# Tracking the Footprints Puzzle: The Problematic Persistence of Science-as-Process in Teaching the Nature and Culture of Science

CHARLES R. AULT JR.
*Department of Teacher Education, Graduate School of Education and Counseling, Lewis & Clark College, Portland, OR 97219-7899, USA*

JEFF DODICK
*Science Teaching Center, The Hebrew University of Jerusalem, Givat Ram, 91904 Jerusalem, Israel*

*Received 25 September 2009; revised 26 February 2010; accepted 11 March 2010*

*DOI 10.1002/sce.20398*
*Published online 29 June 2010 in Wiley Online Library (wileyonlinelibrary.com).*

**ABSTRACT:** For many decades, science educators have asked, "In what ways should learning the content of traditional subjects serve as the means to more general ends, such as understanding the nature of science or the processes of scientific inquiry?" Acceptance of these ends reduces the role of disciplinary context; the "Footprints Puzzle" and Oregon's "Inquiry Scoring Guide" illustrate this point. In the Footprints Puzzle, students are challenged to distinguish observations from inferences to learn about the nature of science or the culture of science. Oregon's Inquiry Scoring Guide separates content knowledge from inquiry skills. Given long-standing discredit of "the" scientific method, modern views emphasize the diversity of inquiry methods and explanatory ideals across disciplines. Paleontologists, for example, reconstruct the behavior of extinct beasts from fossil footprints using methods of inquiry responsive to this aim. Figuring out dinosaur locomotion depends upon making analogies to the limb structure and behavior of extant species. The history of the Footprints Puzzle demonstrates that an enduring adherence to "a process approach" obscures how conceptualization intertwines with methodology.

*Correspondence to:* Charles R. Ault, Jr.; e-mail: ault@lclark.edu

A discipline's concepts themselves, such as "extinction" and "geologic time," function as tools of inquiry in distinctive and productive ways.  *Sci Ed* **94:**1092–1122, 2010

## INTRODUCTION

How best to connect content and process and to teach scientific inquiry has preoccupied science educators for a long time. In a comprehensive perspective spanning more than a century, DeBoer (1991), detailed the evolution of programs balancing the teaching of disciplinary content with the practical needs of nonscientist citizens (e.g., the introduction of a "general science" course in the 1920s). In the post-Sputnik era, Schwab (1962) forged an epistemic link between methods of inquiry and disciplinary structures and J. S. Bruner (1963) challenged educational psychology to acknowledge "that the foundations of any subject may be taught to anybody at any age in some form" (p. 12). Their work gave momentum to the design of new, inquiry-style curricula (Rowe, 1978) in several disciplines and an enduring question, "How should schools depict content and process in scientific inquiry?" Even though, as Chiappetta and Koballa (2006) have suggested, subtle linguistic shifts readily prevent the dichotomizing of content and process (content "with" process and science "as" inquiry, for example; p. 147), the status of "process" as dependent upon or easily disembodied from a discipline remains problematic for classroom practice and state assessment.

From the 1960s through the 1980s, the scales tipped away from content as many science educators advocated "a process approach" (American Academy for the Advancement of Science [AAAS], 1967; Gabel, 1984) to teaching science. This approach treated a small number of content-free skills (typically 14 in AAAS' *Science: A Process Approach* [SAPA]) as representative of the sciences, suggesting that mastery of these skills might enhance student learning in several different subjects. Furthermore, the process approach discouraged mixing content knowledge with training in process skills. Proponents worried that students might become confused by the content, feel discouraged, and therefore lose sight of the process objective. For example, in an exercise for teaching about hypothesis testing students were challenged to define variables influencing rotational speed of a "Whirly Bird" device, isolate an independent (or "manipulated") variable, and test for the system's response (also defined operationally as the dependent or "responding" variable) to the manipulated variable (Gabel, 1984, pp. 87–92; Science Curriculum Improvement Study & Berger, 1970). Developing the ideas of angular momentum and rotational inertia (or any intuitive precursors to these concepts) remained outside the lesson's purview—unnecessary complications that might obscure the logic of experimental design central to the process of scientific inquiry.

Advances in child psychology, studies of school learning, and philosophical skepticism regarding the unity of the sciences have indicated serious flaws in the process approach. From the 1980s through the 1990s, research guided by constructs such as "conceptual change" (Posner, Strike, Hewson, & Gertzog, 1982; E. L. Smith, 1991), "meaningful learning" (Mintzes & Wandersee, 1998; Mintzes, Wandersee, & Novak, 1997), "alternative frameworks" (Driver, 1983), and "metacognition" (Gunstone & Mitchell, 1998) established the centrality of prior knowledge, self-awareness, and social context to forming new concepts as well as the sophisticated logic of children's reasoning in terms of the meanings concepts held for them (Klaassen & Lijnse, 1996). Researchers within these traditions argued that concepts functioned to make phenomena intelligible and served as resources for solving problems. In contrast, when taught "disembodied skills . . . children failed to develop meaningful understanding" (Duschl, Swheingruber, & Shouse, 2007,

p. 215). Despite rejection based upon the accumulation of research indicating its shortcomings, the process approach's "legacy persists in both policy and practice" (Duschl et al., p. 216).

Assumptions about child development, often anchored to Piagetian theory and its prioritization of the dynamic nature of logical operations over more static "figurative knowledge" as well as a fundamental distinction between concrete and formal operations (Piaget, 1969/1970), influenced the design of AAAS' SAPA and other inquiry curricula of the 1960s and 1970s. By 1980, studies of young children's reasoning in situations familiar to them (often where a social purpose was quite apparent) had demonstrated limits to the applicability of Piagetian notions of development (Donaldson, 1978). In a seminal study that departed from Piagetian presuppositions, Nussbaum (1979) determined that children's notion of the earth as a body in space ranged across a spectrum of ideas. He found that children reasoned about up and down and the direction of gravity according to imagery of a flat earth, a flat surface within a sphere, a sphere suspended in space above the "bottom" of space, and a sphere with down being toward the center regardless of the position of the observer on the surface. Children's thinking depended on which notion they held, more than on the stage in their development of an ability to decenter perspective—a finding consistent with Donaldson's research. Children, rather than being deficient thinkers were understood to be capable of explanatory reasoning dependent upon conceptual understandings (Metz, 1995). In a recent study, Metz, echoing Bruner, found children able to think even about their own thinking in the context of uncertainty about designing an inquiry (2004) and concluded that "content and process knowledge should not be separated, but rather—as in the work of scientists—emergent conceptual knowledge can empower the inquiry as inquiry can advance the conceptual knowledge" (Metz, 2008, p. 158).

Despite documentation of problems with the process approach, language such as "observe, infer, predict, classify, define operationally, etc." continues to preface objectives for student learning, often independent of subject. Many educators value, for example, teaching students to distinguish observations from inferences. Across four decades, curriculum authors have revived an exercise for teaching this distinction: the "Footprints Puzzle" (Matthews, Chalmer, Stevenson, Harris, & Dexter, 1973; National Academy of Sciences [NAS], 1998). The Footprints Puzzle, in keeping with the process approach, calls upon students to distinguish observational from inferential statements in a weakly constrained context and not for the purpose of learning about a discipline. Despite the demise of the process approach as a stand-alone curriculum, exercises such as the Footprints Puzzle, though responsive to the aim of promoting children's capacity for reasoning, persist in making generic skills of observing and inferring the starting point for teaching not only the processes of science but also the "nature of science" (Bell, 2008; NAS, 1998) and even the "culture of science" (Settlage & Southerland, 2007). Presuming the ubiquity (not to mention epistemic validity), as did the process approach, of these skills reinforces a message of content-free, methodological unity among sciences and scientific inquiries.

Skepticism regarding the depiction of science as a discrete set of processes common across many subjects stems from answering the question, "In what ways do methods of inquiry depend upon the understandings shared within a particular research community?" Many science educators agree that members of diverse scientific cultures engage in a great variety of enterprises, struggling to improve theoretical insights and design methods of inquiry adequate to answering the questions these structures pose (Duschl, 1990; Duschl & Grandy, 2008). A focus on the unifying process of theory-building and awareness of diversity among approaches to scientific inquiry still leaves open the question, "What to teach?" Answering this question ought to embody how approaches to inquiry distinguish, rather than unite, various disciplines of science while justifying the educational value of this plural portrayal of the "elements of practice"—a portrayal featuring the value of "specialized language," the use of disciplinary "tools," and the role of "representations" (Duschl et al., 2007, pp. 264–268).

Although popular instructional materials and state-administered approaches to assessing inquiry purport to follow reliable research studies and implement U.S. national reforms for teaching science (National Research Council [NRC], 1996; Project 2061 [AAAS], 1990, 1993), the report of the death of the process approach is an exaggeration. Indeed, the U.S. National Science Education Standards (NSES) define "scientific inquiry" in part as "a set of interrelated processes by which scientists and students pose questions about the natural world" (NRC, 1996, p. 214). The NSES feature "scientific inquiry" as a distinct content domain (distinct even from "the unifying concepts and processes of science"). Project 2061's rhetoric stresses the importance of the "themes and habits of mind" common to all the sciences. Asking disciplinary-embedded questions such as "How do paleontologists interpret fossil footprints in order to figure out the behavior of dinosaurs?" needs to illustrate the "common themes and habits of mind spanning all the disciplines" that constitute a "coherent vision of the knowledge and skills [for] every high school graduate" (Rosemon, 2009, p. 3). As unifying constructs displace disciplinary aims, adopting the particular habits of mind of paleontological inquiry and paleontologists becomes an afterthought. Whatever the intrinsic appeal of solving puzzles of dinosaur behavior, science of this nature, according to the reform agenda, ought to serve as the means to a more ambitious end. However, there are good reasons to question this agenda and its continued emphasis on science-as-process.

What is wrong with representing inquiry as generic skills that unite the sciences? How is teaching about approaches to inquiry within particular disciplinary contexts better? Answering with diversity foremost in mind turns attention away from common processes and themes and toward particular cases and fields of inquiry: investigating fossil footprints, for example. A single word foreshadows where this shift of attention might lead: story.

Teaching science as more than process means leading students on a "journey into the unknown, with all the uncertainties that new ventures entail" (Bianchini, 2008). The journey becomes satisfying as the unknowns recede and the uncertainties diminish. Meanings are grasped and their significance felt (Pepi, 1985). Story takes form. Although the study of scientific processes may prompt active learning, processes do not constitute story; they fail to achieve the narrative that brings meaning to experience (J. S. Bruner, 1985; J. Bruner, 2002). They easily lead to naïve, inductive learning that overemphasizes the role of "the pupil as scientist" (Driver, 1973, 1983). Although the "interrelated processes" referred to in the NSES may transform questions into empirically investigable form, "imagery and imagination"—essentials of story—generate the questions that jumpstart meaningful inquiry (Millar & Driver, 1987). Story form, furthermore, provides structure conducive to the reenactment of explanation (Klassen, 2010).

Mary Budd Rowe, a scholar whose contributions to inquiry science remain unsurpassed (e.g., the role of language, wait-time, and fate control; Rowe, 1978), believed in the appeal to students of science as specially crafted stories about the natural world (Bianchini, 2008)—as meaningful interpretations of experiences ("experiments" being a particular type of experience). Paleontological interpretation of fossil dinosaur footprints is one such story. To learn this story means to journey through the landscape of genuine fossil artifacts guided by the imagery of evolutionary thought—to engage in disciplined inquiry. Paleontological inquiry, representative of the importance of context to observing and inferring in particular ways, promises fascinating stories that amplify experience with meaning (Ault & Ault, 2009).

## A BRIEF HISTORY OF RELATED STUDIES

### The Diversity of Scientific Methods and the Centrality of Conceptualization

The standards and benchmarks documents in the United States do pay careful attention to the diversity of different scientific disciplines. Attention now given not only to conceptualization but also to the structure of arguments, patterns of discourse, and underpinning values of scientists—the social crucible of scientific debate—vastly exceeds the representation of science-as-process. Still, the lure of teaching science-as-process remains. Standardization and simplification go hand in hand, facilitating aims of comparing student learning about scientific inquiry made legible as generic skills (Scott, 1998). In fact, the State of Oregon requires that teachers score a sample of student inquiry work in science annually, beginning in third grade, using a "one-size-fits-all" scoring guide (Ault, 2010). The guide lists inquiry skills separate from, yet common to, the content domains of physical, life, and earth sciences.

