

Using Artifact Methodology to Compare Learning Assistants' and Colleagues' Classroom Practices

Stephanie A. Barr, Mike J. Ross, and Valerie Otero

School of Education University of Colorado – Boulder, 249 UCB, Boulder, CO 80309-0249

Abstract. The University of Colorado's LA-Test K-12 research team investigated the classroom practices of former Learning Assistants' who went on to become K-12 teachers. One of the tools used for this analysis of classroom practice was the Scoop Notebook, an instructional artifact package developed to assess teachers' use of reform-oriented practices. In this paper, the authors characterize differences in classroom practices between former Learning Assistants teaching at the secondary level and their colleagues through the collection and analysis of teaching artifacts. Analyses of these artifacts indicate significant differences between LA and non-LA groups. A description of the methodology and implications of using artifact packages to study classroom practice will be discussed, detailing the role of the LA experience in teacher preparation.

Keywords: reformed teaching, learning assistants, teacher preparation, artifact methodology
PACS: 01.40.Fk, 01.40.G, 01.40.jc

INTRODUCTION

The national crisis in physics education is well documented [1]. Only 35% of those teaching physics in the U.S. have a major in physics or physics education [2]. Moreover, few universities graduate more than two physics teachers per year [3]. The NSF-supported Colorado Learning Assistant (LA) model has addressed this crisis by recruiting talented physics majors to become pre-college physics teachers. The program has more than tripled the number of physics teachers graduating from CU Boulder each year (from an average of 0.5 to 2.5 per year). The LA-Test research project was designed to study effects of the LA program on these teachers' practice in K-12 schools. Due to the small number of physics teachers graduating from CU Boulder, we focus this investigation on all STEM disciplines to glean insight into the effects of the LA program on physics teachers.

The overarching goal of the LA-Test K-12 research team is to conduct a quasi-experimental study comparing the practices secondary teachers completing the same teacher certification program with and without the LA experience [4]. Several methods were used to investigate differences between former LAs and a matched sample of K-12 teachers. The study reported here focuses on the data collection and analysis associated with the "Scoop Notebook."

THE SCOOP NOTEBOOK

There have been several methods used to characterize classroom practice. The two most commonly used are case studies and surveys. Each

method presents advantages and limitations and can be used to serve a range of research questions.

Case studies involve hours of classroom observations in order to build very rich and detailed descriptions of classroom practices. They can offer insight into the finer grained and more nuanced aspects of the complex phenomena that constitute teachers' practices. However, classroom observations, and hence case studies, entail laborious demands that make performing a large number of them very time and resource intensive. The resulting sample size issue makes case studies difficult, if not impossible, to generalize to other contexts. Surveys are much less expensive to administer and can easily be used with large samples. They are limited, however, in the kind of information that can be obtained. Typically, they can only assess self-reported information and cannot measure what actually happens in practice.

The Scoop Notebook is an artifact package designed to "scoop" up information from a teacher's classroom to be analyzed in the research lab. While the data are not as rich as classroom observations or case studies, they provide much more substantive, context sensitive information than surveys. Artifact packages, like the Scoop Notebook, decrease the amount of time and resources needed to collect the data while offering a level of detail that is somewhere between case studies and surveys.

The Scoop was developed by Borko and colleagues to document the relationship between artifacts of classroom instruction and the nature of reform classroom practice. According to the designers, the "Scoop Notebook" [is a] one week process in which teachers collect artifacts of instructional practice (e.g., lesson plans, instructional materials, student work),

take photographs of the classroom setup and learning materials, write responses to reflective questions, and assemble the results in a three-ring binder” [5]. A pilot study for this artifact package was conducted in 2005 with Borko et al. [5], who found that the Scoop Notebook is an effective means by which to measure some but not all aspects of classroom practice. Our study is the first large scale application of this methodology.

To support the analysis of the Scoop Notebooks, Borko and colleagues [see 6] developed rating guides for math and science classes using ten dimensions of reform-oriented practice. Since some of the original dimensions apply only to either math or science or vary in name and description, we chose to disregard four categories and to collapse analogous categories for analysis (see Table 1).

