nuanced understanding of NOS to not choose an extreme view.

While we did not assume intervals for the Likert scales, the dichotomous approach created interval scores for each NOS construct across four Likert items. Participants could answer each item incorrectly resulting in a score of zero for a particular NOS construct. Each correct item adds one point to the participants' score for the construct, with a maximum score of four. The overall NOS score was determined by summing the scores from each NOS construct. This led to possible scores between 0 and 32 because there were eight NOS constructs with the highest possible score on each construct being four. The overall Cronbach's alpha of all 32 items was acceptable as discussed below, but when we zoomed in to each NOS construct, only four constructs had acceptable Cronbach's alpha scores. Therefore, we ran *t*-tests for each of the four constructs (with an acceptable Cronbach's Alpha value, as described below) as well as for the combination of all 32 items.

We then wanted to zoom in beyond each construct to each item to explore nuanced changes across participants' views. Therefore, using the four constructs with acceptable Cronbach's alpha scores, we also analyzed each item using McNemar tests to determine what difference there might be in the proportion of participants demonstrating aligned views between the pre- and post-assessments on each particular item.

Establishing Construct Validity Through Constructed Response Data To check for construct validity, a random sample of more than one-third (23) of participants' open-ended responses were analyzed to determine if there was any obvious disagreement between participants' understanding of the NOS constructs based on their Likert responses and their explanations in their open-ended responses. Importantly, we did not "score" participants' open-ended responses. Instead, we searched their responses for utterances that opposed the response they gave in the corresponding Likert items. Therefore, we acknowledge participants may have simply not articulated ideas related to each Likert item. We did not find any such discrepancies in comparing the selected responses to the constructed responses data. If a disagreement had been found, we planned to either change the status of the participant's response to the related Likert item or remove the response. With this increased confidence in the validity of the Likert responses, tests were run using SPSS v.25 for analysis.

Establishing Construct Reliability Through Kuder-Richardson 20 Test The Kuder-Richardson 20 (KR-20) test of reliability was run to determine Cronbach's alpha of each NOS construct assessed by the SUSI instrument, as well as of the overall instrument. The KR-20 test of reliability is useful for testing reliability of dichotomous variables. The acceptable value of Cronbach's alpha for our study is any value near or greater than 0.7, according to Nunnally and Bernstein's (1967) suggestion of 0.5–0.7 as an acceptable range during beginning stages of research.

This test of reliability was done to determine which NOS constructs would have paired-samples *t*-tests run on them according to acceptable values of Cronbach's alpha.

Paired-Samples *t*-Tests Paired-samples *t*-tests were run to determine to what extent there is a difference between pre- and post-measures of the participants' understanding of NOS concepts. Five paired-samples *t*-tests were conducted to answer the first research question. The nature of science concepts assessed were social and cultural impacts on science, collaborative approach to science, scientists' use of imagination and creativity, and scientists' use of different methods. The fifth paired-samples *t*-test was run on the overall NOS score of the entire SUSI instrument (32 items).

McNemar's Tests To answer the second research question, McNemar's tests were run on each item individually for the NOS constructs that had paired-samples *t*-tests run. McNemar's tests were run to determine if there were particular aspects of the participants' thinking about each NOS idea that changed and to provide further insight into *t*-tests results. More specifically, we were looking to see if there is a statistically significant change in the proportion of participants demonstrating accurate views of each NOS idea, as measured by each individual item of the SUSI instrument.

Trustworthiness and Limitations Trustworthiness in our data is rooted in the SUSI instrument used. As Herman et al. (2013) points out, "the ability to obtain, evaluate, and cross compare Likert and qualitative data makes this instrument effective with large- or small-scale studies that assess participants' NOS understanding" (Liang et al., 2008, p. 5). Additionally, our data was analyzed and examined by more than one researcher.

As researchers, we are limited in our understanding of the participants' thinking when combining items into one larger value for the *t*-tests. This potentially loses detail of the participants' thinking regarding specific items in each NOS construct. Additionally, as we

did not include in our analysis the constructed responses of every participant for each NOS construct, the nuances of the participants' thinking were also somewhat lost. This loss was, to some extent, regained when we decided to include the McNemar tests of each individual item. While we are making assumptions about participant interpretation of each item, we found no utterances in participants' open-ended responses to indicate problematic interpretations.

Results and Findings

Kuder-Richardson 20 Test

Cronbach's alpha of the overall instrument was found to be 0.788. The acceptable value of Cronbach's alpha for our study is any value near or greater than 0.7 (Nunnally & Bernstein, 1967). The results of the KR-20 tests of each individual construct are found in Table 1. As described in the next section, paired-samples *t*-tests were done on only the constructs found to have an acceptable Cronbach's Alpha value.