Those who espouse making an understanding of the nature of science central to scientific literacy clearly acknowledge the plurality of methods across subjects and the myth of "the" scientific method (Crowther, Lederman, & Lederman, 2005). Science happens within communities of "epistemic practice"—social bodies that "propose, justify, evaluate, and legitimize knowledge claims within a disciplinary framework" (Kelly, 2008). Appreciating variation among these communities precludes achieving a consensus view of the nature of science. Modern views tend to emphasize cultural, cognitive, epistemological, linguistic, methodological, and sociological features characteristic of diverse research enterprises and recognize that experiments often cannot provide answers to the questions that the historical styles of science, for example, ask (Argamon, Dodick, & Chase, 2008; Cleland, 2002; Dodick, Argamon, & Chase, 2009; Dodick & Orion, 2003; Frodeman, 1995; Gould, 1986; Kitcher, 1993; Kitts, 1977; Mayr, 1985; Orion & Ault, 2007; Schumm, 1991).

Analyses comparing research practices across disciplines reveal not only "that scientific methods are diverse" (Duschl et al., 2007, p. 171) but also that the forms of scientific representation must vary to embody the patterns of the natural world (Driver, Asoko, Leach, Mortimer, & Scott, 1994). For example, fundamental categories of geologic phenomena—faults, deltas, volcanoes—include objects that differ from each other due to unique histories; in contrast, members of chemical categories—elements, isotopes, compounds—have no individual identities that bear upon making reliable predictions (Ault, 1998; Dodick & Orion, 2003; Gould, 1986; Hanson, 1958).

Diverse methods of inquiry and differing forms of representation reflect not only the theory-ladeness of observation but also an even deeper entanglement between inquiry and cognition: "Rarely can a man observe what is not yet for him a conceptual possibility" (Hanson, 1958, p. 175). Or, as Schwab stated, "On the conception, all else depends" (1962, p. 198).

In their analysis of this principle using both historical and psychological examples, Brewer and Lambert (2001) argue that there are instances when theory might influence even perception, particularly "when the perceptual evidence is ambiguous" (p. S176), yet other circumstances when anomalies overcome theory-laden expectations. Theory-laden influences operate with greater impact, according to Brewer and Lambert, as scientists produce and inspect data and communicate findings. For example, Kuhn, in reference to the derivation of physical laws from measurements (measurements being a specific form of observation), argued, "Numbers gathered without some knowledge of the regularity to be expected almost never speak for themselves. Almost certainly they remain just numbers" (1961, p. 175).

A similar conclusion with regard to observations in historical science would seem warranted. Wondering how patterns of discourse along with methodologies might vary from one research community to another, Dodick et al. (2009) conducted a linguistic analysis of thousands of scientific articles from a dozen professional journals, finding stylistic differences between experimental and historical sciences and concluding that these stylistic and rhetorical styles corresponded to distinctive methodological approaches. Geologists, for example, reconstruct earth history from the "bottom up" and realize that this history is "deeply and ineluctably *contingent* and therefore unpredictable even in retrospect" (Rudwick, 2008, p. 560).

Rudwick stresses "the value of attending to the sheer *diversity*" (2008, p. 561) of scientific enterprises:

> The sciences are not all the same, not even all the natural sciences; and we do them no justice and ourselves no favors by continuing to treat physics (or any other single science) as the standard by which all other kinds of knowledge are to be judged either adequate or deficient.

Darwin found strange the assertion that "geologists ought only to observe and not theorise," commenting further that ". . . at this rate a man might as well go into a gravel-pit and count the pebbles and describe the colours. How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service!" (Darwin & Seward, 1903). Before an observation might counter or support a view, however, it must be intelligibly formed—that is to say, a conceptual possibility, an invention of mind.

"Oil is first found in the human mind" (Pratt, as cited in Kastens et al., 2009). The mind converts "colors, textures . . . blotches . . . wiggles" into "explanatory narratives" (Kastens et al., 2009, p. 265). Ethnographers refer to the first stages of this conversion in the mind as "inscription" (Latour, as cited in Kastens et al., 2009): the bounding of nature with artifact, often a symbol. Gowin (1981) and Gowin and Alvarez (2005), who defined concepts as signs or symbols signifying regularity, urged a focus of attention on how concepts guide the recording of events. His concept-event-record triad, in effect, opens inscription to inspection. The investigator, guided by conceptualization, transforms the initial records of events into other representations intended to amplify or clarify patterns, warrant new claims, and modify the conception guiding future inquiry. In geology, records of events become cross-sectional diagrams, correlation charts, and geologic maps. Visual depiction of temporal relationships permeates these artifacts, representing the long view of time characteristic of geology. Geological explanation—the derivation of historical accounts of geologic phenomena—depends upon the order of events in time represented by such artifacts. In brief, the study of geological inquiry has value because it leads to particular habits of mind (i.e., the long view of time), both a "practical tool for decision-makers" and a "fundamental aspect of humanity's self-image" (Kastens et al., 2009, p. 265).

The conception, theory-laden on many levels, directs attention, guides choices, suggests how to proceed, and generates explanatory narratives. It does so within the bounds of shared purposes, conceptual possibilities, and trusted practices of a research community—a guild of storytellers. The primary objection to teaching science as unified by generic processes—to instruction featuring "disembodied skills" expected to generalize across contexts—stems from acceptance of this point of view. Even observing depends upon the conception, if Schwab's dictum holds true—posing a limit to what training in distinguishing observations from inferences might accomplish without attention to conceptualization derived from specific disciplinary expertise.

Studies of scientific enterprises continue to add credibility to Hanson and Schwab's insights. Knorr Cetina (1999) characterized molecular biology and particle physics as distinct "epistemic cultures" adhering to theoretical and procedural commitments suggestive of pervasive "disunity" among the sciences. She flatly rejected the assumption of science as a "unitary enterprise" and instead observes that the "enterprise . . . is not one but many, a whole landscape—or market—of independent epistemic [knowledge producing] monopolies producing vastly different products" (Knorr Cetina, 1999, p. 4). Cartwright (1999) has also captured the limitations and liabilities of the quest to unify the sciences and concluded that physics ought not to stand as the model for other fields to emulate, either in representational forms or methodological commitments. Patterns in the world are disjoint and disjunctive and, when understood in context, lead to a "patchwork, not a pyramid" (Cartwright, 1999, p. 1) of laws and representations expressing order.

Cartwright's work does admit to a unity of attitude among scientists grounded in "conventional empiricism" and committed to the reality of a "messy, mottled world . . . that is the tribunal of our scientific judgments" (p. 6). Cartwright argues strongly against the presumption of a universal scientific method while accepting a unity of attitude. Those who adopt this attitude believe that their "observations are impartial . . . [and can be] (potentially) verified by everyone else," assuming everyone else shares a "basic experience of reality" (Grinnell, 1992, p. 20). Scientific debate may overcome controversy due to "clashes in thought style" (Grinnell, p. 154), but learning science faces limits even a shared scientific attitude may not overcome when basic experiences of reality, molded by belief and conceptualization, differ dramatically. How scientists conceive of events and build structures with which to interpret them matters greatly (Schwab's principle) even though the events, methods, and products of scientific inquiry may have no common structure (Feyerabend, 1993).

In brief and in regard to method of inquiry, philosophers and science education researchers simply do not agree on a unified nature of science (Alters, 1997; Argamon et al., 2008; Cartwright, 1999; Cleland 2001, 2002; Cooper 2002, 2004; Diamond 2002; Dodick et al., 2009; Feyerabend, 1993; Frodeman 1995; Gould 1986, 1989; Mayr 1985; Rudolph & Stewart 1998). Divides may be recognized between experimental and historical methodologies within science (even though both may hold prominence in the same field) or between pragmatist and realist philosophies of science. In part, finding or not finding unity depends upon the level and category of description: Methods may vary, rhetorics differ, and explanatory ideals conflict even as consensus emerges regarding the vital role of conceptualization or "theory building" (Duschl & Grandy, 2008) to all aspects of inquiry.

Consider for a moment how the practices of those attempting to find regularities within complex systems on vast time scales (i.e., paleoclimatology) depart dramatically from those looking for treatments for human disease (i.e., biomedical research). When the patient is the earth and the treatment is the carbon-based global economy, there is only one subject in the study. Has the earth cooled or warmed rapidly during the Holocene Epoch? On what scales and to what extent? The protocol of double-blind, random assignment (with sufficient participants to meet statistical challenges to outcomes) does not fit climate research focusing on a global system. In research about climate change, scientists look for synchronous signals: spikes in oxygen isotope ratios from coral reef carbonates and seafloor sediments. Quite obviously, the protocols for inferring trends in global temperature for the purposes of distinguishing patterns of variability in global temperature on different timescales differ from the protocols for managing variability among human subjects in an experimental design of the efficacy of a medical treatment. In each context, a conception of variability and an understanding of randomness influence the design of the investigation. However, because they refer to very different kinds of regularities, conceptions differ, as do their corresponding methods of inquiry for solving their respective problems.

Windschitl, Dvornich, Ryken, Tudor, and Koehler (2007) point out that the study of complex systems, whether biological, astronomical, meteorological, geological, and so on, typically requires an approach capable of capturing how multiple variables interact probabilistically. They add that in the natural setting patterns of interaction of interest to the investigator unfold without distortions often induced by actively manipulating and controlling variables in the laboratory setting. Windschitl, Thompson, and Braaten (2008) further describe that "a universal scientific method is common in discourse at all levels of science education . . . [and this discourse] . . . subverts young learners' understandings of both the practices and the content of the discipline" (p. 942). They advocate discourse attentive to how models of phenomena within disciplines focus attention on deep thinking about ideas in the context of explanatory reasoning.

A discipline deploys the "specific methods that have proved, in concrete experience, to match the characteristic demands of its own intellectual problems" (Toulmin, 1990, p. 193) and achieve its "explanatory ideals" (Toulmin, 1972). A discipline such as paleontology provides a purposeful context that concepts, through use, depend upon for meaning (Gowin, 1981; Novak, 1977)—a view based upon Wittgenstein's notion of "meaning as use" (Biletzki & Matar, 2010). In sharp contrast with making a distinction between observation and inference the context for learning science, close inspection of concept use requires thoughtful consideration of disciplinary structures and aims as context.

Within disciplines, concepts function as tools of inquiry as well as categories for organizing thought. To ascertain their meaning is to inquire about their use. This philosophical point of view (Wittgenstein's) generates the question, "How do geologists use the concept of 'time' in their inquiries and explanations?" (Ault, 1980). In paleontology, in a very broad sense, the concept of geologic time is a tool of inquiry that functions as a referee among competing histories—if the time relationships do not hold, both in terms of synchrony and sequence (e.g., exactly when each set of superimposed footprints was laid down), then the account becomes suspect (Ault, 1980, 1998; McPhee, 1993). Summarizing conversations with Eldridge Moores (a geologist who has studied the obduction of ocean floor rocks onto continents as tectonic plates collide), McPhee described the use of time to referee findings:

> Where strike-slip faults have sliced a landscape and carried two sides apart, matchups can be traced in time and space. Sedimentary sequences, blue-schist belts, batholithic belts, thrust belts, and mélanges will orchestrally tell what happened. If they are not synchronous, it didn't happen. (McPhee, 1993, pp. 216–217)

Time relationships are just one crucial aspect of the interpretation of fossil footprints. The problem is not just "How old are these fossils?" but "Which tracks came first? How much time passed between the footsteps?" Learning to think about time and with time—with time as a tool of inquiry—is a hallmark of disciplined thinking in the geosciences.

Time limits the observance of events yet their records persist through time. Geoscientists often invoke analogies to overcome the challenge of interpreting past events. Typically, such analogies substitute places for times and in paleontology depend upon comparisons with living creatures. In the study of dinosaur trackways, the concept of "bipedalism" functions to invoke analogies between birds (especially rheas) and theropod dinosaurs and of "quadrupedalism" between hippos and sauropods. Anatomical analogies are essential to drawing inferences about locomotion. The analogies, upon close inspection, differ in nature: the bipedal one indicating a significant degree of shared ancestry and the quadrupedal one invoking similarity of forms based on convergent evolution. Both senses of analogy reflect knowledge of evolution. They are tools of inquiry dependent upon conceptualization, the same as time-as-referee.

Using the footprints puzzle to teach a categorical, linguistic distinction between observational and inferential statements misses an opportunity to teach this principle. The study of trace fossils to reconstruct the behavior of extinct creatures by means of analogy depends upon using the concepts of predator–prey interaction and vertebrate locomotion. At the same time, from the point of view of a learner, their meaning emerges from their function as tools of inquiry. Simplified depictions of unity among the sciences that maintain influential footholds in materials for preparing new teachers and in exercises prompting students to think about scientific reasoning merit careful scrutiny. Often stressed in such materials are the skills of observing and inferring and the existence of a strong, categorical distinction between the two, at the expense of making conceptualization central to either.

