REVIEW

Educational Interventions to Advance Children's Scientific Thinking

David Klahr,[1] Corinne Zimmerman,[2] Jamie Jirout[1]

The goal of science education interventions is to nurture, enrich, and sustain children's natural and spontaneous interest in scientific knowledge and procedures. We present taxonomy for classifying different types of research on scientific thinking from the perspective of cognitive development and associated attempts to teach science. We summarize the literature on the early—unschooled-development of scientific thinking, and then focus on recent research on how best to teach science to children from preschool to middle school. We summarize some of the current disagreements in the field of science education and offer some suggestions on ways to continue to advance the science of science instruction.

Science education aims to advance children's knowledge about the natural world and to help them master procedures for discovering, assessing, revising, and communicating that knowledge. We believe that science interventions can be most effective when they are consistent with what research in cognitive development has revealed about children's thinking and learning. This is not the only lens through which to view science education literature, nor is it one usually used by science educators, who necessarily focus on the complexities of the knowledge they are attempting to convey and the constraints imposed by the realities of classrooms and schools.

Psychologists have been investigating the development of basic cognitive skills that support scientific literacy for more than 50 years (*1–4*), making it possible to design theoretically grounded educational interventions that can advance children's scientific thinking. Three necessary components for any such intervention are: a statement of the knowledge to be acquired, a set of instructional activities that are consistent with what is known about the constraints of human thinking and learning, and an assessment process.

Here we describe some ways in which research in cognitive development has advanced our understanding of children's scientific thinking, and review how this research interfaces with science instruction at two different developmental phases: preschool (including infancy) and K-8 science.

A Taxonomy for Classifying Interventions in Science Education

Scientific thinking can be characterized in terms of two principal features: (i) content, which includes an array of domain-specific topics, such as physics, chemistry, biology, Earth sciences, and so on, combined with a smaller set of domain-general concepts, such as equilibrium, time, feedback, and causality; and (ii) processes, including formulation of hypotheses, design of experiments and observations, and evaluation of evidence. (*5*).

This framework can be used to classify different types of psychological investigations of scientific thinking (Table 1). The two rows in Table 1 are intended to emphasize the fact that "science educators aim to convey not only the content of science" (row 1) "but also the processes whereby scientific knowledge is acquired, refined, revised, extended, and disseminated, including modes of argumentation and the social and professional context of the scientific enterprise" (row 2). (6). Research on domain-specific hypotheses (cell A) assesses young children's knowledge about the Sun-Moon-Earth system, in which children progress, between first and third grade, from a variety of geocentric beliefs to a variety of heliocentric beliefs (Table 2). Even by third grade, most children's models are only partially correct (7). One of the challenges of science instruction is that rather than being empty vessels into which knowledge can be poured, novice science learners bring to the classroom many misconceptions, including some that may require radical reconceptualization. (8).

Studies in cell F, focusing on how children evaluate abstract evidence patterns, reveal

Table 1. Categorization of types of foci in psychological studies of children's scientific thinking.

Type of knowledge	Type of scientific processes: Forming hypotheses	Designing and running experiments and observations	Evaluating evidence
Domain-specific	A	B	C
Domaingeneral	D	E	F

Table 2. Distribution of children's beliefs about the relative motion of the Sun, Earth, and Moon. Numbers indicate the number of children in each grade holding the various beliefs about the motion of the Earth, Moon, and Sun (*7*).

Earth motion	Moon motion	Sun motion	Grade 1	Grade 3	Total
1. Rotates, revolves around Sun	Rotates, revolves around Earth	None	0	1	1
2. Rotates, revolves around Sun	Revolves around Earth	None	1	5	6
3. Rotates, revolves around Sun	Moves parallel to Earth around Sun	None	0	1	1
4. Rotates, revolves around Sun	None	None	2	2	4
5. Rotates, revolves around Sun and Moon	None	None	0	1	1
6. Rotates	Rotates	Rotates	1	0	1
7. Rotates	None	None	0	2	2
8. Rotates	Revolves around Earth	Revolves around Earth	0	1	1
9. None	Revolves around Earth	Revolves around Earth	0	2	2
10. Rotates	Rotates, up and down	Rotates, up and down	1	0	1
11. Rotates	Up and down	Up and down	2	1	3
12. None	Up and down	Up and down	9	3	12
13. None	None	None	2	0	2
14. Rotates, revolves around Sun	Moves with Earth around Sun, up and down	Rotates, up and down	1	0	1
Total			19	19	38

[1]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [2]Department of Psychology, Illinois State University, Normal, IL 61790, USA

that preschoolers can distinguish conclusive from inconclusive evidence patterns and that they can be trained to correctly interpret even complex patterns. *(9)*. Studies in cells B and E focus on the logic of unconfounded experiments. In cell E, investigators examine children's ability to learn about the conceptual and procedural basis of experimental design, without concern for underlying domain-specific knowledge (*10*), whereas studies in cell B explore the interaction between domain-specific knowledge and the logic of experimentation *(11)*.