Paired-Samples *t*-Test

All paired-samples *t*-tests were statistically significant ($p < .05$) when looking at differences between the pre- and post-measures of the NOS concepts. The paired-samples *t*-test analyzing differences in the participants' understanding of all assessed NOS concepts combined was also found to be statistically significant. To determine effect size, we calculated Eta squared values and used Cohen (1988): small = .01, moderate = .06, large = .14. For each of the significant paired-samples *t*-tests, participants on average scored higher on the post-measure than on the pre-measure. Table 2 shows the results of the pair-sample *t*-tests and effect sizes for all comparisons.

McNemar's Test

McNemar's tests were conducted on each of the 16 items of the constructs of the SUSSI that had acceptable Cronbach's Alpha value of greater than or near 0.7 (Nunnally & Bernstein, 1967). These McNemar's tests were used to determine if the proportion of participants

Table 1 Kuder-Richardson 20 (KR-20) tests—summary of results for pre-/post-measures of nature of science concepts ($n = 60$)

NOS construct	Cronbach's alpha value
Prior Knowledge	.442
Tentative	.437
Laws and Theories	.161
Social and Cultural	.701
Collaborative	.718
Creative and Imagination	.857
Scientific Method	.680
Supernatural	.257
Overall NOS	.788

Table 2 Paired-samples *t*-tests—summary of results for pre-/post-measures of nature of science concepts (*n* = 60)

	Pre-measure		Post-measure		<i>t</i>	<i>df</i>	<i>p</i>	Eta statistic	
	<i>M</i>	SD	<i>M</i>	SD				Value	Effect size
Social and Cultural	2.700	.357	3.517	.930	3.928	59	< .001*	.207	Large
Collaborative	2.233	1.454	3.600	.785	6.597	59	< .001*	.425	Large
Creative and Imagination	1.2 67	1.561	2.833	1.520	6.163	59	< .001*	.392	Large
Scientific Method	1.783	1.316	3.233	1.100	7.542	59	< .001*	.491	Large
Overall NOS	16.317	.986	23.017	4.355	9.898	59	< .001*	.628	Large

* *p* < .05

demonstrating accurate views for each item of the SUSSI instrument was significantly different from pre- to post-test. Fifteen of the 16 McNemar's tests were statistically significant. Table 3 demonstrates the results of McNemar analysis for each SUSSI item tested. The percentages were rounded to the nearest whole percent, therefore may not equal 100%.

Discussion and Implications

Review of Findings

Participating inservice elementary teachers' views of assessed NOS concepts were less aligned to literature at the beginning of the study. Based on the results of our *t*-tests, overall NOS understanding and each of the four NOS concepts demonstrated statistically significant growth from the pre- to post-assessment. While these results align well to existing literature, our analysis of individual items provides for more detailed analysis of NOS learning as described below.

Details of Changes in Participants' Views of Specific NOS Constructs

Other studies (e.g., Abd-El-Khalick & Akerson, 2009; Akerson et al., 2000; Akerson & Abd-El-Khalick, 2003; Akerson, Cullen, & Hanson, 2009; Akerson, Weiland, Rogers, Pongsanon, & Bilican, 2014; Deniz & Akerson, 2013) provide fine-grained understanding of participants' thinking through qualitative data collection. These studies use representative quotes that, by their very nature, make fine-grained comparisons across studies difficult. Given the nature of the SUSSI instrument, and our relatively large sample size in comparison to qualitative studies, we hope to shed light on more fine-grained details of NOS learning that could be compared across studies through quantitative assessment. Below, we discuss nuanced changes in our participants' understanding and related literature.

Social and Cultural Supporting the work of Bell, Mulvey, & Maeng (2016), Herman & Clough (2016) and Akerson et al. (2000), the participants in our study recognized and increased their understanding of cultural impact on what and how science is done, even after starting out with many participants beginning the PD program with already aligned views.