## The Oversimplification and Decontextualization of Observation and Inference

Norris (1984, 1985) found suspect the assumptions about the nature of observation in the design activities for teaching process skills: the existence of a fundamental linguistic (and epistemological) distinction between observational and inferential statements, the exclusive origin of observations in sense perceptions, and the belief that observation is "the simplest of all the intellectual activities of scientists" (Norris, 1985, p. 817). The inheritance of the "sense-derived" notion of observation and the sharp distinction between observational and inferential statements makes the nature of science standards as subject to this criticism as was the original process approach. Lederman (1998), for example, has claimed that the distinction has importance that goes beyond process in that it foreshadows a parallel distinction between scientific laws and theories.

Critiques of the process approach have stressed how "human imagination and social processes" (Leach, Driver, Millar, & Scott, 1997, p. 148) combine to become beliefs in what actually exists and influence the act of observing. Stephen J. Gould's *The Mismeasure of Man* (1981) has provided perhaps the most familiar and pernicious example from the history of science of how belief may distort perception and skew interpretation regardless of intention.

When nineteenth-century Europeans undertook investigations of human skulls, they presumed without question the superiority of Whites. Phrenology became a way of explaining this superiority, ranking human races on several measures, and inferring social and cognitive traits based upon the physical forms of skulls (Shorto, 2008). That phrenological data might displace Whites from positions of superiority was beyond the realm of possibility. Brewer and Lambert (2001) credit Gould for tracing how scientific racism (a theoretical bias so deep it shaped belief in what was possible to observe) led, unconsciously, to biased perceptions and treatments of the raw data. Gould's study provides a robust example of inference run astray and deserves an audience in part because of its example of how the bias of belief may blind observation. Yet the story is worth knowing primarily for its insights into the origin and consequences of scientific racism, not simply as a means for looking at observation and inference.

Millar and Driver (1987) convincingly demonstrated that turning process into a content domain confuses means with ends—the proper aim being to depict the sciences according to their characteristic concepts, methods, and purposes rather than a common methodology. Their fruits are understandings, useful in appropriate contexts, and responsive to the demands characteristic of different problems. For example, the derivation of singular, as opposed to universal, statements, is a demand characteristic of geological problem solving (Kitts, 1977), as it is of all other historical sciences. Such statements are about particular events and are, by nature, probabilistic and "sketchy" (Kitts, p. 25)—thus complicating the observation–inference distinction. "The goal of geology is the derivation and testing of singular descriptive statements about the past . . . Retrodiction, not prediction, is the most characteristic inference . . . ” (p. 39). To retrodict is to infer the past. Retrodictions are inferences about what might likely be found in the record of the earth’s past and in what temporal order, given trusted, singular statements about its history. Kitts concludes that geologic claims do not, and cannot, achieve the status of “laws” in the same sense as understood within other disciplines—the singular and the lawful may not be in contradiction, but they are different and both styles of claiming are of value in proper contexts. Making process an “end” obscures this understanding of the nature of geological inquiry, glossing over the particular challenges to making trusted statements about the earth’s past and discounting the importance of observing “within a framework” as well as classifying and hypothesizing “within defined domains” (Millar & Driver, 1987, p. 56).

A decade after criticism of the process approach took hold, Lederman (1998) editorialized that subject matter teaching in science often lacked meaningful purpose and context. Isolated and reduced to factual propositions, teaching content invited rote learning and failed to introduce students to understandings of the nature of scientific inquiry that ought to be widely shared among citizens in a democratic state. Lederman proposed teaching a conception of the nature of science in which distinguishing between observation and inference, extending this distinction into the realm of theories and laws, and emphasizing the tentative nature of scientific knowledge played central roles. Lederman’s proposal again made content a means to a more abstract end. With the advent of Bell’s *Teaching the Nature of Science Using Process Skills*, means and ends united in circular fashion to demote content almost completely. According to Bell, learning to compose and classify statements as observations or inferences directly teaches students the nature of science. This relegation of content to interchangeable parts status negates the distinction between historical and experimental styles of scientific inquiry. Learning to observe and infer as independent skills and assuming a sharp distinction between the two initiates the problem. The problem continues as science content serves merely to achieve more abstract, and universal, ends, hiding how story (“theory” will do as a synonym) makes concepts meaningful and methods of inquiry reasonable.

The “problem with process” is one of neglecting what scientists from different fields actually do to solve the problems peculiar to their interests. The challenge, assuming diversity among the sciences, of “What to Teach?” remains. What does understanding science as practiced in a particular context offer that matters to a citizen in a democratic state, irrespective of his or her status as a future scientist? What is the meaningful purpose and context for learning school science, if not a nature of science upgrade to science-as-process? Of what value are the separate journeys, the different stories, the diverse ways of storytelling? The representation of the sciences unified by common processes (AAAS, 1967; Gabel, 1984) obscures how scientists deploy imagination and imagery, rely upon relevant understandings, and engineer methods of inquiry suitable within particular contexts. Devaluing content discourages the curriculum designer’s search for the imagery of the natural world that might govern productive thinking, prove capable of enriching experience, and contribute to widely shared, informed visions depended upon for democratic governance.

To substantiate this conclusion we offer an analysis of a particular paleontological, fossil footprints problem (Bird, 1941) transformed first into an exercise for teaching science-as-process (Matthews et al., 1973; Padilla, 2000) and later used to illustrate the nature (Bell, 2008; NAS, 1998), and the culture (Settlage & Southerland, 2007) of science. The exercise may be found on page 89 of *Teaching Evolution and the Nature of Science*, where there is a drawing of two sets of intersecting dinosaur tracks (NAS, 1998, pp. 87–89). This drawing represents the discovery of fossil footprints within the Glen Rose Formation (Early Cretaceous age of approximately 100 million years) along the Paluxy River at the

Davenport Ranch in Texas (Lockley, 1991; Bell ascribes the source for the drawing as younger Cretaceous trackways from Alberta, Canada).

The Paluxy tracks themselves created a minor tourist sensation in the 1930s when local residents claimed the discovery of human footprints (probably augmented) next to the dinosaur tracks—a discovery worthy of charging admission (Kitcher, 1982, p. 122; Kuban, 1989, accounts for these "elongate" tracks in terms of impressions made by bipedal dinosaurs varying their walk and stance, as do modern birds; see also American Museum of Natural History ([AMNH], 2010). Most observers delight in observing footprints exposed by the Paluxy, the Connecticut, or the Purgatoire rivers, learning to bask in the reality of Mesozoic time and the scale of evolutionary change. The real story, the science of uncovering dinosaur behavior (and, perhaps, the sociology of exploiting fossils for creationist profit), arguably has more value as a teaching resource than the disembedded lesson in scientific inquiry provided by the Footprints Puzzle.

## THE STORY OF THE FOOTPRINTS PUZZLE

### Origins of the Puzzle

The Paluxy River fossils of dinosaur footprints have had a remarkable influence on science education spanning decades. Incredibly, 40 tons of rock, excavated from the Paluxy site, were boxed, exported, and subsequently displayed at the American Museum of Natural History (Bird, 1941) with the result that the public has had ample opportunity to witness the actual footprints. Generations of science educators have taught lessons using the puzzle based on these footprints to illustrate the distinction between observation and inference and the importance of this distinction to the nature of science.

In the mid-1960s, the National Science Foundation ambitiously funded several projects intended to embody authentic practices of inquiry and the structures of knowledge within leading disciplines (DeBoer, 1991). *Investigating the Earth* (Harris, 1973; Matthews et al., 1973) was the flagship effort in the earth sciences produced by the Earth Sciences Curriculum Project (ESCP). As originally piloted, the ESCP text (American Geological Institute [AGI], 1964/1965) included a passage on fossil traces of "tracks, trails, burrows, and impressions" helpful in reconstructing behaviors and environments. The authors wrote: "Fossil footprints are especially interesting. Before skeletons of dinosaurs were discovered, the three-toed footprints of ancient two-legged animals were thought to have been made by ancient birds" (AGI, 1964, p. 19–14). Interestingly, modern birds are considered to be avian dinosaurs—modified, sister groups of (nonavian) dinosaurs (Prum & Brush, 2003), who themselves may have had protofeather features (Zheng, You, Xu, & Dong, 2009).

The Footprints Puzzle exercise in *Teaching Evolution and the Nature of Science* (NAS, 1998) first appeared in the 1973 edition of *Investigating the Earth*. Teachers who use this exercise are instructed to uncover the drawing panel by panel. In the first panel, two sets of prints angle toward each other. In the second panel, the footprints intersect and intertwine in a circular pattern. In the third panel, a single set of tracks departs the scene. Students share observations and draw inferences about what may have happened, given the footprints as evidence. The possible interpretations are, of course, multiple and varied—with many competing ideas equally valid. Unfortunately, the interpretive exercise for students "bears no resemblance to the fossil footprint evidence" (Lockley, 1991, p. 230; for a side-by-side comparison of Bird's 1941 sketch of these tracks and the 1973 ESCP diagram, see Figure 1).

The ESCP textbook introduced "investigating a footprint puzzle" with the phrase, "suppose you discovered a set of fossilized tracks ... " (Matthews et al., 1973, p. 372). Students read that such tracks were "common in certain parts of New England and in the

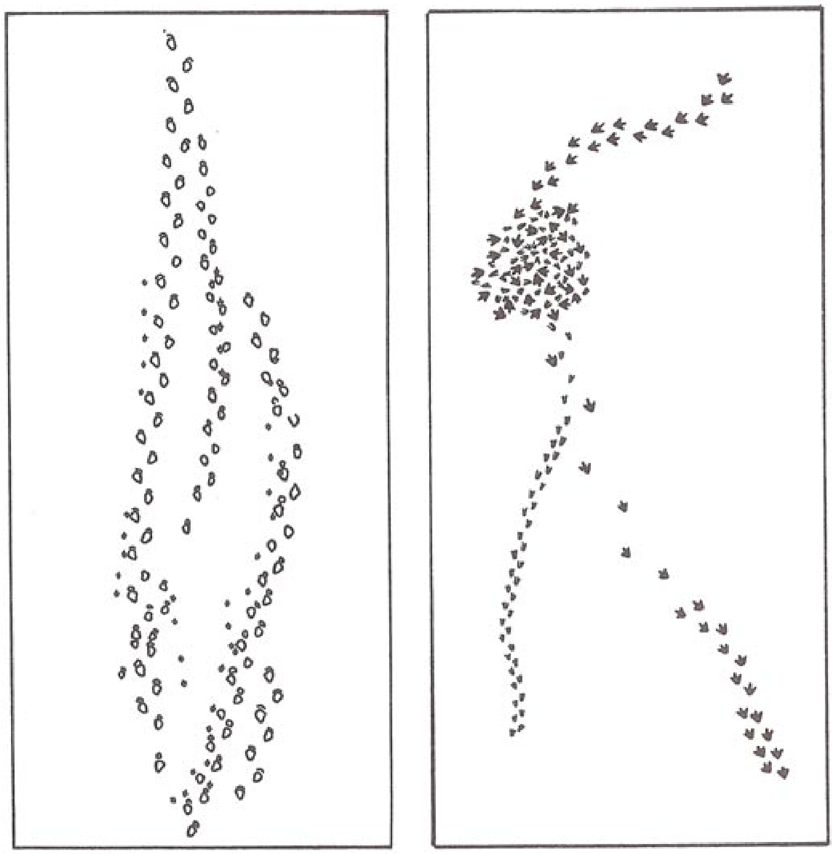

**Figure 1.** Comparison of Bird's original field sketch (left) with the textbook version of the Paluxy River trackways from Lockley (1991, p. 171). © Cambridge University Press 1991. Reprinted with the permission of Cambridge University Press.

southwestern United States," thus freeing the puzzle from claiming to represent a particular locality (whether Alberta, Colorado, Connecticut, or Texas)—a step in the direction of disembedding the artifact from its temporal and geographic context. Questions prompted students to think about whether the two sets of tracks were made at the same time as well as to reconstruct what might have actually happened.