In some laboratory studies of children's scientific thinking, and in most science education contexts, children negotiate the entire cycle of inquiry (cells A through F) while engaged in self-directed exploration of multivariable systems that simulate the processes of scientific discovery. Such studies enable researchers to examine the dynamic interaction between domain-general strategies and developing conceptual knowledge (*12*, *13*). This research has identified several factors that influence the development of scientific thinking skills, including the following:

1) The amount, strength, and veridicality of prior knowledge (*14*). For example, most children believe that heavy objects sink faster than light objects. When investigating the sink rates of objects of different size, shape, and density, children often fail to isolate weight as a possible causal factor, because they believe that they already know its causal status, or if they do so and find unexpected results, they often attempt to explain them away *(15)*.

2) The specific domain of inquiry. For example, fifth-graders exhibit greater metastrategic understanding and make more valid causal inferences when reasoning about physical, rather than social, domains (*1*6).

3) The perceived goal of inquiry; i.e., whether children approach multivariable tasks with a scientist versus an engineering mindset. The former aims to uncover causal regularities, and the latter aims to produce effects (*1*7).

Phases of Scientific Thinking in the **Early** Years

The issues associated with nurturing, enriching, and sustaining children's interest in scientific knowledge and procedures differ with the phase of development

Preschool* science *assessments and interventions. The enthusiastic wonder with which both children and scientists approach the world around than may account for the alluring notion of "the scientist in the crib" (18). However, research on early cognitive processes reveals that thinking processes follow a developmental trajectory involving the acquisition and coordination of many component *skills*. Although very young children have competencies that support aspects of scientific thinking (*19*), many children leave school having failed to learn much about science Even for those who go on to advanced careers in science, many years of intense training are necessary to become a "real" scientist.

Much of the literature (*1*) on infants' acquisition of fundamental knowledge focuses on aspects of the physical world, such *as* momentum (*20*), solidity (*20*), and gravity (*21*), but there is research on infants' understanding of the biological *(22)* and social worlds as well. However, there is no consensus on how scientific the thinking of young children really is. Some researchers support the "child as a scientist" position (*19*), whereas others challenge this view *(10)*. Efforts to train scientific thinking in young children have yielded mixed results. Although there is no evidence that interventions in the first 18 months can accelerate the course of these developmentally primary *(23)* processes to produce "baby Einsteins" (*24*), there is evidence that preschool children can be trained to improve their control of some mental processes that are widely agreed to be important for learning and understanding science (and mathematics): self-@ation, cognitive flexibility, and inhibitory control *(25)*.

Another general cognitive, and motivational, aspect of scientific thinking is curiosity. Children bring a spontaneous curiosity to the natural world (4). However, the construct of curiosity has proven difficult to operationalize. One broad approach to preschool science education, perhaps influenced by Piagetian theory, presumes that preschoolers traverse a fixed sequence of stages with respect to scientific thought. This perspective tends to constrain efforts to include much scientific content in the preschool curriculum. For example, a study of 20 Midwestern middle-class preschools found that less than 5% of instructional activities were explicitly designed to promote science learning *(26)*. The other approach presumes that preschool programs should aim to nurture children's natural scientific curiosity because, it is argued, "Real science begins with childhood curiosity" *(27)*. The goal of such interventions is to help children develop *early* forms of the complex concepts involved in scientific reasoning (*28*).

