

HISTORICAL EVIDENCE AND EVIDENCE OF LEARNING

SAM WINEBURG, AVISHAG REISMAN, and BRADLEY FOGO

In his magisterial *Idea of History*, R. G. Collingwood likened historical understanding to “a web of imaginative construction.” Yet, unlike the poet or novelist, whose imaginations soar in boundless flights of fancy, the historian must never wander far from the “fixed points” of historical evidence:

The historian’s picture of his subject, whether that subject be a sequence of events or a past state of things, thus appears as a web of imaginative construction stretched between certain fixed points provided by the statement of his authorities; and if these points are frequent enough and the thread spun from each to the next are constructed with due care . . . the whole picture is constantly verified by appeal to these data, and runs little risk of losing touch with the reality which it represents.¹

For Collingwood, it is fidelity to evidence that keeps the historian honest, the constant connection to the gritty reality of the documentary record. The historian, to be sure, must possess a fecund imagination, but creativity, alas, has limits. The historical story is tethered to evidentiary warrants that ground it and give it shape. As readers of history, we, in fact, judge historical accounts by their fidelity to these bits of “reality,” and we deploy (implicitly, if not explicitly) a set of criteria — coherence, plausibility, verisimilitude — to arrive at conclusions about the believability of a given historical account.

In honing our judgments about historical believability, we rely on the extensive body of historiographic writings by historians and philosophers of history.² But where do we turn for guidance when our topic is not the epistemological status of historical narratives, but the truth claims about what young people have *learned* from these narratives? If we use notions of coherence, plausibility, and verisimilitude to judge historical accounts, what tools do we use to assess claims by

education researchers that learning has indeed occurred? Here, the body of historiographic writings provides less aid. Like all learning researchers, the history education researcher is an applied social scientist and must ply and be judged by the criteria that define that field.

Social scientists do not find data by searching the documentary record but create data by venturing into classrooms and capturing what transpires in them by assigning tests, distributing questionnaires, collecting assignments, arranging interviews, recording observations, and so on. In this sense, research on history education appeals less to notions of “imaginative construction” when making claims about learning than to the traditional criteria of the applied social scientist: reliability, validity, accuracy, parsimony, and rigor in the handling of empirical data.

The articles that appear in this special issue attest to the range of topics and treatments that characterize history education today. By browsing these articles – and the extensive literature they build on—it is apparent that they represents a mere sliver of an energetic and burgeoning field. It is hard to imagine that only twenty-five years ago, one could have attended the annual convention of the American Educational Research Association and searched in vain for a single session on history education. The collection before us testifies to a growing and flourishing field of research.

It is precisely for this reason – the rapid growth in the number and kinds of researchers engaged in history education – that it is useful to pause and take stock. Each of the articles here contributes to our understanding of the issues and possibilities of history education across a range of levels — from the elementary school classroom to pre-service teacher education, all the way up to the continuing professional education of experienced teachers. But we would be remiss as commentators if we sufficed with a congratulatory pat on the back. To continue to make progress, we must ask questions that lead to greater self-awareness about the claims we make as applied social scientists. By asking about method and warrant, it is our hope to spur both the authors of these papers and the field as a whole to think harder about what we mean when we claim to have observed learning in these different settings.

Jane Bolgatz of Fordham University takes us into an urban classroom characterized by a high concentration of poverty (98 percent of students qualify for free lunch). These students, almost entirely black and Hispanic, earn failing scores on district and state tests. It is a

school, at first glance, that we think we know well. But rather than presenting yet another desultory image of sullen students struggling to stay awake with endless worksheets and pilfered copies of district tests, we find, instead, a refreshing ray of light: a fourth-grade teacher who engages her young charges in an innovative curriculum in which questions of historical evidence, perspective, causality, and narrative are not only present, but central.