By 2000, Prentice-Hall had published a revised version of the Footprints Puzzle, converting it from a geologic to a "mysterious tracks in the snow" exercise (Padilla, 2000, pp. 26–27). The exercise featured "observing, inferring" as its focus and further disembodied the tracks from Bird's original discovery. (Lockley, 1991, cites a 1989 Prentice-Hall *General Science* text by Hurd, Mattheis, and Synder as an example of the Footprints Puzzle activity, but we were unable to confirm this citation, either through a search of various editions of the cited text or direct correspondence with Prentice-Hall. Lockley has shared that the puzzle reached his desk via a student and may, in fact, have come from the ESCP, making his citation an error; nevertheless, Prentice-Hall did incorporate the puzzle as "A Tale Told by Tracks" into its *Explorer Science: Part B Animals* "Real World Lab" activities in 2000).

In the ironically classified "lab" activity, Prentice-Hall sets the context as walking in the park while suspecting that many animals, though unseen, are about. The footprints no longer look dinosaurian; instead, they are more mammalian. The three-panel presentation remains, as does the pattern of converging trails that circle together in the middle panel. The instructions tell students to write down their observations and for each observation to draw one or more inferences.

In 2007, John Settlage and Sherry Southerland renewed attention to the Footprints Puzzle by featuring it in *Teaching Science to Every Child: Using Culture as a Starting Point*. In an effort to uncover cultural aspects of science, they presented observing, inferring, and classifying as the "foundation of scientific inquiry" (Settlage & Southerland, 2007, p. 32). Having provided a measure of context by describing the diagram as animal tracks in a field of snow (as was done in the Prentice-Hall series), Settlage and Southerland presented the Footprints Puzzle as a way for prospective teachers to "practice observing and inferring" (p. 40).

In 2008, Randy Bell's book, *Teaching the Nature of Science through Process Skills*, added to the Footprints Puzzle's visibility. His book has earned positive reviews and received enthusiastic acceptance among teachers, especially at the elementary level, for making science accessible to novices and teachers uncomfortable with their own scientific knowledge. Bell adapted his exercise, "Trailing Fossil Tracks," from the 1973 ESCP text, the same source as used for the NAS's 1998 publication.

## The Footprints Puzzle as an Icon of Process and the Nature of Science

Bell clearly depicts the puzzle as a fossil, the drawings (exposed in three successive illustrations) resembling actual rock exposures. Having established a reasonable modicum of context, Bell prompts the students to revise their inferences as observations accumulated while imagining themselves on a fossil-hunting field trip in "the wilds of Alberta, Canada" (p. 73).

Nevertheless, Bell's diagram of footprints remains an oversimplification, the same as its ESCP source. Rather than a true record of empirical events (multiple trackways that do not circle about each other so obviously), the diagram neglects how an actual record of the event—a scientist's sketch of the original site or a photograph of the actual impressions—would inevitably include subtleties and ambiguities capable of influencing interpretation. Intended to prompt inferences (Did the tracks occur at the same time?) from unambiguous observations (Do the tracks cross?), the Footprints Puzzle diagram highly constrains subtlety or ambiguity in the depiction of both the impressions made and the paths taken by the wandering, mystery beasts. The diagram makes clear that only one set of tracks departs the scene.

On the most general level, Bell introduces students to the questions, "How do scientists know so much about things they cannot directly observe?" and "What the Earth was like long before there was anyone there to describe it?" (p. 73). His preamble adds that "much of science deals with phenomena and entities that cannot be described through direct observation." His "Trailing Fossil Tracks" provides the teacher with some suggestions about what to expect from students regarding their interpretations of the tracks in terms of their dinosaur origin. He also instructs teachers to remind "students that observations should be limited to what you can directly observe." Bell has admirably problematized observation itself, though the exercise abstracts observing and inferring as generic process skills and posits a sharp distinction between them. Time-as-referee—a conceptual tool of inquiry—makes an entrance but recedes in importance in keeping with the goal of teaching the observation–inference distinction.

For example, a student might observe, according to Bell, that there "are two different kinds of dinosaur tracks." He suggests that students, at first, will likely "embellish" their observations with implicit inferences (such as there being "two different kinds"). Presumably, embellishments are unwarranted. A better observation, he suggests, would be "the impressions have three distinct appendages on one side" (p. 74). The footprints in the diagram are bird-like and invite anyone with familiarity with bird feet to remark on the three, distinct impressions in one direction—perhaps "toes"—and the single bulge in the opposite direction—perhaps a "heel." In his attempt to avoid inference, Bell refers to the probable toes as "distinct appendages." "Impressions have appendages" is an observation; "there are two kinds of tracks" is not, according to the examples guiding teacher feedback to students. Unfortunately, there is very little to distinguish these two statements into discrete types—one with and one without any implicit inference. We do not see any meaningful difference between the status of the two statements.

In a second example of observation, Bell shares that "the two sets of tracks appear to converge on a common point." Of course, the statement might just as well read, "The two sets of tracks appear to move away from a common point." Someone is inferring the direction of travel in both instances—from the side of the presumed foot that contains the inferred appendages (often referred to as "toes"—assuming this inference is warranted). No one, prompted by the context of being in a science class, will likely infer that several kids on pogo sticks, with footprints at the base, made track impressions on a beach to fool people. Imposing some boundaries seems inescapable. Admittedly, Bell does provide moderate contextual information and even adds for teachers the information that the fossil footprints in the activity were those of *Albertosaurus* and *Edmontosaurus*. Hence he suggests the idea of imagining fossil hunting in Alberta—a limited allusion to thinking about how process and content might work in tandem in distinctive ways to solve a particular problem within a specific discipline.

In the assessment section of the Footprints Puzzle exercise, Bell poses the rhetorical question, "How could you be absolutely sure about what happened on the rock slab more than 65 million years ago?" The expected answer is, of course, that no one can be sure—the nature of science is tentative. The aim of the exercise remains to teach how "scientists know so much about things they cannot directly observe" such as "what the Earth was like" in the distant past. Students do not reach a single conclusion because the evidence is inconclusive. "However, the process of observing, inferring, developing explanations, and weighing the explanations against the evidence is completely in line with the work of scientists. This process works in terms of producing explanations that fit the evidence, even if scientists can't prove that such explanations are 'true.'" (p. 78). Bell's approach emphasizes an indeterminate process, not the actual science for improving upon explanation in a particular context. It teaches students to adopt a habit of mind—tentativeness—but without experiencing an authentic reenactment of inquiry.

In each instance of the Footprints Puzzle as an exercise for teaching about the processes of science, the disciplinary context and its characteristic methods melt away, the vacuum as easily filled by humor as by school science. Don Martin, a cartoonist for *Mad Magazine*, did exactly that. In a series of panels titled "On the Trail with a Zoologist," he first drew a bird-watching scientist excited by an encounter with a single set of fair-sized, three-toed tracks (Martin, 1970, pp. 7–8).

The bird-watcher starts off in the same direction as the tracks and soon encounters where they cross with another set of much, much larger ones. In dire fear, the zoologist runs ahead following the smaller set of tracks, hoping to escape the monstrous creature suggested by the giant footprints. Soon the zoologist stands face to face with a towering, clawed, reptile, perched on, for its size, rather scanty feet. Nearby he notices a cute little bird strutting along on oversized feet. The zoologist observed, and inferred—not too wisely, as it turned out.

Similar to Don Martin's comic, the curriculum authors of the ESCP (1973) and its linear descendent (NAS, 1998) Footprint Puzzle have fashioned a fictional episode of science that imposes a judgment about what students ought to learn at school in science class. In place of a disciplinary context, this judgment presumes a unity among the sciences in terms of how scientists think and reason, design investigations, frame questions, and communicate findings.

The depiction of scientific thought and language as neatly falling into the categories of "observation" and "inference" appears suspect. Unfortunately, the Footprints Puzzle's pedigree, inherited from the process approach, is one that exemplifies the effort to treat the complexities of different sciences in this fashion. In various guises, those who invoke the puzzle as a means to teach the distinction between observation and inference as a hallmark of "the" nature of science tend to strip it of authenticity and the rich array of resources for solving problems—for example, the use of analogy with extant bipedal birds or quadrupedal hippos to understand how extinct dinosaurs walked. The resources of rich imagery and relevant knowledge—content—become implicit inferences to weed out, not patterns of ideas to inspect when puzzled by the artifacts of interest. Assuming that observation is "theory-laden" (Hanson, 1958; Kuhn, as cited in Brewer & Lambert, 2001) suggests that exercises that subtract context and content not only deplete the resources for making inferences but also misrepresent how scientists observe. This approach may even entice students into thinking of the claims of science as arbitrary and relative.

In the National Academy's *Teaching About Evolution and the Nature of Science* (1998), the pathways and tracks are drawn to emphasize the potential for making inferences about interactions. The drawing focuses observations. These observations admit to several inferences: a large dinosaur encountered and ate a prey species, for example. Or perhaps there was a corpse that two scavengers visited at different times—and one was able to fly away. Or maybe the two trackways came from Terror Birds, not dinosaurs at all. Or perhaps one was a juvenile, begging for food, and later hitched a ride on its parent's (mother's?) back. If tracks in the snow (not fossils), then maybe they reveal a moment when a baby duck hopped on to the back of a stork, having imprinted at birth on another species. Once imaginative speculation kicks in, the inference making (and assumption testing) fuels delight. There is really no need to introduce as background that these are dinosaur trackways or animal tracks in the snow—or even footprints at all (perhaps they are imitation footprints created by Creationists to fool tourists). The science becomes inviting and accessible—yet fake.

Fake because the record of events—the diagram of intersecting footprints—is a creative piece of fiction, an inaccurate representation of an actual fossil artifact. It has been made more dramatic to invite speculation, but at the same time has lost its connection to the real event. The inferences prompted by the exercise may resemble one or another aspect of the nature of science (indeed, comparing and contrasting the implications of multiple, competing hypotheses reflects problem solving in historical sciences such as paleontology); however, the exercise lacks authenticity and context. Or, more precisely, the exercise has replaced authentic data in a particular context of inquiry with a stylized diagram in an educational context. Observers may interpret the diagram by imposing whatever context they are able to justify. Very little evidence exists to constrain any inference—and there is little direction or purpose to making an inference other than as an exercise in making up a convincing story to accompany a drawing. However, focusing attention on the role played by imposing a context—and thereby invoking prior knowledge—has the potential to enhance the activity.

Across four decades, the Footprints Puzzle has emphasized a generic view of the relationship between observation and inference, strongly suggesting exaggeration in any reports of the demise of teaching science-as-process. Only incidentally does the lesson uncover an approach to solving paleontological problems, the challenge to thinking presented by actual data, the limit to inference stemming from the discipline's characteristic "historical-comparative method" (Gould, 1986), the disciplined use of appropriate analogies to make inferences, or the "retrodictive" and "singular" style (i.e., predicting the order of events in the past with time-as-referee) of making inferences (Kitts, 1977) endemic to the geosciences.

## THE FOOTPRINTS PUZZLE AND THE QUEST FOR UNITY

### Unity and the Reform Agenda

Project 2061's signature document *Science for All Americans* accepts that "scientists differ greatly from one another in what phenomena they investigate and in how they go about their work" (1990, pp. 3–4), yet ends with emphasis on unifying themes (such as systems thinking) and habits of mind common to all sciences. The NSES carefully acknowledge that in the vision of the Standards "inquiry is a step beyond 'science-as-process'" (NRC, 1996, p. 105). This acknowledgment continues, however, to reveal an entrenched quest for unity. The "new vision" expects students to "combine processes and scientific knowledge as they use scientific reasoning and critical thinking" to develop an understanding of the nature of science (p. 105). Generic senses of "reasoning" and "thinking," without characterization by disciplinary context, join the equally undifferentiated "nature" and "process," implying that students are to find ways to apply generic reasoning and thinking skills in tandem with equally generic inquiry processes regardless of the scientific subject. In addition, the NSES promulgate three separate approaches to unification: unifying concepts and processes in science, science as inquiry, and the history and nature of science. Rather than acquiring insight and expertise across an array of distinctive enterprises, the aim of distinguishing a unified nature of science continues to dominate:

> Most who seek to define science for classroom purposes would likely insist that their objective is to accurately represent, if only generally, just how science works—an understanding of which is deemed useful in modern society to be sure, but that is itself essentially free from social or political bias. The most common approach to this task has been to abstract from the complex practices of science some set of universal descriptors, or underlying assumptions that figure in all scientific work. (Rudolph, 2002, p. 65)

This universalist agenda epitomizes an enduring strain in the culture and history of science education. Debate about whether to present scientists as unified by commitment to shared methods of inquiry and thought or to depict for students the differences among their phenomena of interest and relevant methods of inquiry within the sciences preceded the teaching of science in public schools.