This developing interest in the feasibility of early science instruction has led most *states* in the United States, as well as high-level national advisory panels, to formulate science standa preschool education in which curiosity central role *(29)*. But preschool teachers dilemma because there is no consensus what curiosity is or how to measure it *(3i*

Nevertheless, science is finding a place school curricula that encourage teachers tend, stimulate, encourage, and draw on ch curiosity *(31)*. Procedures *to* produce such e

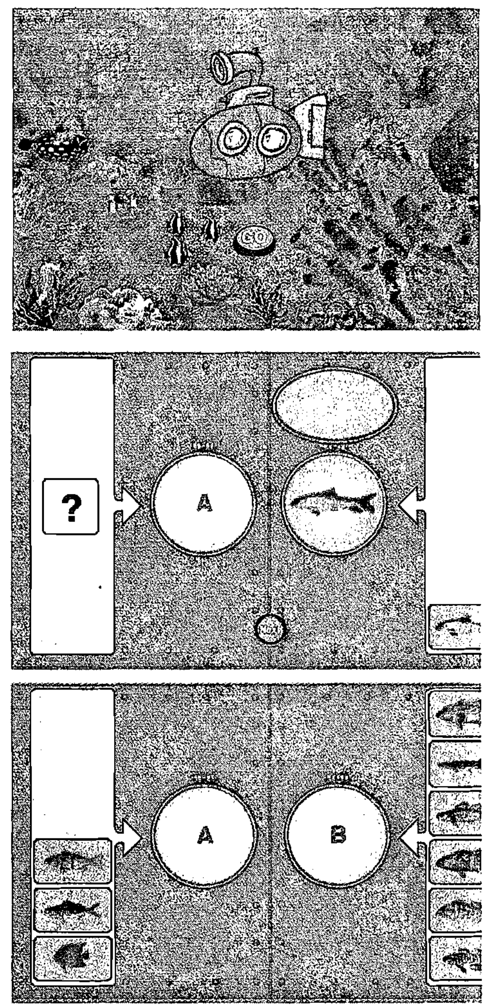

Fig. 1. Curiosity game for preschoolers. Children choos of two windows to open in order to see what kind o outside the submarine. For each of several trials, th adjacent to each initially closed window shows one to si a question mark. The number of possible fish correspond amount of uncertainty associated with each window. middle panel shown here, the window on the Left has ma uncertainty and the window on the right has the m uncertainty (if children choose it, they know for sure wh will appear). The middle panel contrasts two levels of tainty: window A will reveal one of three fish, window reveal one of six fish. Children work their way through a tree of 18 trials contrasting varying levels of uncertai riosity is indicated by the amount of uncertainty the child throughout the task (36).

must address questions of content, delivery, and assessment. Unfortunately, these curricula lack clear procedures for assessing their curiosity-increasing effects. The first two questions are the easiest to answer because they concern inputs (instruction) rather than outputs (measures of changes in curiosity), and preliminary answers can be found in the following three preschool science curricula.

One program, the *Young Scientist Series* (*32*), provides professional development tools for teachers, building on prior knowledge and encouraging scientific thinking and behavior. Assessments of its effectiveness focus primarily on instructional support rather than student outcomes (33). *Science Start*, another preschool program emphasizing professional development, aligns content with existing science standards and integrates science instruction with language and literacy, [illegible] and social studies (*31*, *34*). It emphasizes scientific vocabulary development, as well as planning and problem-solving skills. The effectiveness of the language development portion of the program has been empirically supported (*35*), although its impact on other aspects of children's scientific thinking has not been assessed. *Preschool Pathways to Science* incorporates basic research on children's ability to engage in relatively complex thinking. It provides children with a mental structure, creating a base of knowledge on which to build when experiencing new information. It focuses on teaching the vocabulary and processes of observing, predicting, and observing to check predictions (36).

Thus, the question of how to assess the impact of preschool science programs on children's curiosity remains. Operationally defining curiosity is a first step. Recent work suggests that it can be assessed using a measure of children's exploratory preference for different levels of uncertainty, in a computer-based game in which children choose to explore among situations varying in the amount of information available (Fig. 1). The validity and reliability of this measure of curiosity indicate that it is, in fact, related to children's basic inquiry skills (28).

Elementary and middle-school children. K-8 curriculum developers have traditionally underestimated the developmental readiness of children to engage in scientific thinking. Children entering school have already learned a substantial amount about the natural world, and they possess reasoning processes that support causal inference and evidence interpretation (*4*). However, much of children's scientific content knowledge is implicit, often including mistakes and misconceptions (Table 2). The instructional challenge is to diagnose and remediate these misconceptions while simultaneously building on correct knowledge. Examples of how to do this in specific content areas are available for K-8 science teachers (*37*).