DATA AND METHODOLOGY

The LA-Test K-12 research team collected Scoop Notebooks from participating teachers during each year of the study. Per Borko and colleagues’ guidelines, teachers selected a representative week of instruction for the Scoop Notebook, beginning the “Scooping” process at the beginning of the unit [6].

All Scoop Notebook artifacts were analyzed by two raters according to rubrics detailed in the *Quick Reference Guide for CRESST SCOOP Rating*. Each dimension is rated on a five-point scale, *low* (1), *medium* (3), and *high* (5). The associated scoring rubric contains both detailed and general descriptions of each level and specific classroom examples of *high*, *medium*, and *low* practice. Each scorer of a two person team scored Scoop Notebooks independently and then recorded the highest score observed based upon all included artifacts. Differences between scores for an individual notebook were negotiated to achieve a consensus set of scores.

The nineteen participating teachers submitted a Scoop Notebook included eleven former LAs and eight non-LAs. All participants were classroom teachers in middle and high schools. In addition to former LAs who completed a teacher certification program in secondary math or science, a comparison group was established by recruiting teachers who completed the same teacher certification program, but did not participate in the LA program. Teachers were matched across groups according to teaching content area (physical science, life science, or math), teacher preparation program (licensure only or licensure with master’s), and school context (urban, suburban, or rural). This matched sample served as a “control” for the overall study in order to determine if there were substantial differences between the teaching practices

TABLE 1. Collapsed (Borko et.al) Scoop Dimensions

Borko et.al Science Dimensions	Borko et.al Mathematics Dimensions	LA-Test Study Dimensions
Collaborative Grouping	Collaborative Grouping	Collaborative Grouping
Structure of Lessons	Structure of Lessons	Structure of Lessons
Cognitive Depth	Cognitive Depth	Cognitive Depth
Scientific Discourse Community	Mathematical Discourse Community	Math/Sci. Discourse Community
Use of Scientific Resources	Use of Mathematical Tools	Use of Math Tools/Sci. Resources
Assessment	Assessment	Assessment
Explanation/ Justification	Explanation/ Justification	Explanation/ Justification
Connections/ Applications	Connections/ Applications	Connections/ Applications
Hands-on Inquiry		
	Problem solving	
	Multiple Representations	

of LAs and non-LAs.

Each Scoop package was treated as an individual unit of analysis. We collected two Scoop packages (over three years) from ten of the participants, three from one participant and one from the remaining eight. All together thirty-seven notebooks were collected and analyzed (see Table 2). The scores were analyzed by calculating the mean scores along each dimension in order to make comparisons among subgroups.

TABLE 2. Number of Scoop Notebooks Analyzed

	<u>LA</u>	<u>non-LA</u>
Science	10	4
Math	6	9

RESULTS

We analyzed the Scoop Notebooks along the various dimensions to determine any differences between LAs and non-LAs mean scores with respect to reform-oriented classroom practices. Results indicate some notable differences across the LA and non-LA groups for the aggregated math and science teachers. As seen in Figure 1, LAs mean scores exceeded those of non-LAs in seven of the eight

categories, outperforming their colleagues at statistically significant levels in the following dimensions: Collaborative Grouping, Discourse Community, and Explanation/Justification ($p < 0.02$).

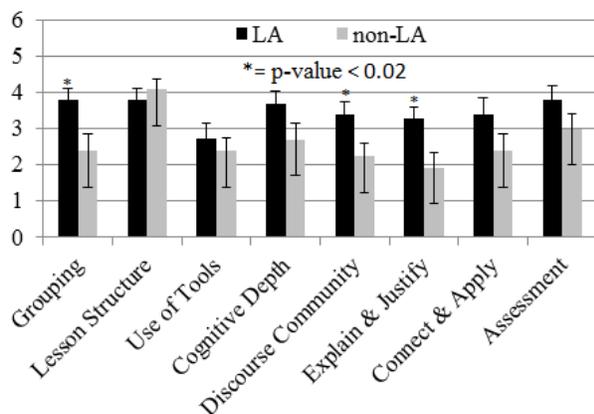


FIGURE 1. Scoop Mean Scores for All Teachers: LA (n=14) and non-LA (n=15).