### Unity or Disunity as the Archetypal Question in Science Education

At the dawn of the nineteenth century, physics had already earned a position of positivist privilege as a science more advanced than biology (Comte, 1830/2000). Comte speculated that the advanced (more lawful) status of physics was due, in part, to there being greater complexity to biological phenomena and hence more difficulty in ascertaining the laws governing interactions. However, Comte's conclusion that biology was less experimental than physics disturbed Huxley (Huxley, 1854/1901). Comte believed that the comparative method, artfully developed and employed by Cuvier, was especially distinctive and dominant in biology. Comte attributed the ascendancy of the comparative method to Cuvier's careful observations of anatomy with the power to classify living and fossil creatures in keeping with presumptions about the natural order of relationships—turning whales, for example, from fish to mammal. The comparative method made a distinction between the structure and function of an organ, finding ways to unmask natural affinities disguised by diverse adaptations. Using these very same homologies—a conceptual refinement to the comparative methodology pervasive within biology—readily transferred from Cuvier's catastrophic to Darwin's evolutionary approach to solving problems of affinity and ancestry among living things.

Huxley, on the other hand, argued against Comte's discipline-by-discipline portraiture and in favor of the unity of science under the banner of experimental method, claiming biologists were as experimental as anyone else. In part, his position was a general reaction toward physics' condescending attitudes toward biology. Furthermore, he concluded, comparison was "the essence of every science" and required to "discover a relation of cause and effect" (Huxley, 1854/1901, pp. 48–49). Kitcher, in defending the validity of evolutionary science against creationist judgments, has echoed this argument: "There is no basis for separating the procedures and practices of evolutionary biology from those fundamental to all sciences" (Kitcher, 1982, p. 5).

Huxley, lecturing even before the publication of Darwin's *Origin of Species*, had voiced a compelling rationale for incorporating the study of science into the general education of elites (high school for the general public would come decades in the future). Its value, he believed, lay in the objective generation and justification of empirical claims pertaining to all physical events, each such claim demonstrable in material terms and restricted to natural causes. Huxley, too, was a positivist; he disagreed with the founder of positivism's views on the distinctiveness (and implicit hierarchy of trustworthiness, mathematics, and physics at the apex) of different sciences. Yet both aspired to present science as an approach to inquiry with universal pretensions, a privileged position many have assigned to scientific methodology for centuries.

During the twentieth century, John Dewey's popularizers distilled views about scientific methodology into a short list of steps comprising "the" scientific method (Rudolph, 2005). Strong criticism of this epistemological underpinning of science teaching rapidly appeared in the post–World War II era. Vannevar Bush, patriarch of the federal government's post–World War II commitment to funding scientific research, in a review of *Science and Common Sense*, applauded how the book's author, James B. Conant (President of Harvard University), had made it "crystal clear that there is no such thing as the scientific method" (Bush, as cited in Rudolph, 2005, p. 341; for further discussion, see Chiapetta & Koballa, 2006, p. 97),

The ascendancy of a universal characterization of science in curriculum design, the dominance of teaching controlled variable/experimental design as method of inquiry, and the assessment of student performance of generic inquiry skills reflects the bureaucratization of a particular political and professional culture while misrepresenting the diversity of scientific enterprises. Debating the nature of scientific unity has been endemic to teaching science since its inception. However, historians and philosophers of science now clearly counsel attention to diversity among scientific enterprises. In honoring this wisdom, educators must be cognizant of the political pressures that tend to favor the quest for a universal depiction of science in schools, justifying, for example, the decontextualization of observation and inference skills in exercises such as the Footprints Puzzle.

## UNITY AND CODIFICATION OF INQUIRY AT THE STATE LEVEL

### Legibility, Accountability, and Decontextualization

The application of standards serves to standardize behavior; standardizing makes complex phenomena more legible or decipherable. State power and citizen opportunity depend upon making legible the many systems operating throughout society; legibility makes possible the collection of the data required to administer a modern state (Scott, 1998). Accountability presupposes legibility; the illegible is unaccountable. Among the systematized, legible domains of the modern state are property and transport, names and addresses, weights and measures, income and taxes, labor and products, and populations and education.

Unfortunately, making things legible often entails making their representations simplistic. Ultimately, the legible representation begins to stand for the real thing as abstractions replace realities, and the arranged order of the virtual world supplants the untamed complexities of the organic. The forest plan stands for the real forest. The virtual river, modeled by computation, stands for the real river. Grade level performance eclipses child development. A fictional cartoon, the Footprints Puzzle, replaces field experience and genuine artifact.

To make the learning of science legible for purposes of instruction and assessment, pressures mount to standardize the processes of science independent of context. Legibility trumps complexity; state interests override idiosyncratic, local interests and teacher discretion. State control presumes standardization across school districts and therefore rejects particularity in lessons and responsiveness to diverse interests among students (Seltzer-Kelly, 2008). Instead of learning to appreciate singular statements, derived within a particular context and responsive to the demands characteristic of solving such problems, students taught science are expected to extrapolate from their own reasoning insights into the nature and process of science in general.

The drive for legibility, fueled by the political aim of comparing academic performance across schools and nations, readily seizes upon the oversimplification of science as method, process, inquiry, or nature. Making education accountable means making it measurable, which in turn means making it legible. Fossil footprints, instead of being unique impressions of real behavior by extinct beasts, transmogrify into training exercises for mastering a basic tenet: the observation–inference distinction.

According to Scott (1998), centralized authority siphons power from local constituencies to coordinate actions on behalf of larger state interests. Educational standards empower political bureaucracies to colonize local educational communities, subverting local and personal interests and priorities to nationalized agendas; laudable, of course, in many instances (e.g., ending disparities in school performance correlated with race, class, gender, or ethnicity) but not all. For education, this power shift reduces teacher and school site autonomy and limits local discretion over what to teach, such as appreciation of place, interpretation of local landscapes, and community engagement in "place-based" education (Gruenewald, 2003; Gruenewald & Smith, 2008; Semken 2005; G. A. Smith, 2002, 2007). Place-based pedagogy stresses how important geological and ecological (as well as historical and sociological) stories intersect in the particular places where students live. This approach departs rather dramatically from pedagogy devoted to implementing "standards-based" education espoused in national reform documents that define learning "for all" and that states invoke to hold school performance accountable (Ault, 2008).

The quest for universal aspects of science obscures how methods of investigation and conceptual understandings mutually interact in productive and distinct ways and how specific methods in different sciences were derived as a response to overcoming particular disciplinary problems. Local interest in particular topics—aeronautics in Texas, river channel migration in Missouri, migratory waterfowl in North Dakota, plate margin geology in Oregon, alpine ecology in Colorado, fossil footprints in Connecticut—must struggle for an invitation to the standards party, receiving one if able to demonstrate how a particular topic enhances learning a general standard. With regard to inquiry skills, any science will do if children learn to observe, infer, and recognize the difference.

## Oregon's Inquiry Scoring Guide

Classroom teachers in Oregon must score a sample of student scientific inquiry annually beginning in third grade. For all grade levels, the Oregon Scientific Inquiry Scoring Guide prompts ratings in four categories: "forming a question or hypothesis, designing an investigation, collecting and presenting data, and analyzing and interpreting results" (Oregon Department of Education [ODE], 2009a, p. F-1,2). The Scoring Guide calls for ranking at one of six levels within each of the four categories and defines each rank with descriptors reflecting either "application of scientific knowledge, nature of scientific inquiry, or communication" (ODE, 2009a, p. F-2).

Having codified inquiry as a content domain, the State of Oregon acknowledges unfinished business: "These process standards are intended to be interwoven with content in the three science disciplines" (ODE, 2009b, p. 2) and "The standards are not the curriculum" (p. 1). Ultimately, teachers must accomplish the interweaving task—must spin bureaucratic straw into achievement gold. The three disciplines are, of course, physical, life, and earth/space science. State accountability policy has set an agenda that reinforces the bifurcation of science into propositional representation of content and generic characterization of inquiry. This agenda does not address a reciprocal relationship between methods of inquiry and knowledge structured for solving problems within purposeful, discipline-derived contexts.

"What all students are expected to know and be able to do" (ODE, 2009b, p. 1) descends directly from "national trends" embodied in the National Science Education Standards (NRC, 1996) and *Science for All Americans* (Project 2061/AAAS, 1990). "Recent science education reports have recommended that science standards be organized by a small number of big ideas which are essential for all people to understand" (ODE, 2009b, p. 1). Oregon is simply following the national lead in a quest for the best big ideas and most encompassing processes with which to organize school science.

The quest to unify school science reflects belief in the unity of method and universality of representation across the sciences—the opposite of Cartwright's patchwork, not pyramid, imagery. In place of this quest, school science ought to seek the imagery and tools respectful of and responsive to "life in the messy world that we inevitably inhabit" (Cartwright, 1999, p. 18). As Cartwright warns,

> The problem is that our beliefs about the structure of the world go hand-in-hand with the methodologies we adopt to study it. The worry is not so much that we will adopt the wrong images with which to represent the world, but rather that we will choose wrong tools with which to change it. (1999, p. 12)

We assert that the Oregon Inquiry Scoring Guide is just such a wrong tool for changing the world of science teaching and offer an alternative approach respectful of context, illustrated using the case of actual trace fossil inquiry.

## FOSSIL FOOTPRINTS IN AN AUTHENTIC CONTEXT

As a child in the 1820s, Mary Anning left an enduring mark on paleontology with discoveries of both an ichthyosaur and a plesiosaur (Rudwick, 2008). Observations of one of Mary Anning's famous fossils led to the inference that it was a type of crocodile, then fish, then perhaps a kind of platypus, or maybe a reptilian salamander—before settling on the ichthyosaur (literally, "fish-lizard") label. Inference in each case appealed to comparison, in terms of anatomy, to better known creatures, a method championed by Cuvier that transformed nineteenth-century taxonomy.

Emphasis on disciplinary context calls for an understanding of the relationship between methods of inquiry and the nature of the phenomena of interest. As summarized by Kitcher, "... we discover more about the world while simultaneously learning how to investigate the world ... " (1993, p. 202). This is the image of concept and method rowing together in the same boat. Cuvier's comparative method, applied to fossils such as Mary Anning's croc-fish-salamander-platypus-lizard ichthyosaur, generates plausible inferences. The role of prior knowledge of anatomy, subject to revision, functioning as a tool of inquiry stands as an archetypal example of how scientists, facing a particular problem, have engineered appropriate methods in a particular discipline, leading to improved understanding. Teaching how paleontologists do science ought to make this point clear, rather than treating the science as a means for abstracting a universal principle.

### Improving the Understanding of Trackways

Lockley analyzed discrepancies in the interpretation of the tracks found in Dinosaur State Park along the banks of the Paluxy River in Texas. Their discoverer, Roland T. Bird (a most appropriate name for a dinosaur researcher), hypothesized a theropod's (bipedal, carnivorous dinosaur) attack on a sauropod (elephantine, brontosaur type). Lockley, after some careful analysis, found this conclusion "fanciful at best and probably wrong." For example, he found no compelling reason to believe that both sets of tracks were put down at the same time, and noted that the sauropod's tracks indicated no change in gait—something one might expect of a prey pursued by a predator.

Real tracks invite stories about real dinosaurs, thus prompting interest in broader insights about Mesozoic life. For example, comparisons of the *Brontopodus* form tracks (sauropod impressions) from the Early Cretaceous Paluxy River site (Glen Rose Formation, approximately 100 million years ago) with those exposed by the erosive action of the Purgatoire River in Colorado (Late Jurassic, Morrison Formation, approximately 145 million years ago) support the idea that these beasts, separated in life by tens of millions of years, ambled in similar fashion along ancient shorelines. Tracks, it seems, add cues to the interpretations of fossil behaviors in time and geographic space. Living creatures, humans included, continue to enjoy this habit of walking shorelines, whether of lakes or oceans (Lockley, 1991, pp. 127–129).

The fossil footprints along the Paluxy River recorded the march of 12 sauropods and perhaps 3 theropods. (There is another Dinosaur State Park in Connecticut where extensive sets of footprints left by bipedal, largely carnivorous dinosaurs, with affinities to ancestors of modern birds, lie exposed in the Early Jurassic [approximately 200 million years ago] East Berlin Formation strata near the Connecticut River.) In the actual fossils, the footprints of at least three theropods overlap those of the sauropods. Presumably, the carnivorous dinosaurs were following a potential meal. At one point, a set of tracks left by one of the carnivores merges with the trackway of a sauropod and then disappears. Maybe, for a few steps, it walked within the footprints of the much more massive dinosaur. This mysterious and abrupt end to one set of theropod tracks was apparently seized upon as the basis for the Footprints Puzzle textbook diagram. The diagram exaggerates the intersection of the pathways and presents just two sets of tracks from among the 3 predator and 12 prey trackways.