The expanding range of substantive topics in science is daunting. By some estimates, there are thousands of concepts that could be taught (38). Therefore, rather than focus on the content of interventions for teaching either domain-specific or cross-cutting concepts, we review current research about how best to teach science. This active and contentious (39) research area is important because the way that science is taught is inextricably connected to what students learn about the nature of science itself. The controversy over inquiry approaches is characterized by several dichotomies, the most common of which is direct instruction versus discovery learning (40). Most influential science curriculum publications lean heavily toward inquiry (*30*), whereas many researchers from a cognitive science tradition argue that a guided form of explicit instruction is consistent with decades of research on the parameters and structures of the human cognitive system (*41*, *42*).

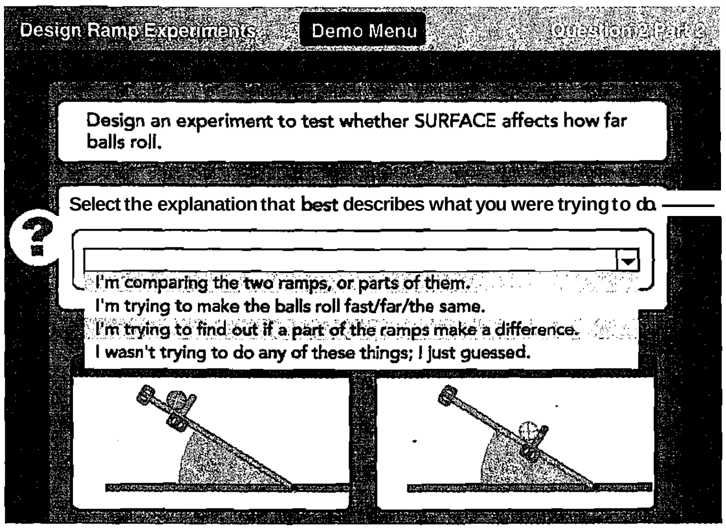

Fig. 2. TED (Training in Experimental Design) is an intelligent computer-based tutor for teaching children how to design unconfounded experiments (*53*). In this screen shot, children are being asked to design an unconfounded experiment to determine whether the type of surface makes a difference in how far a ball rolls.

Educational interventions as engineering artifacts. Instructional design and curriculum development can be viewed as the engineering application of the basic science of cognition: Based on the best available science, one crafts a complex artifact, ranging from a problem set to a lesson plan to an entire curriculum, and then measures performance in non-idealized circumstances (real classrooms with real teachers and students) (43). The design, implementation, and assessment of these artifacts may be influenced by theoretical stances, but ultimately an operational definition of the teaching method must be provided, so that others can replicate, modify, and assess it. However, because this is difficult, interventions are often given broad nonspecific labels, such as teacher-centered, student-centered, discovery, direct instruction, or hands-on.

These broad, and vague, labels for different types of interventions can be replaced with descriptions of instructional methods that are presented in sufficient detail to be replicated. Studies from our lab assess the impact of different approaches to teaching children from second to sixth grade how to design unconfounded experiments. This central domain-general topic, often called the control of variables strategy (CVS) in the literature, is included in the National Research Council's (NRC's) science education standards (29); Benchmarks *for Science Literacy* (*44*); and high-stakes science tests at state, national, and international levels.

In our studies, we used materials in which four two-level factors could be varied to determine whether or not those factors are causal with respect to an outcome. Our contexts have included ramps, springs, sinking objects, and pendulums and have been instantiated in both physical and virtual worlds (45) and currently include an adaptive computer-based tutor (Fig. 2), in which four potentially causal factors can be contrasted or controlled: surface texture, run length, ramp height, and ball type. The learner is asked to design experiments to investigate specific questions (such as, does surface texture make a difference in how far a ball will roll?), and the system diagnoses learners' responses and adaptively decides on the next instructional component.

In one of our studies (*9*), we contrasted three interventions labeled discovery learning, Socratic instruction, and direct instruction. Because each of these terms on its own could cover a huge variety of instructional interventions, we provided an unambiguous operational definition for each method (Fig. 3). Indeed, it is essential to state the details of the three approaches in order to assess and replicate them. The explicit information contained in Fig. 3 enables discussions of differential effectiveness to be grounded in well-defined aspects of the instructional manipulations.