Comparisons of the science Scoop Notebooks, shown in Figure 2, revealed no statistically significant differences. However, the mean scores for former LAs, now teaching science, exceeded that of their non-LA counterparts in five of the eight dimensions: Collaborative Grouping, Mathematical/Scientific Discourse Community, Explanation/Justification, Connections/Applications, and Assessment.

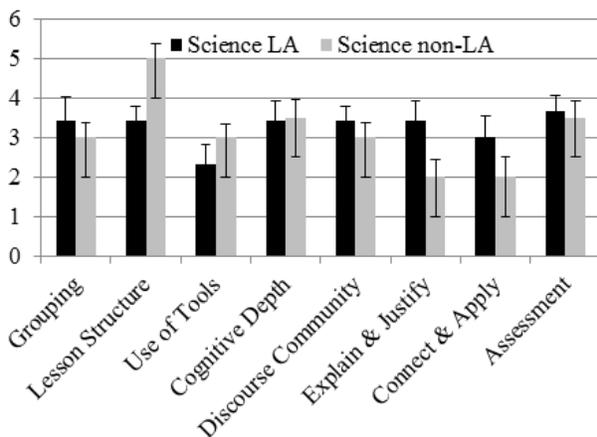


FIGURE 2. Mean Scores for Science Teachers: LA (n=10) and non-LA (n=4).

In Figure 3, the comparison of math LAs scores to math non-LAs also shows them outscoring their counterparts in every dimension. However, these differences in means are only statistically significant in the Collaborative Grouping ($p < 0.002$), Mathematical/Scientific Discourse Community ($p < 0.02$), and Explanation/Justification dimensions ($p < 0.05$). The

Cognitive Depth, Connections/Applications, and Assessment dimensions, though not statistically significant, differed beyond the standard errors.

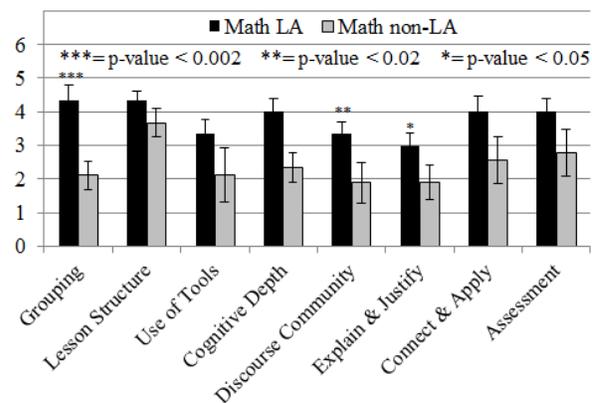


FIGURE 3. Scoop Mean Scores for Math Teachers: LA (n=6) and non-LA (n=9).

DISCUSSION

Overall, the aggregated data show that LA mean scores were consistently higher than non-LA mean scores in seven of the eight categories. Of these categories, the three statistically significant dimensions can be related directly to common aspects of the LA experience. Collaborative Grouping is a primary learning goal of the LA seminar as well as a principle means of instructional organization for many, if not all, of the LA run undergraduate sessions. Discourse Community and Explanation/Justification are also a very large part of the LA program. The program is designed such that LAs develop pedagogical content knowledge through their work with content-specific learning teams. LAs also meet weekly with the lead instructor of the course to review the content, anticipate questions, and analyze assessment data. This improved content knowledge has also been shown in comparisons of LAs to other undergraduates [7].

The LA experience also emphasizes teaching for deeper conceptual understanding and the practice of challenging student assertions by requiring explanation and justification. This experience may very well be related to the higher LA means for Explanation/Justification and for Discourse Community in addition to the significance in differences between math LA and non-LA mean scores for Cognitive Depth. One may also speculate that the social and collaborative nature of the LA course as well as the LA run undergraduate sessions may have impacted the scores for Discourse Community as well. Finally, the differences in the Assessment category scores favoring

LAs may be attributable in part to the emphasis on formative assessment in the LA seminar. LAs are educated on the theoretical and practical bases for models of formative assessment, explicitly encouraged to informally assess student understanding in a variety of ways and to use that information to help move students toward greater understanding.