Bolgatz's study, drawing on anthropological methods of participant observation and discourse analysis, falls into the tradition of "wisdom of practice," in which rich examinations and documentations of "wise practitioners" become part of the professional research base for teaching.³ The teacher in this study, Ms. Agosto, does not shy away from the tough questions asked by her fourth-graders – such as when they asked her about how one actually knows whether the bones archeologists have unearthed from a Revolutionary-era graveyard come from blacks or whites. Rather than stifling her students' curiosity, Agosto joins them as a fellow historical enquirer: "I don't know if I really know the answer to that," she tells them. "Let's come up with some ideas."

Using excerpts from field-notes, brief quotes from classroom exchanges, and her own powers of analysis, Bolgatz limns a picture of a classroom in which young historians – eight- and nine-year-olds – take on issues well beyond the stereotyped images of the "expanding environments" curriculum of the elementary school. And, at first blush, this is truly a place in which extraordinary things occur. Toward the end of her article, Bolgatz states: "Data from this study demonstrated that *all students* can practice solid historical thinking skills." [emphasis added]

This claim places us at the heart of a dilemma. On one hand, reading Bolgatz's descriptions, we clearly are in the hands of a talented teacher who, with a rare combination of subject matter knowledge, creativity, and skill, creates a zestful classroom environment. On the other hand, as fellow researchers, we are compelled to ask questions about the empirical warrant for Bolgatz's claims. On what basis can she claim that "all students" demonstrated "solid historical thinking skills"? Of the twenty-five students in this classroom, there are many who remain silent in this research report, and many who appear to utter only a few sentences. Even those who do speak do not unambiguously demonstrate "solid historical thinking skills," nor is what the author means by this term sufficiently spelled out for us to judge. Ample

research has shown that classroom speech is notoriously difficult to interpret as a lone index of learning and must be supported (“triangulated”) by different and varied data sources.⁴ We passionately hope that the author is correct in her assessment about the abilities and growth in learning by these students, but the *demonstration* of that growth awaits a more comprehensive follow-up than what we are provided with here.

The following two articles confront the demands of the digital age and promote the use of visual media in history education. The authors, Bruce Fehn of the University of Iowa and David J. Staley of Ohio State University, believe multimedia projects afford students opportunities to wrestle with the interpretive nature of historical narrative and the persuasive power of visual images. Their case is compelling. One cannot deny the increasing presence of technology in the classroom, nor can one argue against the value of arming students with critical *visual* literacy skills in an age where they are constantly bombarded with images. Images tell stories, and multiple images tell multiple stories depending on how they are arranged. In short, the authors believe that the thoughtful implementation of visual media in the history classroom will help both students and teachers gain fluency in an increasingly ubiquitous cultural medium.

The authors’ embrace of technology and visual media is refreshing and bold. As Staley notes, “most historians see themselves as writers,” and history teachers are often reluctant to assign multimedia projects. Staley responds to teachers’ hesitation by designing a rubric or “heuristic” for evaluating visual projects. Drawing from the literature on museum exhibitions, Staley’s criteria for evaluating visual displays include the formulation of a central argument or “big idea,” and the careful consideration of the image sequence and word choice.

The heuristic serves as a useful starting point for teachers who have been visual-media-shy in the history classroom. Yet, other than describe the tool, Staley offers little pedagogical guidance for the teacher who hopes to incorporate visual media in the classroom. As it stands, his heuristic could assist graduate students viewing a museum exhibit, but fails to consider the scaffolding required to teach students to use visual media.⁵ Nor does Staley show how the careful sequencing of visual images constitutes, by itself, a demonstration of historical thinking. Certainly, images can tell a story, but historical claims require evidentiary warrants. Visual analogies are powerful, as Staley points out, because they can point out “similarities in the midst

of apparent difference.” But history is the study of context and insists on the significance of time and space. Students are all too quick to make glib comparisons between historical events: Vietnam is like Iraq, Japanese internment is like the Holocaust, Hitler is like Saddam Hussein. Staley seems to celebrate the potential for visual media to erect such facile comparisons — for example, the “connection” linking Abu Ghraib to lynching. However, the greatest challenge facing history teachers is how to draw students’ attention to the particular, not the general; to the warrant, not the claim. For Staley’s heuristic to have value in the history classroom, it must move evidence from the periphery to the center.