Based on our data, the social and cultural influences on science seem to be fairly intuitive for our participants. This finding differs from the results of Mesci and Schwartz (2017) who noted their participants struggled to grow in their understanding of how culture influences

Table 3 Summary of results for McNemar tests of nature of science concepts ($n = 60$)

# of Likert question	Likert item	Not aligned at pre and post	Aligned at pre and post	Changed from not aligned to aligned	Changed from aligned to not aligned	<i>p</i> -value
Social and Cultural						
1	Scientific research is not influenced by society and culture because scientists are trained to conduct pure, unbiased studies.	6 (10%)	31 (52%)	20 (33%)	3 (5%)	< .001*
2	Cultural values and expectations influence what science is conducted and accepted.	1 (2%)	42 (70%)	12 (20%)	3 (5%)	.013*
3	Cultural values and expectations influence how science is conducted and accepted.	3 (5%)	39 (65%)	14 (23%)	4 (7%)	.031*
4	All cultures conduct scientific research the same way because science is universal and independent of society and culture.	1 (2%)	31 (52%)	19 (32%)	8 (13%)	.052
Collaborative						
5	Scientists usually work collaboratively with other scientists when conducting research.	2 (3%)	31 (52%)	27 (45%)	0 (0%)	< .001*
6	Scientists usually work with other scientists, but only to share results.	6 (10%)	22 (37%)	27 (45%)	5 (8%)	< .001*
7	Scientists usually work alone when conducting research.	2 (3%)	32 (53%)	22 (37%)	4 (7%)	.001*
8	Scientific knowledge usually emerges from discussions and social interactions among scientists.	4 (7%)	39 (65%)	16 (27%)	1 (2%)	< .001*
Creativity and Imagination						
9	Scientists use their imagination and creativity when they collect data.	12 (20%)	11 (18%)	32 (53%)	5 (8%)	< .001*
10	Scientists use their imagination and creativity when they analyze and interpret data.	16 (27%)	13 (22%)	28 (47%)	3 (5%)	< .001*
11		12 (20%)	19 (32%)	26 (43%)	3 (5%)	< .001*

Table 3 (continued)

# of Likert question	Likert item	Not aligned at pre and post	Aligned at pre and post	Changed from not aligned to aligned	Changed from aligned to not aligned	<i>p</i> -value
12	Scientists do not use their imagination and creativity because these conflict with their logical reasoning. Scientists do not use their imagination and creativity because these can interfere with the need to be unbiased.	12 (20%)	15 (25%)	26 (43%)	7 (12%)	< .001*
Scientific Method						
13	Scientists use different types of methods to conduct scientific investigations.	0 (0%)	43 (72%)	14 (23%)	3 (5%)	.013*
14	Scientists follow the same step-by-step scientific method.	5 (8%)	22 (37%)	32 (53%)	1 (2%)	< .001*
15	When scientists use the scientific method correctly, their results are true and accurate.	17 (28%)	12 (20%)	29 (48%)	2 (3%)	< .001*
16	When investigating the natural world, experiments are the only way scientists develop good scientific knowledge.	16 (27%)	22 (37%)	20 (33%)	2 (3%)	< .001*

science. Despite having a high proportion of participants that were aligned at both the pre- and post-test, we still saw three of the four items (items 1, 2, and 3) make significant gains because nearly all remaining participants developed accurate views. Our study involved teachers from an urban district in which cultural influence on learning has been a part of their district's PD for many years. Perhaps teachers were primed to recognize the ways in which culture influences thinking and making connections between teacher cultural bias and scientific cultural bias could be a mechanism to support teachers in learning this NOS construct.

Collaborative Herman and Clough (2016) note that teachers need to acquire an expanded view of the ways in which scientists are collaborative. Specific parts of scientists' work that Herman and Clough (2016) defined as done collaboratively include peer review and scientists' communication influencing their colleagues' thinking. Both of these areas of collaboration are included in the SUSSI (items 5–8). Participants in our study moved beyond the specific misconception that scientists are only collaborative when sharing results. As with the cultural influences construct, we observed a high proportion of participants begin and remain aligned from pre- to post-test. Where we saw the most growth in the proportion of participants was on items 5 and 6 with 45% improving their understanding of both, which describe how scientists work together beyond just sharing results.

Creativity and Imagination Typically, the creative NOS seems to be one of the easier NOS concepts to learn. Multiple studies have found participants improve their understanding of the role of creativity and imagination in science (Akerson et al., 2000; Bell, Matkins, & Gansneder, 2011; Bell et al., 2016; Donnelly & Argyle, 2011; Herman & Clough, 2016; Ward & Haigh, 2017). Donnelly and Argyle (2011) even found the majority of their participants to be aligned at the pre-assessment (27/31) and again at the post-assessment (34/35) with participants noting the use of creativity to design new methods of testing as well as interpreting data. Our study differed from Donnelly and Argyle (2011) in that we did not necessarily start with the majority of the participants aligned to more accurate thinking, but rather saw much growth from the participants across each item (9–12) targeting the creative NOS. Although qualitative approaches are often more sensitive than quantitative approaches, our study seems to have demonstrated growth where Donnelly and Argyle (2011) did not.