Lockley reminds us that the textbook sketches are "pure fantasy" and that "no such real fossil footprint evidence exists" (Lockley, 1991, p. 171). In the 1998 NAS diagram (as in Bell's 2008 book), as two animals "draw near, the assumption is made that the predator speeds up and attacks the other animal, its poor victim. They run around in circles and the poor victim succumbs, perhaps to dizziness" (Lockley, p. 171). Lockley notes that a perceptive student will wonder why the presumed prey walked along rather steadily, then chose to run in a circle to escape its predator.

The faux footprints prompt inauthentic inferences. Lockley poses questions calling for comparisons useful to the interpretation of the actual fossil artifacts. Each question conjures imagery capable of generating inferences through analogy. Where today do huge beasts record their comings and goings in shoreline muds? How do patterns of footprints due to migration differ from those left by predators attacking prey? What distinguishes the marks of fighting among males of the same species from the signs of being hunted by another? How do footprints mark the switch between the stalk and the kill? Which footprints indicate juveniles at play or adults courting mates? Have solitary animals walked in succession along the same trail or have they passed together as a herd? Are three-toed tracks evidence of dinosaurs, birds, or mammals?

In a search for footprints in the mudflats of Lake Manyara, Tanzania, Lockley and his colleagues discovered modern trackways indicative of attack:

> Hippos are notoriously pugnacious, and territorial battles are commonplace. There could be little doubt as to the exact significance of the several hundred yards of variable trackways we studied. We also found the trackways very educational from a paleontological perspective. They showed us that an attack reveals evidence of running, slipping and sliding, and sudden changes in direction. We would not expect to find blood and dung preserved in the fossil record [as left behind by the fighting hippos], but the trackways would be presumably irregular, unusual, and significantly different from those exposed at Dinosaur State Park in Texas. (Cohen, Lockley, Halfpenny, & Mitchel, 1989, as cited in Lockley, 1991, p. 172)

Roland T. Bird, credited with unearthing the Paluxy tracks in the 1930s, offered an interpretation that has found favor in guidebooks and informed the initial display (since revised) of these tracks at the American Museum of Natural History. Unconstrained by analogy to real modern-day footprints imprinted in African mud by "pugnacious" hippo behavior, Bird imagined that one of the predatory carnosaurs (a theropod type of dinosaur) sandwiched the left-rear flank of a brontosaur (a sauropod-type of dinosaur) in one hold-on-for-dear-life bite. Pegged by its teeth or gripping with its claws and carried by its prey, the carnosaur left the ground, only later to find its footing and yield its hold. Apparently, the ESCP text authors found this account compelling.

Lockley found little, if any, reason to accept Bird's speculation. First, there were no changes in gait or direction—an observation inconsistent with the behavior of both predator and prey during an attack. The tracks crossed in such close proximity that Lockley inferred simultaneity to be impossible because the tail of the brontosaur would have blocked the path of the carnosaur, keeping the appetizing left flank out of reach.

In brief, the two species, assembled in small groups, walked separately. Maybe the carnosaurs were stalking the brontosaurs; maybe they were patrolling a territorial hunting ground some days after the brontosaurs had meandered through. Lockley's knowledge of behavior and anatomy, his comparative style of reasoning (informed by hippo behavior), and his techniques for measuring properties of trackways to determine gaits, combined to both constrain and generate compelling inferences. Bird worked in comparable fashion; through point and counterpoint, their contextualized and disciplined discourse has improved our understanding of these and similar tracks.

At the American Museum of Natural History in New York City, thousands of visitors have viewed this megabeast hunt. For many years, the label accompanying the exhibit accepted Bird's interpretation of an attack in real time. For dramatic effect, the exhibit designers posed two dinosaur skeletons as the imagined trackmakers. One was the theropod *Allosaurus*;

the other, its stalked prey, *Brontosaurus* (equivalent to *Apatosaurus* in recent taxonomy). From the point of view of purists, convenience compromised the exhibit—the available skeletons, although authentic, dated from the Late Jurassic, not the Early Cretaceous. As Lockley lamented, the bones and footprints were out of synchrony by about 50 million years. Exhibit designers, following Lockley's lead, recently updated the interpretation of these trackways (AMNH, 2010).

The science education community promulgated a diagrammatic "fantasy" of fossil footprints (imposing a concentric spiral made from two sets of footprints in place of loosely braided, parallel paths left by several dinosaurs), divorcing this exercise from context to teach the nature of science as anchored in observation and inference. The museum representation of this scenario intended to foster public understanding of science glossed over the facts of geologic time. Meanwhile, paleontologists, through comparisons of limb anatomy, locomotion, and predator–prey behavior, improved upon early interpretations of the Paluxy trackways. Their work discredited notions of human tracks existing alongside those of dinosaurs. The inference of possible stalking behavior displaced that of attack and escape. The real story of the puzzling Paluxy footprints makes the practice of science within a particular discipline quite interesting.

### Science Books with Authentic Fossil Footprint Contexts

A "deep time" book for secondary students, *From Dinosaurs to Darwin* (Dodick & Orion, 2002), focuses attention on Bird's famous photograph of the Paluxy River dinosaur tracks (Bird, 1941; see Bird's Figure 5, p. 78). A set of three-toed carnosaur tracks on the left intersects the sauropod tracks near the upper portion of the photograph (Figure 2). The text prompts students with this authentic artifact and provides background on how large herbivorous dinosaurs and smaller carnivorous ones probably walked (on all four legs or just their hind limbs, respectively). Following a series of lessons on inferring dinosaur behavior from structures and their functions, the authors pose the question, "What in your opinion happened at the site found at Paluxy Creek approximately 70 million years ago?"

The exercise was not isolated to teach a general thinking skill. Instead, the authors embedded it within a robust assortment of lessons intended to foster appreciation of geologic time and evolutionary change. Instead of teaching "Which processes define inquiry as scientific?" Dodick and Orion helped students answer, "What is the value of understanding a specific event embedded in geologic time?" Rather than worrying about aspects of the nature of science all students should learn, they focused attention on how a paleontologist figures out the behavior of extinct beasts from fossil footprints—as well as answers to other paleontological questions. The topic of fossils in Texas is almost painfully specific, yet the exercise faithfully corresponds to the idea that science is what scientists do. This approach may diversify the appeal of science by diversifying the presentation of what scientists do.

A children's book, in reference to a photograph of a fossil trackway, asks, "Do you think the dinosaur that made these tracks was walking or running?" (Gray, 2007, p. 11). Another children's science book (Weidner Zoehfeld, 2007) details how scientists use footprints to learn about dinosaurs. Weidner Zoehfeld's describes one set of "front footprints to look as if they could have been made by the fingertips of a human giant wearing mittens!" (p. 19). The book proceeds to compare three-toed, carnivorous dinosaur footprints with those left by modern birds—the bipedal analogy. These trade books elevate the use of imagery, analogy, and knowledge to help young children understand how scientists reason.

**Figure 2.** Roland T. Bird's 1941 photograph of the Paluxy River trackways reproduced in an Israeli science textbook (Dodick & Orion, p. 73). Neg./Transparency no. 324393 (photo by Roland T. Bird), courtesy the Library, American Museum of Natural History.

## CONCLUSION AND IMPLICATIONS

### Summary

Lockley's research provides compelling and fascinating reconstructions of the behavior of extinct creatures from fossil footprints. To teach a generic process skill, school science has obliterated much of the detail and disciplinary expertise that made his story interesting and his methods productive

This juxtaposition points to several conclusions. First, interpreting fossil footprints may be strongly constrained by the purpose of inferring behavior and reconstructing habitat or weakly constrained by the educational aim of teaching a distinction between observation and inference. In the authentic context, analogy (as the geologic principle of substituting present events for past ones) contributes to productive reasoning. In the latter context, analogical thinking may be misconstrued as making implicit inferences and blurring the intended observation–inference distinction.

Second, observing inevitably imposes boundaries; purely observational statements are an illusion. Framing an observational statement requires the use of concepts to organize thought—a conceptual starting point that opens or closes the possibilities of imagination. Implicit relationships—inferences—lurk in the words that assert the most fundamental observations. Perhaps, the decontextualized diagram reveals not footprints but shadows cast by a swarm of insects, or a scattering of dark leaves against white snow, or the tracks of stilt-walking children in Halloween costumes. Lacking the constraints of context and purpose, there are virtually no implausible categories for framing either observational or inferential statements—and little basis for categorizing a statement as inherently one type or the other.

The significance of this distinction to representations of the "nature" (Bell, 2008) and the "culture" (Settlage & Southerland, 2007) of science extends these implications to bring into question the quest to unify science and to make content the means of achieving more abstract ends. Reform documents do honor diversity of practice, yet valuing science-as-process persists. Moreover, teaching science as inquiry, reinforced by separating inquiry skills from content, tends to gloss over approaches taken by historical sciences and leave the impression that a rather uniform, controlled variable, experimental method pervasively dominates all disciplines. This dominance even leads some scientists to believe that children "have an instinct for controlled experimentation" (Editors of *Scientific American*, 2010, p. 28).

## Oversimplying Scientific Styles

In his foreword to *Inquiry and the National Science Education Standards*, Bruce Alberts reflected that among both children and scientists "Inquiry is in part a state of mind—that of inquisitiveness" (Olson & Loucks-Horsley, 2000, p. xii). He furthermore concluded, completely consistent with the process approach and the Whirly Bird example in the Introduction, that "one skill that all students should acquire . . . is the ability to conduct an investigation where they keep everything else constant while changing a single variable" (p. xiii). Subsequently, the first chapter compared an investigation of stands of dead cedar trees on Washington's coast (Atwater et al., 2005) with a fifth-grade class's efforts to figure out why three trees growing side-by-side looked so different (one had lost leaves, one had yellow leaves, and one had green leaves).

Having puzzled over the curious distribution of stands of dead cedar trees, Atwater imagined how several different processes might produce this pattern. The task of his research became one of ruling out several plausible, competing hypotheses and confirming others. His style of geologic inquiry indeed shared with that of experimental design the goal of ruling out competing hypotheses and advancing empirical support for one that can withstand attempts to contradict it. However, sleuthing the geologic history responsible for the demise of stands of cedars required Atwater to think in distinctive ways. The particular example of inquiry featured in *Inquiry and the National Science Education Standards*—geomorphological fieldwork—illustrated the historical rather than an experimental style of science and reinforced awareness of variation among scientific enterprises. In fact, geologic expertise alone would not suffice: Atwater turned to Japanese historiography, indigenous voices, and computer modeling to reach his most telling conclusion: the occurrence of a great subduction zone earthquake along the Cascadia subduction zone on January 26, 1700, that left in its wake the puzzling stands of dead cedar trees (Atwater et al., 2005). The scale and complexity of coastal deposits, buried forests, and remnants of "sand geysers" precluded "keeping everything constant while changing a single variable" and exemplifies the pedagogical power of a particular place.

Oversimplifying science "as an activity that is capable of producing verifiable knowledge by means of a carefully prescribed experimental method" (Rudolph, 2007) fosters public misunderstanding of scientific enterprises—especially how they differ. What scientists do and how scientists think encompasses more than steps in a stereotypical scientific method (Chiapetta & Koballa, 2006, pp. 96–99). Reducing science to process devalues this insight with pernicious consequences. Rudolph, a historian who has examined the origin and influence of "the" scientific method in school science (Rudolph, 2005), finds the politicization of the science of global climate change, with school boards, for example, prohibiting screening of *An Inconvenient Truth* ". . . a ringing indictment of our national efforts to teach the public about science in its most meaningful sense. Too many of our citizens simply do not understand how it is that researchers figure out what's going on in the world" (Rudolph, 2007, p. 2):

> It's this misunderstanding about how science is done that has been and continues to be exploited by various business and political interest groups. The situation with global warming is a telling case in point. Given that the majority of the public hold an oversimplified view of science—as an activity that is capable of producing verifiable knowledge by means of a carefully prescribed experimental method—it's not surprising that those who seek to undermine public faith in the claims made by climatologists have highlighted the uncertainties in their work.

Rudolph, in counterpoint to the premise of unity, asserts,

> We need to help students understand the variety of methods and techniques that scientists use to explore the diverse phenomena in the world—that is, the process of knowledge construction as it's actually practiced (in all its localized instances) rather than the facile stereotype of some non-existent, singular scientific method. (2007, p. 3)

Rudolph fears a public receptive to demagoguery, this receptiveness predicated upon a widely shared acceptance of the mythology of the scientific method. To the degree that "the" scientific method has morphed to become process skills that illustrate the nature of science, the acceptance of mythology persists.