Bruce Fehn’s study brings visual media into the teacher preparation classroom. Fehn examines his own pedagogy and seeks to determine whether certain curricular interventions have helped his students create more interactive PowerPoint presentations. He discovers that if left to their own devices, students initially produce PowerPoint presentations “mimicking narrative structures featured in PowerPoint lectures frequently observed in classrooms and lecture halls.” The first round of student presentations contain few slides and recapitulate standard historical narratives in neat bullet points and generic story arcs. Following two interventions, however, Fehn finds that students’ presentations include many more slides, fewer words, and creative sequences that suggest alternative historical interpretations and narratives.

We question the relationship between Fehn’s measurement criteria and his students’ understanding of historical narrative. Fehn shows that students’ PowerPoints include the greatest number of slides when they are barred from using words. He also observes that a number of these longer slideshows contain “complex narrative structures,” where authors’ “‘experimented’ with juxtaposition, cluster, sequence, or anomalies.” Such a claim raises the question of whether a sequence of incongruous images can demonstrate students’ understanding of “complex narrative structures.” We remain uncertain about the utility of using isolated features as proxies of historical understanding.

Although good presentations may include fewer words and more slides, the inverse is not necessarily true: presentations with fewer words are not necessarily more effective. The effort to capture historical understanding in this way recalls attempts to use compositional connectives—thus, therefore, in light of the fact—as indicators of an essay’s coherence. True, good essays tend to contain these aspects.

But we can imagine insipid essays filled with such connectives as well. The same holds true for the formalisms in this study. Fehn is faced with what researchers call a problem of “construct validity.”⁶ The measures he uses to operationalize quality bear an ambiguous relationship to the construct he seeks to measure.

Finally, we offer a note of caution to both Fehn and Staley. We are in the midst of a literacy crisis. Students in our inner cities are failing to learn how to read.⁷ While we welcome the authors’ enthusiasm for the role of technology and multimedia productions into the history classroom, some teachers will be eager to embrace visual media for the simple reason that students have an easier time working with images than with words. We fear that a decreased emphasis on writing and reading will exacerbate the literacy gap between rich and poor — not ameliorate it.

Kevin Bolinger of Indiana State University and Wilson J. Warren of Western Michigan University explore the “apparent gulf between professionals’ advice and actual teachers’ practices.” Their study recognizes the imperviousness of teacher practice to the efforts of researchers and curricular reformers. As Larry Cuban captured in his apt metaphor in 1979, reform efforts affect teacher practice like a passing storm: “Hurricane winds sweep across the sea, tossing up 20-foot waves . . . while on an ocean floor there is unruffled calm.”⁸ For over a century, the authors argue, “best practices” in social studies have always included instruction that promotes higher-order thinking and inquiry-based learning, regardless of whether the promulgators were committed to citizenship education, “constructivist” methods, or standards-based instruction. Their study, however, asks whether methods that promote thinking and interpretation actually appear in classroom practice.

Using survey data, the researchers find a tension between the methods that teachers most commonly use and the methods they believe are most effective. In particular, the study finds that though teachers may value methods such as role play and debate, they are most likely to lecture. The authors conclude, “The results of this study suggest that among both elementary and secondary social studies teachers passive methods are used more frequently than active, and arguably, more authentic methods.”

A core assumption guides this study. The authors presume that method represents a teacher’s disciplinary understanding. Research does show that students learn best when “knowledge is acquired not

from the information communicated and memorized but from the information that students elaborate, question, and use.” But the dichotomy between “active” and “passive” *learning* does not map neatly onto *method*.⁹ According to the authors’ formula, lecture is passive; role play is active. However, we certainly have all attended provocative, interactive lectures, and participated in vacuous role plays. The authors privilege method at the expense of a careful examination of teachers’ subject matter and pedagogical content knowledge. There are many paths to wisdom in history teaching, and it is in the *execution* of method – not its mere practice – that holds the key to understanding expertise in teaching.¹⁰

The final two articles in this collection examine history education within a system of high stakes accountability. Timothy Kelly, Kevin Meuwissen, and Bruce VanSledright interrogate the content of the national and Virginia history standards, while Stephanie van Hover, David Hicks, and William Irwin explore the effects of Virginia’s high stakes history tests on beginning teachers’ “notions of historical thinking.” The papers focus primarily on the state of Virginia; however, these authors raise important questions that should be addressed in all states where content standards and high stakes tests demarcate the teaching and learning of history.