At each grade level, direct instruction was the most effective for immediate learning, near-transfer assessments, far-transfer assessments (in new contexts), and remote transfer assess-

Three instructional conditions

Aspect	"Direct"	"Socratic"	"Discovery"
Materials	Ramps, springs, sinking objects		
Goal setting	By teacher: Can you find out whether X makes a difference in how far the ball rolls?		
Physical manipulation of materials by child	Yes	Yes	Yes
Design of each experiment	Teacher	Student	Student
Probe questions	Yes	Yes	No
Explanations	Yes	No	No
Summary	Yes	No	No
Execution of experiments	No	Yes	Yes
Observation of outcomes	No	Yes	Yes

Fig. 3. An operational definition of the generic terms "Direct Instruction," "Socratic Instruction," and "Discovery Leaning," used in an experiment to teach second-, third-, and fourthgrade children how to design unconfounded experiments (*9*). Each column—corresponding to one type of instruction—contains the values of the essential features listed in the rows. For example, the "Probe Questions" row indicates that there are probe questions for two of the conditions, but not for the "Discovery" condition, and the "Execution" row indicates that students do not *execute* experiments in the Direct condition, but they do in the other two. Sufficient detail is provided so that other researchers can explore replications and modifications of each type of instruction. The column headings are convenient generic labels, but they are not intended to be universally accepted definitions.

ments (after delays of months or even years). Subsequent studies have replicated the general finding that explicit instruction was most effective in the short and long term, in both carefully controlled single-classroom studies as well as large-scale interventions (36 classrooms with nearly 800 total students) (46). Similar studies from other labs have demonstrated that children can learn CVS from less–directed instruction, given extensive scaffolding (i.e., guided instruction/discovery). However, children take much longer to reach mastery in that case, and they are no better at transferring knowledge to new contexts than children who received more explicit instruction (47).

Kuhn and colleagues (48) have also investigated children's ability to learn about CVS, but with a broader focus in which students use computer-based experimental design contexts to explore ways to promote the metacognitive and metastrategic skills involved in differentiating and coordinating theory and evidence. Kuhn maintains that such skills differentiate individuals with more or less sophisticated scientific thinking and represent one of the ways in which children *are* not necessarily intuitive scientists. For example, fifth-graders classified as either high or low academic achievers were explicitly taught metastrategic knowledge of the CVS (*49*). Students interacted with a computerized task to determine how five variables affected seed germination. After an initial investigation of the task, the control group was taught about seed germination, whereas the experimental group was given a metastrategic knowledge intervention. The intervention consisted of describing the CVS and discussing which features of a task indicate when and how the CVS should be used. Students receiving the intervention showed both strategic and metastrategic gains that were still apparent in transfer tasks administered 3 months later. Low academic achievers showed the greatest gains. Thus, although metalevel competencies may not develop routinely, they can be learned via explicit instruction.

With respect to the issue of unambiguous operational definitions, we note that the descriptions of the three types of CVS instruction used in our research (Fig. 3) are less complex than the descriptions necessary to define many other methods used in science instruction, such as modeling, explanation building, group work, argumentation, etc. Nevertheless, we believe that in order to replicate, evaluate, and fully interpret educational experiments, it is necessary for researchers to strive toward such clarity (59.

Converging Trends to Improve the Quality of Science Education

The current state, and likely future, of science education have been profoundly influenced by three NRC reports crafted by experts from the learning sciences, cognitive and developmental psychology, and science education that summarize the *state* of the art of knowledge about human cognition and learning (*41*), lay the groundwork for the integration of psychological models and psychometric procedures (*51*), and challenge the existing state of educational research by setting forth clear guidelines for increasing the scientific rigor of the discipline (52). All of this bodes well for the future of this field and suggests that we will continue to see substantial progress toward solving many of the challenging issues [redacted]g effective science education for our children (*54*).