## The Problem with Process

The problem with process is its lack of potential to cultivate imagery, provide purpose, and amplify the emotional experience of inquiry. A process approach ignores the synergistic relationship between inference and observation at a basic level and between discovering "more about the world while simultaneously learning how to investigate the world" (Kitcher, 1993, p 202) at a more advanced one. Not all science reduces to "the" experimental method; historical approaches are productive, too.

Dodick et al. (2009) conclude that the accumulating evidence and wisdom reflecting variation among styles of science should convince educators to "design more pluralistic curricula" and explore whether this pluralistic approach widens the appeal of science to students. Research ought to examine how to make various subjects inviting, accessible, useful, and authentic for novice learners without resorting to a one-size-fits-all depiction of scientific inquiry that bifurcates content and process (e.g., Oregon's Inquiry Scoring Guide).

Both those advancing the unity of the sciences and those drawn to the importance of distinctive "epistemic cultures" do hold in common a high level aim: teaching the basis for trusting or distrusting bodies of knowledge. Their debate has consequences for curriculum design, for assessment of student learning, and for pedagogy in the classroom. Agenda-setting premised upon the quest for unity among the sciences privileges "the scientific method, the processes of science, or the nature of science" over disciplinary-contextualized approaches to inquiry (and their distinctive features) as well as children's imaginative thinking (and their spontaneous questions) in the design of the curriculum. The reform agenda for school science ought to acknowledge the complexity of disciplined thinking while simultaneously valuing the imagery and narrative necessary to making such thinking meaningful. To value the role of imagery and narrative is to recognize the resemblance of storytelling to theory building (Klassen, 2010). Children, from the primary grades on, can be encouraged to engage, for example, in theorizing through "science talks" as a way to establish context for investigative activity (Gallas, 1995). Disciplines offer direction. Properly introduced, they propose, not impose, ways of thinking while extending the child's experience (Dewey, 1990).

### The Value of Disciplinary Thinking

Interpreting the impressions made by different types of dinosaurs along Mesozoic shorelines calls for disciplined thinking and the thinking of specific disciplines. The science is empirical, though not necessarily experimental. Investigators trained in historical-comparative methods search for salient analogies across time and geography, across species and behavior—invoking the imagery of fighting hippos and running rheas to interpret fossil footprints.

In paleontology, several examples of method and conception functioning in tandem as tools of inquiry stand out: the "long view" of time, using sequence and synchrony in time to referee among competing claims, capturing historical contingency with singular statements, substituting place for time (or living creatures for extinct ones), visually representing temporal relationships, and developing historical narratives as explanatory ideals. These ideas amalgamate content and process in a purposeful context, blurring any sharp distinction between the two. With new knowledge about digigrade (toe-perched) and plantigrade (flat-footed) walking, the gaits of rheas and emus evoke the imagery of stalking carnosaurs. Observing a chicken takes on new meaning, a pleasure no training in a process skill offers.

The quest for universals of science on which to predicate widespread literacy has historically found a receptive audience among educators anxious to find ways of elevating the intelligence of the general population. No one wishes to abandon this enterprise in a society premised upon the democratization of intellect as a means of blocking the tyranny of elites, yet the quest caricatures inquiry. It engages children in spurious tasks, such as classifying statements with little context, as observations or inferences. Contrasting the history of the Footprints Puzzle, a training exercise in distinguishing observations from inferences, with the history of actual paleontological inquiry underscores the limits of this quest. Demoting content to the means for acquiring scientific habits of mind and generic skills of inquiry holds enduring and seductive appeal. As antidote, science teaching ought to balance the positivist traditions of universal laws, empirical induction, and decontextualized logic with the resources of imagery, language, and variegated methods responsive to context and adapted to particular problems.

The struggle to define compelling aims for school science, derived from the particular and distinctive aspects of disciplinary inquiries, is a daunting task. Among them ought to be amplifying the emotional content of experience, enhancing the capacity to assume responsibility, and diversifying the appeal of the sciences. Toward these ends, good theory-based stories, both the product of and a guide to inquiry, may help students grasp the imagery proven useful to understanding the natural world in particular and disciplined ways.

One final story deserves telling—an exemplar of how the childhood appeal of fossils prompted interest, amplified experience, and encouraged social responsibility in the life of a young woman (Ault, 2008):

Along the banks of the Purgatoire River in Colorado, a class of third and fourth graders sat down next to an exposure of dinosaur footprint fossils that extended upstream and downstream, encoding patterns of behavior on a single, steamy day during the Late Jurassic. The Jurassic landscape had long departed, the scene now occupied by arid grasslands eroded down to bedrock by the river. Gazing upon the winding pathway of the exposed tracks, the children wondered what might have happened on that particular day so long ago when the rock was only mud. The experience left them curious about dinosaur stories etched in stone, footprint by footprint.

Decades later, one of these former fourth-graders shared in an e-mail that she had recently visited the site with her own children and that she had also, at age 13, found an *Allosaurus* skeleton, now displayed at the Denver Museum of Nature and Science, in chase of a *Stegasaurus* (Johnson & Troll, 2007). As an adult, she became politically active based on her concern for the preservation of and public access to the Purgatoire River tracks:

> I got an anti-Piñon Canyon expansion resolution passed in my precinct that will go on to the Boulder county Democrats and then hopefully on to inclusion in the Colorado state Democratic platform.... The army's map of the area of interest, at http://www.pinoncanyon.com/images/mapareaofinterestweb.jpg, shows the new larger army site engulfing the dinosaur trackway plus other paleontological sites... No one would get to see the dinosaur tracks ever again if the Army goes through with the expansion. (I. Wood, personal communication, February 5, 2008)

For her, preserving real trackways with the interest of future generations in mind—her children among them—matters greatly. The story of her achievements, anchored to knowledge of actual fossils, embodies the importance of learning a particular subject in a disciplined way. Her story stands in stark contrast with the imposition of a questionable epistemological premise about observational and inferential statements disembodied from original context, idealized as universals of the nature of science, stylized in a fictitious manner, and represented in school science as the Footprints Puzzle.

## REFERENCES


Alters, B. J. (1997). Whose nature of science? Journal of Research in Science Teaching, 34, 39–55.

American Association for the Advancement of Science/Commission on Science Education. (1967). Science—a process approach. New York: Xerox.

American Geological Institute/Earth Sciences Curriculum Project. (1964/1965). Investigating the earth. Text prepared at the University of Colorado for experimental use in secondary school earth sciences courses during the year 1964–1965. Boulder, CO: Johnson Publishing Company (1964) and Washington, DC: American Geological Institute (1965). Subsequently published in Boston: Houghton Mifflin (1967).

American Museum of Natural History. (2010). Trackways. Retrieved February 16, 2010, from http://www.amnh.org/exhibitions/dinosaurs/trackways/lessismore.php.

Argamon, S., Dodick, J., & Chase, P. (2008). Language use reflects scientific methodology: A corpus-based study of peer-reviewed journal articles. Scientometrics, 75(2), 203–238.

Atwater, B., Satoko, M., Kenje, S., Yoshinobu, T., Kazue, U., & Yamguchi, D. K. (2005). The orphan tsunami of 1700. Professional Paper 1707 of the U.S. Geological Survey. Seattle: University of Washington Press.

Ault, C. R., Jr. (1980). Children's concepts about time no barrier to understanding the geologic past. Ph.D. Thesis, Cornell University. Dissertation Abstracts International, 41, 04A.

Ault, C. R., Jr. (1998). Criteria of excellence for geological inquiry: The necessity of ambiguity. Journal of Research in Science Teaching, 35, 189–212.

Ault, C. R., Jr. (2008). Achieving *querencia*: Integrating a sense of place with disciplined thinking. Curriculum Inquiry, 38, 605–637.

Ault, C. R., Jr. (2010). One size fits none? Journal of Science Teacher Education, 21, 1–5.

Ault, T., & Ault, C. R., Jr. (2009). On the trail of Darwin's megabeasts. American Paleontology, 17(1), 17–19.

Bell, R. L. (2008). Teaching the nature of science through process skills. Boston: Pearson/Allyn and Bacon.

Bianchini, J. (2008). Mary Budd Rowe: A storyteller of science. Cultural Studies of Science Education, 3, 799–810.
Biletzki, A., & Matar, A. (2010). Ludwig Wittgenstein. In E. N. Zalta (Ed.). The Stanford encyclopedia of philosophy (Spring 2010 Edition). Retrieved February 9, 2010, from http://plato.stanford.edu/archives/spr2010/entries/wittgenstein/.
Bird, R. T. (1941). A dinosaur walks into the museum. Natural History, 47(2), 74–81.
Brewer, W. F., & Lambert, B. L. (2001). The theory-ladenness of observation and the theory-ladenness of the rest of the scientific process. Philosophy of Science, 68. Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers (Sept., 2001), S176–S186. Chicago: The University of Chicago Press. Retrieved from Academic Search Premier database.
Bruner, J. (2002). Making stories. Cambridge, MA: Harvard University Press.
Bruner, J. S. (1963). The process of education. New York: Vintage.
Bruner, J. S. (1985). Narrative and paradigmatic modes of thought. In E. Eisner (Ed.), Learning and teaching the ways of knowing (pp. 97–115). Chicago: The National Society for the Study of Education.
Cartwright, N. (1999). The dappled world: A study of the boundaries of science. Cambridge, England: Cambridge University Press.
Chiappetta, E. L., & Koballa, T. R. (2006). Science instruction in the middle and secondary schools: Developing fundamental knowledge and skills (6th ed.). Upper Saddle Hill, NJ: Pearson Merrill Prentice-Hall.
Cleland, C. (2001). Historical science, experimental science, and the scientific method. Geology, 29, 987–990.
Cleland, C. (2002). Methodological and epistemic differences between historical science and experimental science. Philosophy of Science, 69, 474–496.
Comte, A. (2000). The positive philosophy of Auguste Comte. Volume II. (H. Martineau, Trans.). Kitchener, ON: Batoche Books. (George Bell & Sons, 1896, version). (Original work published 1830.) Retrieved January 30, 2010, from http://socserv2.mcmaster.ca/~econ/ugcm/3ll3/comte/Philosophy2.pdf.
Cooper, R. A. (2002). Scientific knowledge of the past is possible: Confronting myths about evolution and the nature of science. American Biology Teacher, 64, 476–481.
Cooper, R. A. (2004). Teaching how scientists reconstruct history: Patterns and processes. American Biology Teacher, 66, 101–108.
Crowther, D. T., Lederman, N. G., & Lederman, J. S. (2005). Understanding the true meaning of the nature of science. Science & Children, 43(2), 50–52. Retrieved February 4, 2010, from http://www3.nsta.org/main/news/stories/science_and_children.php?news_story_ID=51055.
Darwin, F., & Seward, A. C. (Eds.). (1903). *More letters of Charles Darwin. A record of his work in a series of hitherto unpublished letters* (Vol. 1). London: John Murray. Letter of September, 18, 1861, from Charles Darwin to Henry Fawcett (p. 195). Originally published in Stephens, L. (1885). Life of Henry Fawcett. Retrieved February 5, 2010, from The Complete Work of Darwin Online Web site http://darwin-online.org.uk/.
DeBoer, G. E. (1991). A history of ideas in science education: Implications for practice. New York: Teachers College Press.
Dewey, J. (1990). The school and society; and, The child and the curriculum. Chicago: University of Chicago Press.
Diamond, J. (2002). Guns, germs and steel: The fates of human societies. New York: W.W. Norton.
Dodick, J., Argamon, S., & Chase, P. (2009). Understanding scientific methodology in the historical and experimental sciences via language analysis. Science & Education, 18, 985–1004.
Dodick, J., & Orion, N. (2002). From dinosaurs to Darwin: Evolution from the perspective of (deep) time: A curriculum for the science for all program. Rehovot, Israel: Weizmann Institute of Science.
Dodick, J., & Orion, N. (2003). Geology as an historical science: Its perception within science and the education system. Science & Education, 12, 197–211.
Donaldson, M. (1978). Children's minds. New York: W. W. Norton.
Driver, R. P. (1973). The representation of conceptual frameworks by young adolescent science students. Ph.D. Thesis, University of Illinois at Urbana–Champaign. Dissertation Abstracts International, 34, 11A.
Driver, R. P. (1983). The pupil as scientist? Milton Keynes, England: Open University Press.
Driver, R. P., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23(7), 5–12.
Duschl, R. A. (1990). Restructuring science education. New York: Teachers College Press.
Duschl, R. A., & Grandy, R. E. (2008). Consensus: Expanding the scientific method and school science. In R. A. Duschl & R. E. Grandy (Eds.), Teaching scientific inquiry (pp. 304–325). Rotterdam, The Netherlands: SensePublishers.
Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). Taking science to school: Learning and teaching in grades K-8. Washington, DC: National Academy Press.
Editors of Scientific American. (2010). Start science sooner. Scientific American, 302(3), 28.