Scientific Method Bell et al. (2016), Herman and Clough (2016), and Mesci and Schwartz (2017) all describe general growth in many of their participants' thinking regarding the scientific method. More specifically, participants tended to move from believing science follows a step-by-step scientific method to accurately understanding scientists use many empirical methods in their work (e.g., item 15). Seventy-two percent of participants in our study began the program believing there are many methods of doing science (item 13), yet tended to believe the steps of the scientific method must be followed as prescribed (item 14). Given that the participants acknowledged scientists may use a variety of methods (item 13), teaching them to consider how there is no single scientific method (item 14) was fairly easy. However, these participants still held tight to their misconception of experiments (item 16). Teacher educators may want to consider starting with the seemingly intuitive idea that scientists use many different strategies (item 13), then move toward helping teachers understand that scientists do not have to follow step-by-step procedures (item 14), and finally go further to help teachers differentiate between experimental versus observational science (item 16).

Final Thoughts and Future Work

Our results demonstrate that participants did grow in their understanding of NOS constructs as a result of the year-long PD. This finding is clearly demonstrated by the overall and construct-specific *t*-tests and large effect sizes. Using McNemar tests, we were able to gain a deeper understanding of the nature of the changes in participant views.

While some studies with fewer participants (e.g., Akerson et al., 2009; Akerson & Abd-El-Khalick, 2003; Herman & Clough, 2016) are able to provide deep analysis of individual participant thinking, the insights drawn from this larger study may provide teacher educators with more transferable, yet still nuanced, insight for helping teachers develop more informed NOS views. Perhaps, even larger sample size studies could generate additional insights. However, given the SUSSI only had four constructs with acceptable Cronbach's Alpha levels, if more quantitative work is to be done, additional development must be undertaken. That is, based on the limited scope of the SUSSI constructs with acceptable reliability, we do not recommend its use as a quantitative instrument beyond those four constructs. Therefore, the search continues for a strong quantitative NOS instrument that would allow for robust large-scale studies. We believe the recently published RFN Questionnaire (Kaya et al., 2019) holds promise. Yet, before large-scale studies could be carried out and compared, the instrument should be subjected to reliability testing at the construct level as we have done here with the SUSSI or perhaps undergo factor analysis.

References

- Abd-El-Khalick, F., & Akerson, V. L. (2004). Learning as conceptual change: factors mediating the development of preservice elementary teachers' views of nature of science. *Science Education*, 88(5), 785–810.
- Abd-El-Khalick, F., & Akerson, V. L. (2009). The influence of metacognitive training on preservice elementary teachers' conceptions of nature of science. *International Journal of Science Education*, 31(16), 2161–2184.
- Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: a critical review of the literature. *International Journal of Science Education*, 22(7), 665–701.
- Akerson, V. L., & Abd-El-Khalick, F. (2003). Teaching elements of nature of science: a yearlong case study of a fourth-grade teacher. *Journal of Research in Science Teaching*, 40(10), 1025–1049.
- Akerson, V. L., & Abd-El-Khalick, F. S. (2005). "How should I know what scientists do?—I am just a kid": fourth-grade students' conceptions of nature of science. *Journal of Elementary Science Education*, 17(1), 1–11.
- Akerson, V. L., Abd-El-Khalick, F., & Lederman, N. G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching*, 37(4), 295–317.
- Akerson, V. L., Cullen, T. A., & Hanson, D. L. (2009). Fostering a community of practice through a professional development program to improve elementary teachers' views of nature of science and teaching practice. *Journal of Research in Science Teaching*, 46(10), 1090–1113.
- Akerson, V. L., & Hanuscin, D. L. (2007). Teaching nature of science through inquiry: results of a 3-year professional development program. *Journal of Research in Science Teaching*, 44(5), 653–680.
- Akerson, V. L., Morrison, J. A., & McDuffie, A. R. (2006). One course is not enough: preservice elementary teachers' retention of improved views of nature of science. *Journal of Research in Science Teaching*, 43(2), 194–213.
- Akerson, V. L., Weiland, I., Rogers, M. P., Pongsanon, K., & Bilican, K. (2014). Exploring elementary science methods course contexts to improve preservice teachers' NOS of science conceptions and understandings of NOS teaching strategies. *EURASIA Journal of Mathematics, Science & Technology Education*, 10(6), 647–665.
- Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education*, 95(3), 518–542.