A more interesting aspect of our findings, however, is that there are no statistically significant differences for science LAs. This question is even more remarkable in light of the following two outcomes: (1) Gray and colleagues [4] findings which show former science LAs out scoring their colleagues on the Reformed Teacher Observation Protocol (RTOP) and (2) statistically significant findings in the LA/non-LA and math subgroup comparisons in this paper. Most likely this discrepancy is attributed to the small number of Scoop Notebooks that have been scored compared to the RTOP data which consists of 3 times as many observations as Scoop Notebooks. There are, however, some aspects of the CU LA program, as well as the kind of reporting required for the Scoop artifact package that may explain the lack of statistically significant differences in the science category.

The math LAs that participated in this study came from a more unified program that was directed and run by a single instructor who holds a PhD in math education. She has an extensive background in secondary education, including five years teaching high school math and science including in an inner city school. Since this is a standard methodology of assessment in teacher education circles, we believe that this instructor may have provided more mentoring (intentionally or not) in the preparation and presentation of materials in a portfolio fashion. As was noted earlier, significant differences were found for science teachers in classroom observations [4]. Therefore, we believe that there also may be something in the Scoop reporting method, or in formal lesson plan and reflection presentation that can account for this discrepancy.

The experience for science LAs was markedly different from that of the math LAs who participated in this study. In science, five departments participate in the LA seminars with over ten different classes and there is a lack of standardization with regard to who is directly in contact with the science LAs. Many of the science LA programs are run by professors with little to no experience in K-12 or discipline-based education research (DBER). The instructors teaching the LA supported courses may or may not be familiar with findings from PER and related fields. Instead, the focus of the physics LA training centers around training in the use of the UW tutorials [8] and other methods that prepares them to be successful in collaborative settings.

These diverse experiences amongst CU LA participants may affect these teachers' understanding of how to provide as detailed reflections and lesson plans as their math counterparts. Finally, there are limitations to the data collected in the Scoop Notebook. Scorers are bound by the artifacts and reflections provided by the participants, so it is difficult to tell whether a lack of experience with writing reflections or presenting information need for the artifact package in a way the truly represents their reform classroom practices.

Of course, the differences for any of these LA and non-LA Scoop scores are not in themselves the basis for causal inferences about the impact of the LA program, but they do offer considerable evidence, particularly when taken with other findings from the larger study, that the LA program does impact the practice of teachers along a significant number of reform-oriented dimensions.

ACKNOWLEDGMENTS

We thank the LA-Test K-12 team for their contributions to the collection and scoring of the Scoop data. This work is partially funded by NSF grant # ESI – 00554616.

REFERENCES

1. <http://www.phystec.org/PhysicsToday.pdf>
2. <http://www.aip.org/statistics>
3. T. Hodapp, "National Task Force on Teacher Education in Physics" 2010, SMTI National Conference, Cincinnati, OH.
4. K. Gray, D. Webb, and V. Otero, "Are Learning Assistants Better K-12 Science Teachers?" 2009, edited by C. Henderson et al., AIP Press, Physics Education Research Conference Proceedings. Melville, NY: AIP Press. p. 157
5. H. Borko, B. Stecher, A. Alonzo, S. Moncure, S. McClam, Educational Assessment, *10*(2), 73- 104 (2005).
6. Borko, H., Stecher, B., and Kuffner, K. (2007). Using Artifacts to Characterize Reform-Oriented Instruction: The Scoop Notebook and Rating Guide. CSE Technical Report 707. Los Angeles: The Regents of the University of California.
7. V. Otero, S. Pollock, and N. Finkelstein, American Journal of Physics, *78*(11), 1218-1224 (2010).
8. L. McDermott, P. Shaffer, and the Physics Education Group, *Tutorials in Introductory Physics* _Prentice-Hall, Saddle River, NJ, 2002.