Feyerabend, P. (1993). Against method. London: Verso.
Frodeman, R. (1995). Geological reasoning: Geology as an interpretive and historical science. Geological Society of America Bulletin, 107, 960–968.
Gabel, D. F. (1984). Introductory science skills (2nd ed). Prospect Heights, IL: Waveland Press.
Gallas, K. (1995). Talking their way into science. New York: Teachers College Press.
Gould, S. J. (1981). The mismeasure of man. New York: W. W. Norton.
Gould, S. J. (1986). Evolution and the triumph of homology, or why history matters. American Scientist, 74, 60–69.
Gould, S. J. (1989). Wonderful life. London: Hutchison Radius.
Gowin, D. B. (1981). Educating. Ithaca, NY: Cornell.
Gowin, D. B., & Alvarez, M. C. (2005). The art of educating with V diagrams. New York: Cambridge University Press.
Gray, S. H. (2007). Dinosaur tracks. New York: Scholastic.
Grinnell, F. (1992). The scientific attitude. New York: Guilford Press.
Gruenewald, D. A. (2003). The best of both worlds: A critical pedagogy of place. Educational Research, 32(4), 3–12.
Gruenewald, D. A., & Smith, G. A. (2008). Place-based education in the global age: Local diversity. Mahwah, NJ: Erlbaum.
Gunstone, R. F., & Mitchell, I. J. (1998). Metacognition and conceptual change. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), Teaching science for understanding (pp. 133–163). San Diego, CA: Academic Press.
Hanson, N. R. (1958). Patterns of discovery. Cambridge, England: Cambridge University Press.
Harris, M. F. (1973). Investigating the earth. (American Geological Institute Earth Science Curriculum Project, Rev. ed.). Boston: Houghton Mifflin.
Huxley, T. H. (1901). On the educational value of the natural history sciences. In T. H. Huxley (Ed.), 1901, Science and Education. New York: P.F. Collier and Sons. (Original work published 1854)
Johnson, K., & Troll, R. (illustrator). (2007). Cruisin' the fossil freeway. Golden, CO: Fulcrum.
Kastens, K. A., Manduca, C. A., Cervato, C., Frodeman, R., Goodwin, C., Liben, L. S., et al. (2009). How geoscientists think and learn. EOS, Transactions, American Geophysical Union, 90, 265–266.
Kelly, G. (2008). Inquiry, activity, and epistemic practice. In R. A. Duschl & R. E. Grandy (Eds.), Teaching scientific inquiry (pp. 99–117). Rotterdam, The Netherlands: SensePublishers.
Kitcher, P. (1982). Abusing science: The case against creationism. Cambridge, MA: MIT Press.
Kitcher, P. (1993). The advancement of science. New York: Oxford University Press.
Kitts, D. B. (1977). The structure of geology. Dallas, TX: Southern Methodist University Press.
Klaassen, C. W. J. M., & Lijnse, P. L. (1996). Interpreting students' and teachers' discourse in science classes: An underestimated problem? Journal of Research in Science Teaching, 33, 115–134.
Klassen, S. (2010). The relation of story structure to a model of conceptual change in science learning. Science & Education, 19, 305–317.
Knorr Cetina, K. (1999). Epistemic cultures: How the sciences make knowledge. Cambridge, MA: Harvard University Press.
Kuban, G. J. (1989). Elongate dinosaur tracks. In D. Gillette & M. Lockley (Eds.), Dinosaur tracks and traces (pp. 52–72). New York: Cambridge University Press.
Kuhn, T. S. (1961). The function of measurement in modern physical science. ISIS, 52(2), 161–193. Retrieved February 5, 2010, from http://www.jstor.org.watzekpx.lclark.edu/stable/pdfplus/228678.pdf.
Leach, J., Driver, R., Millar, R., & Scott, P. (1997). A study of progression in learning about "the nature of science": Issues of conceptualization and methodology. International Journal of Science Education, 19, 147–166.
Lederman, N. G. (1998). The state of science education: Subject matter without context. Electronic Journal of Science Education, 3. Retrieved February 4, 2010, from ERIC database http://wolfweb.unr.edu/homepage/jcannon/ejse/ejsev3n2.html.
Lockley, M. (1991). Tracking dinosaurs. New York: Cambridge University Press.
Martin, D. (1970). On the trail with a zoologist. Mad Magazine #132. New York: E.C. Publications.
Matthews, W. H., III, Chalmer, J. R., Stevenson, M. F., Harris, D. T., & Dexter, W. A. (1973). Investigating the earth. Boston: Houghton Mifflin.
Mayr, E. (1985). How biology differs from the physical sciences. In D. J. Depew & B. H. Weber (Eds.), Evolution at the crossroads: The new biology and the new philosophy of science (pp. 43–63). Cambridge, MA: MIT Press.
McPhee, J. (1993). Assembling California. New York: The Noonday Press/Farrar, Straus and Giroux.
Metz, K. (1995). Reassessment of developmental constraints on children's science instruction. Review of Educational Research, 65(2), 93–127. Retrieved from ERIC database.
Metz, K. (2004). Children's understanding of scientific inquiry: Their conceptualization of *uncertainty* in investigations of their own design. Cognition and Instruction, 22(2), 219–290. Retrieved from ERIC database.

Metz, K. (2008). Narrowing the gulf between the practices of science and the elementary school science classroom. The Elementary School Journal, 109(2), 138 – 161.

Millar, R., & Driver, R. (1987). Beyond processes. Studies in Science Education, 14, 33 – 62.

Mintzes, J. J., & Wandersee, J. H. (1998). Reform and innovation in science teaching: A human constructivist view. In J. J. Mintzes, J. H. Wandersee, & J. D. Novak (Eds.), Teaching science for understanding (pp. 30 – 58). San Diego, CA: Academic Press.

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (1997). Meaningful learning in science: The human constructivist perspective. In G. D. Phye (Ed.), The handbook of academic learning: Construction of knowledge (pp. 405 – 447). San Diego, CA: Academic Press.

National Academy of Sciences. (1998). Teaching about evolution and the nature of science. Washington, DC:Author.

National Research Council. (1996). National Science Education Standards. Washington, DC: National Academy Press.

Norris, S. P. (1984). Defining observational competence. Science Education, 68, 129 – 142.

Norris, S. P. (1985). The philosophical basis of observation in science and science education. Journal of Research in Science Teaching, 22, 817 – 833.

Novak, J. D. (1977). A theory of education. Ithaca, NY: Cornell.

Nussbaum, J. (1979). Children's conceptions of the earth as a cosmic body: A cross age study. Science Education, 63, 83 – 93.

Olson, S., & Loucks-Horsley (Eds.). (2000). Inquiry and the National Science Education Standards: A guide for teaching and learning. Committee on Development of an Addendum to the National Science Education Standards on Scientific Inquiry, National Research Council Center for Science, Mathematics, and Engineering Education. Washington, DC: National Academy Press.

Oregon Department of Education. (2009a). Office of Assessment and Information Services. Science Test Specifications and Blueprints 2009 – 2010. Benchmark 2, F1-2. Retrieved August 19, 2009, from http://www.ode.state.or.us/wma/teachlearn/testing/dev/testspecs/asmtscitestspecsg5_0910.pdf.

Oregon Department of Education. (2009b). Oregon K-HS science content standards. Retrieved February 24, 2010, from http://www.ode.state.or.us/teachlearn/subjects/science/curriculum/2009_adopted_k-h_science_standards_updated(11.13).pdf.

Orion, N., & Ault, C. R., Jr. (2007). Learning earth sciences. In S. A. Abell & N. G. Lederman (Eds.), The handbook of research on science education (pp. 653 – 687). Mahwah, NJ: Erlbaum.

Padilla, M. J. (2000). A tale told by tracks. In Prentice-Hall science explorer: Animals (student ed., pp. 26 – 27). Needham, MA: Prentice-Hall.

Pepi, D. (1985). Thoreau's method: A handbook for nature study. Englwood Cliffs, NJ: Prentice-Hall.

Piaget, J. (1970). Science education and the psychology of the child (D. Coltman, Trans.). New York: The Viking Press. (Original work published 1969)

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Toward a theory of conceptual change. Science Education, 66, 211 – 227.

Project 2061 (American Association for the Advancement of Science). (1990). Science for all Americans. New York: Oxford University Press.

Project 2061 (American Association for the Advancement of Science). (1993). Benchmarks for science literacy. New York: Oxford University Press.

Prum, R. O., & Brush, A. H. (2003). Which came first, the feather or the bird? Scientific American, 288(3), 84 – 93.

Roseman, J. E. (2009). Science for all Americans: Timely or timeless? NSTA Reports, 20(7), 3.

Rowe, M. B. (1978). Teaching science as continuous inquiry (2nd ed.). New York: McGraw-Hill.

Rudolph. J. L. (2002). Portraying epistemology: School science in historical context. Science Education, 87, 64 – 79.

Rudolph, J. L. (2005). Epistemology for the masses: The origin of "the scientific method" in American schools. History of Education Quarterly, 45, 341 – 376.

Rudolph, J. L. (2007). An inconvenient truth about science education. Teachers College Record, ID Number: 13216. Date published: February 9, 2007. Retrieved March 31, 2008, from http://www.tcrecord.org/content.asp?contentid=13216.

Rudolph J. L., and Stewart, J, (1998). Evolution and the nature of science: On the historical discord and its implication for education. Journal of Research in Science Teaching, 35, 1069 – 1089.

Rudwick, M. J. S. (2008). Worlds before Adam: The reconstruction of geohistory in the age of reform. Chicago: University of Chicago Press.

Schumm, S. A. (1991). To interpret the earth. Cambridge, England: Cambridge University Press.

Schwab, J. (1962). The conception of the structure of a discipline. The Educational Record, 43, 197 – 205.

Science Curriculum Improvement Study & Berger, C. F. (1970). Subsystems and variables: Teacher's guide. Chicago: Rand McNally.

Scott, J. C. (1998). Seeing like a state. New Haven, CT: Yale Press.

Seltzer-Kelly, D. (2008). Deweyan Darwinism for the twenty-first century: Toward an educational method for critical democratic engagement in the era of the Institute of Education Sciences. Educational Theory, 58, 289 – 304.

Semken, S. (2005). Sense of place and place-based introductory geoscience teaching for American Indian and Alaska Native undergraduates. Journal of Geoscience Education, 53, 149 – 157.

Settlage, J., & Southerland, S. A. (2007). Teaching science to every child: Using culture as a starting point. New York: Routledge.

Shorto, R. (2008). Descartes' bones: A skeletal history of the conflict between faith and reason. New York: Doubleday.

Smith, E. L. (1991). A conceptual change model of learning science. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), The psychology of learning science (pp. 43 – 64). Hillsdale NJ: Erlbaum.

Smith, G. A. (2002). Learning to be where we are. Phi Delta Kappan, 83, 548 – 594.

Smith, G. A. (2007). Place-based education: Breaking through the constraining regularities of public school. Environmental Education Research, 13, 189 – 207.

Toulmin, S. E. (1972). Human understanding. Princeton, NJ: Princeton University Press.

Toulmin, S. E. (1990). Cosmopolis: The hidden agenda of modernity. New York: Macmillan/The Free Press.

Weidner Zoehfeld, K., & Washburn, L. (illustrator). (2007). Dinosaur tracks. New York: Collins.

Windschitl, M., Dvornich, K., Ryken, A., Tudor, M., & Koehler, G. (2007). A comparative model of field investigations: Aligning school science inquiry with the practices of contemporary science. School Science and Mathematics, 107, 382 – 390. Retrieved from ERIC database.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. Science Education, 92, 941 – 967. Retrieved from ERIC database.

Wood, I. (2008). Personal correspondence (e-mail). Re: Ft. Carson and footprints. Received Tuesday, February 5.

Zheng, X. T, You, H. L, Xu, X., & Dong, Z. M. (2009). An Early Cretaceous heterodontosaurid dinosaur with filamentous integumentary structures. Nature, 458, 333 – 336.