- American Association for the Advancement of Sciences [AAAS]. (1989). *Science for all Americans: Project 2061*. New York, NY: Oxford University Press.
- Bell, R. L., Matkins, J. J., & Gansnedler, B. M. (2011). Impacts of contextual and explicit instruction on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching*, 48(4), 414–436.
- Bell, R. L., Mulvey, B. K., & Maeng, J. L. (2016). Outcomes of nature of science instruction along a context continuum: preservice secondary science teachers' conceptions and instructional intentions. *International Journal of Science Education*, 38(3), 493–520.
- Capps, D., & Crawford, B. (2013). Inquiry-based instruction and teaching about nature of science: are they happening? *Journal of Science Teacher Education*, 24(3), 497–526.
- Clough, M. P. (2006). Learners' responses to the demands of conceptual change: considerations for effective nature of science instruction. *Science & Education*, 15(5), 463–494.
- Clough, M. P. (2007). Teaching the nature of science to secondary and post-secondary students: questions rather than tenets. *Pantaneito Forum*, 25.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cofré, H., Núñez, P., Santibáñez, D., Pavez, J. M., Valencia, M., & Vergara, C. (2019). A critical review of students' and teachers' understandings of nature of science. *Science & Education*, 28, 205–248.
- Dagher, Z. R., & Erduran, S. (2016). Reconceptualizing the nature of science for science education. *Science & Education*, 25(1–2), 147–164.
- Deniz, H., & Akerson, V. L. (2013). Examining the impact of a professional development program on elementary teachers' views of nature of science and nature of scientific inquiry, and science teaching efficacy beliefs. *Electronic Journal of Science Education*, 17(3), 1–19.
- Donnelly, L., & Argyle, S. (2011). Teachers' willingness to adopt nature of science activities following a physical science professional development. *Journal of Science Teacher Education*, 22(6), 475–490.
- Herman, B., & Clough, M. P. (2016). Teachers' longitudinal NOS understanding after having completed a science teacher education program. *International Journal of Science & Mathematics Education*, 14, 207–227.
- Herman, B. C., Clough, M. P., & Olson, J. K. (2013). Teachers' nature of science implementation practices 2-5 years after having completed an intensive science education program. *Science Education*, 97(2), 271–309.
- Holliday, G., Lederman, J., & Lederman, N. (2014). “Wow! Look at that!”: discourse as a means to improve teachers' science content learning in informal science institutions. *Journal of Science Teacher Education*, 25(8), 935–952.
- Irzik, G., & Nola, R. (2011). A family resemblance approach to the nature of science for science education. *Science & Education*, 20(7–8), 591–607.
- Irzik, G., & Nola, R. (2014). New directions for nature of science research. In *International handbook of research in history, philosophy and science teaching* (pp. 999–1021). Dordrecht: Springer.
- Kaya, S. (2012). An examination of elementary and early childhood pre-service teachers' nature of science views. *Procedia – Social and Behavioral Sciences*, 46, 581–585.
- Kaya, E., Erduran, S., Aksoz, B., & Akgun, S. (2019). Reconceptualised family resemblance approach to nature of science in pre-service science teacher education. *International Journal of Science Education*, 41(1), 21–47.
- Kruse, J., Kent-Schneider, I., Voss, S., Zacharski, K., & Rockefeller, M. (2020). Investigating the effect of NOS question type on students' NOS responses. *Research in Science Education*, 1–18.
- Lederman, N. G. (1992). Students and teachers' conceptions of the nature of science. A review of the research. *Journal of Research in Science Teaching*, 29(4), 331–359.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521.
- Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M., & Ebenezer, J. (2008). Assessing preservice elementary teachers views on the nature of scientific knowledge. A dual response instrument. *Asia-Pacific Forum on Science Learning and Teaching*, 9(1), 1–20.
- McComas, W. F. (2004). Keys to teaching the nature of science. *Science Teacher*, 71(9), 24–27.
- Mesci, G., & Schwartz, R. (2017). Changing preservice science teachers' views of nature of science: why some conceptions may be more easily altered than others. *Research in Science Education*, 47(2), 329–351.
- National Research Council [NRC]. (1996). *National science education standards*. Washington, DC: The National Academies Press.
- NGSS Lead States. (2013). *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press.
- Nunnally, J., & Bernstein, I. H. (1967). *Psychometric theory*. New York: McGraw-Hill Higher, Inc..

- Shim, M. K., Young, B. J., & Paolucci, J. (2010). Elementary teachers' views on the nature of scientific knowledge: a comparison of inservice and preservice teachers. *Electronic Journal of Science Education, 14*(2).
- Ward, G. G., & Haigh, M. (2017). Challenges and changes: developing teachers' and initial teacher education students' understandings of the nature of science. *Research in Science Education, 47*(6), 1233–1254.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.