Van Hover, Hicks, and Irwin introduce us to seven history teachers who juggle the myriad demands of classroom teaching. These teachers demonstrate many characteristics of capable and enthusiastic young teachers — they aspire to develop students’ critical thinking skills, engage classrooms with a variety of instructional strategies, and inspire new interest in the study of history. The authors find that the state’s standards and tests force these teachers to “to cover an enormous amount of content at a rapid pace.” Indeed, every one of the teachers “perceived the time pressure imposed by the SOL tests as extreme,” and claimed that the tests “precluded them from covering topics in depth.” Drawing on interviews, observations, and classroom documents, the authors paint a disheartening picture of a high-stakes environment that saddles beginning teachers with yet another burden.

This study suggests that Virginia’s system of accountability inhibits teaching for historical thinking. Once again, however, we are confronted with the question of evidentiary warrant. First, the authors presumably believe that the seven participants are representative of beginning teachers in Virginia. The reader is left wondering at the reasoning behind the authors’ sampling: Why seven teachers, why not ten,

or six, or even one? How do the concerns of these teachers represent those of the greater pool of beginning teachers in Virginia? In small-scale qualitative research, sampling is “theoretical.”¹¹ Given that small samples preclude generalization to an entire population, it is incumbent on the researcher to specify the “theory” guiding the sampling scheme.

Second, we question the authors’ claim that “all seven beginning teachers elucidated rich and interesting conceptions of history and historical thinking.” The evidence backing this claim consists of short quotes from six of the seven teachers. In the words of one of these teachers, historical thinking means “the skills associated with history – analyzing, synthesizing”; for another, it entails “looking at people and decisions and questioning those decisions, figuring out why those decisions were made and, in the scheme of things, how they impact our society today.” On the basis of such airy generalities, we are a long way from concluding that these young teachers possessed “rich and interesting conceptions of historical thinking.”

It seems clear that institutional demands—in particular, the state’s system of accountability—concerned the teachers in this study. What is less certain, however, is how the demands of testing and standards interacted with the subject matter knowledge that the teachers brought into their classrooms. Had the teachers demonstrated a firm grasp of the conceptual and practical elements of historical thinking, we could reasonably infer, with the authors, that teacher practice was compromised by institutional demands. However, without compelling evidence, we can only speculate.

Timothy Kelly, Kevin Meuwissen, and Bruce VanSledright of the University of Maryland offer a similarly discouraging portrait of history education in Virginia. The authors critique the embedded “conceptions” of history in the *National History Standards* and Virginia’s *History and Social Science Standards of Learning*. They contend that both documents feature narratives of “Western growth and dominance” and “the exceptionality and progress of the United States.” Whereas the national standards at least allow for multiple interpretations of the past, the Virginia standards “present history as a finished product ready to be delivered and consumed in predictable fashion.” The authors claim that the Virginia standards include “token” nods towards historical thinking skills, but primarily consist of names, dates, and places for students to “commit to memory.” Therefore, the Virginia Standards of Learning assessments - multiple-

choice, high stakes exams - only reinforce this approach to history and promote teaching for the short-term retention of historical facts.