References and Notes

1. J. Piaget, *The Child's Construction of Reality* (Routledge and Kegan Paul, London, 1955).
2. R. F. Kitchener, *Sci. Educ.* 2, 137 (1993).
3. D. Kuhn, R. S. Siegler, Eds, *Handbook of Child Psychology* (Wiley, Hoboken, NJ, ed. 6, 2006).
4. R. A Duschl, H A Schweingruber, A Shouse, *Taki to School: Learning and Teoching Science in Grac* (National Academies Press, Washington, DC, 200
5. D. Klahr, H. A. Simon, *Psychol. Bull.* 125, 524
6. N. S. Newcombe et *al.*, Am. *Psychol.* 64, 538
7. A Samarapungavan, S. Vosniadou, W F. Brew *Cogn. Dev.* 11 491 (1996).
8. M. T. H. Chi, J. *Learn. Sci.* 14, 161 (2005).
9. D. Klahr, Z. Chen, *Child Dev.* 74, 1256 (2003
10. Z. Chen, D. Klahr, *Child Dev.* 70, 1098 (1999
11. T. Jaakkola, S. Nurmi, *J. Comput. Assist. Learn.* 24, 2
12. A Keselman, *J. Res. Sci. Teach.* 40, 898 (200
13. K Metz, J. *Leorn. Sci.* 20, 50 (2011).
14. S. Vosniadou, *International Hondbook of Rese Conceptual Change* (Routledge, New York, 20
15. D. E. Penner, D. Klahr, *Child Dev.* 67, 2709 (
16. D Kuhn, S. Pearsall, *Cogn. Dev.* 13, 227 (19
17. C Zimmerman, *Dev. Rev.* 27, 172 (2007).
18. A Gopnik, A Meltzoff, P. Kuhl, *The Scientist i Minds, Brains, and How Children Leorn* (Willia New York, 1999).
19. R. Baillargeon, *Curr. Dir. Psychol. Sci.* 13, 89
20. G Gredebäck, C von Hofsten, *Infancy* 6, 165
21. L B. Cohen. C H Cashon, in *Handbook of Ci Psychology: Vol. 2. Cognition, Perception, and* W. Damon, R. M. Lerner (Series Eds.) and D. R. S. Siegler (Volume Eds.) (Wiley, New York, e pp. 214–251
22. K. Inagaki, G Hatano, in *Essays in Developmenta* (Psychology Press, New York, 2002), pp. 19–66.
23. D. C. Geary, *The Origin of Mind: Evolution of Cognition, and General Intelligence* (America Psychological Association. Washington, DC, 2
24. J. S. DeLoache *et ol.*, *Psychol. Sci* 21, 1570
25. A Diamond, W. S. Barnett, J. Thomas, S. Mu *Science* 318, 1387 (2007).
26. T. Tu, *Early Child. Educ. J.* 33, 245 (2006).
27. K. Conezio, L French, *Young Children* 57, 1
28. R. Gelman, K. Brenneman, *Early Child. Res.* 150 (2004).
29. National Research Council *Inquiry and the Nah Education Standards: A Guide for Teaching and* (National Academies Press, Washington. DC, 20
30. G Loewenstein, *Psychol. Bull.* 116. 75 (199
31. L French, *Early Childhood Res. 0.* 19, 138
32. I. Chalufour, K Worth, *Discovering Nature w Children* (Redleaf Press, St Paul MN, 2003)
33. J. Gropen, N. Clark-Chiarelli, 5. B. Ehrlich, Y paper presented at the Conference of the Sc Research in Educational Effectiveness, Washi 3 March 2011.
34. P. D. Pearson, E. Moje, C. Greenleaf, *Science* 328
35. S. M. Peterson, *Discourse Process.* 46, 369
36. J. Jirout, D. Klahr, paper presented at the bier of the Society for Research in Child Developme Province of Quebec, Canada. 1 April 2011.
37. S. Michaels, A. W. Shouse, H. A. Schweingruber *Science! Putting Research to Work in K-8 Scienc* (National Academies Press, Washington, DC, 2
38. W. M. Roth, *Educ. Res. Rev.* 3, 30 (2008).
39. S. Tobias, T. M. Duffy, *Constructivist Theory Ap Instruction: Success or Failure?* (Routledge, Nev
40. R. E. Mayer, Am. *Psychol.* 59, 14 (2004).
41. D. Bransford, A. L. Brown, R. R. Cocking, *How* (National Academies Press, Washington, DC, 2
42. P. A. Kirschner, J. Sweller, R. E. Clark, *Educ.* 75 (2006).
43. A Brown, J. *Leorn. Sci.* 2, 141 (1992).
44. American Association for the Advancement *Benchmarks for Science Literacy* (Oxford Ur Oxford, 1993).
45. D. Klahr, L M. Triona, C Williams, *J. Res. S* 183 (2007).
46. R. F. Lorch *et al.*, *J. Educ. Psychol.* 102, 9
47. D Dean Ir., D. Kuhn. *Sci. Educ.* 91, 384 (
48. D. Kuhn, *Education for Thinking* (Harvard U Cambridge, MA, 2005).
49. A Zohar, B. Peled, *Learn. Instr.* 18, 337 (