If, as the authors claim, it is the case that standards promote the memorization of fixed narratives, their case must be heard. It is precisely for this reason that we wished they had provided us with a more systematic analysis. The authors base their conclusions, in large part, upon cursory examinations of the structure and organization of these standards documents and limit their analysis to four national standards and three Virginia standards. There are no coding schemes to identify and quantify the types and amounts of history covered by these standards, and no detailed consideration of how the Virginia tests measure student knowledge. Where are the empirical data that would shed light on this issue? What is the exact amount of content that students and teachers are held accountable for? Are there observations of classrooms that might provide us with ideas about how these policies truly play out in practice? Even restricting ourselves to the written standards documents, we might ask how often themes of “exceptionality” and “progress” appear? What is the ratio of “unicultural” to multicultural standards? How exactly do the Virginia tests relate to these standards? In the absence of empirical data, we are left, for the most part, to take the authors at their word.

Together both Virginia studies highlight the pressing need to evaluate the validity of our accountability systems. We encourage the authors to continue with such important projects. However, without empirically sound arguments, such endeavors, instead of influencing the future of history education, will be too easily dismissed as further examples of the hortatory literature found on all fronts of the history wars.

All told, the variety we encountered in reading these reports attests to a vigorous and healthy field of inquiry. As we plan new studies into how learners grapple with the complexities of historical thinking, whether as fifth-graders or as seasoned professionals, we must not let our desire for a certain state of affairs outstrip our ability to demonstrate it empirically. If there is any doubt, we must redouble our efforts so that our methods of capturing learning in the history classroom withstand the toughest challenges our critics can mount. This vigilance will ensure our continued progress as a field.

NOTES

1. R. G. Collingwood, *The Idea of History* (New York: Clarendon Press, 1946), 242.

2. See, for instance, Collingwood, *The Idea of History*; Jack H. Hexter, *The History Primer* (New York: Basic Books, 1971); Louis O. Mink, "The Autonomy of Historical Understanding," *History and Theory* 5 (1966): 24-47; Inga Clendinnen, *Reading the Holocaust* (Cambridge: Cambridge University Press, 2002).

3. Lee S. Shulman, "Knowledge and Teaching: Foundations of the New Reform," *Harvard Educational Review* 57 (February 1987): 1-22; Samuel S. Wineburg and Suzanne M. Wilson, "Models of Wisdom in the Teaching of History," *Phi Delta Kappan* 70 (September 1988): 50-58.

4. Pamela Grossman, Samuel Wineburg, and Stephen Woolworth, "Toward a Theory of Teacher Community," *Teachers College Record* 103 (December 2001): 942-1012; Matthew B. Miles and M.A. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook* (Thousand Oaks, CA: Sage, 1994); Frederick Erickson, "Qualitative Methods in Research on Teaching," in *Handbook of Research on Teaching*, ed. Merlin C. Wittrock (New York: Macmillan, 1986), 119-61.

5. Allan Collins, John Seely Brown, and Ann Holum, "Cognitive Apprenticeship: Making Thinking Visible," *American Educator: The Professional Journal of the American Federation of Teachers* 15 (Winter 1991): 6-11, 38-46.

6. Lee J. Cronbach and Paul E. Meehl, "Construct Validity in Psychological Tests," *Psychological Bulletin*, 52 (July 1955): 281-302.

7. National Center for Education Statistics, *Trial Urban District in Reading Report*, 2003 <<http://nces.ed.gov/nationsreportcard/reading/results2003/districtresults.asp>>.

8. Larry Cuban, "Determinants of Curriculum Change and Stability, 1870-1970," in *Value Conflicts and Curriculum Issues: Lessons from Research and Experience*, eds. Jon Schafferzick and Gary Sykes (Berkeley, CA: McCutchan, 1979), 142.

9. John D. Bransford, Ann L. Brown, and Rodney R. Cocking, eds., *How People Learn: Brain, Mind, Experience, and School* (Washington D.C.: National Academy Press, 1999).

10. Wineburg and Wilson, "Models of Wisdom in the Teaching of History."

11. Michael Quinn Patton, *Qualitative Research and Evaluation Methods* (Thousand Oaks, CA: Sage, 2002).



COPYRIGHT INFORMATION

TITLE: Historical Evidence and Evidence of Learning
SOURCE: Int J Soc Educ 22 no1 Spr/Summ 2007

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited. To contact the publisher:
<http://www.bsu.edu/classes/cantu/journal.htm>