

Virtual vs. Physical Materials in Early Science Instruction: Transitioning to an Autonomous Tutor for Experimental Design

David Klahr¹, Lara Triona², Mari Strand-Cary¹, & Stephanie Siler¹

1, Department of Psychology, Carnegie Mellon University. Pittsburgh, PA

2. California State University, Fresno. Fresno, CA

The spread of computer based instructional materials makes it important to determine the relative merits and effects of virtual materials vs. physical materials in early science instruction. In this paper we first lay out a framework for comparing key aspects of this virtual-physical issue, and then we describe three studies addressing it. In two studies with middle school children we found that children using virtual and physical materials made equally large gains in their knowledge while learning a complex procedure (control of variables) under conditions of direct instruction (Study 1) and while learning about specific physical effects in a engineering design challenge in a discovery learning content (Study 2). These results suggest that simply replacing the physical materials with virtual materials does not affect the amount of learning or transfer when other aspects of the instruction are preserved. In the third study we describe our progress in creating a virtual tutor for teaching experimental design procedures and concepts to middle school children.

The widespread availability of computers makes them an appealing option for presenting instructional materials in laboratory science. The advantages of computer-based science instructional materials include portability, safety, cost-efficiency, minimization of error, amplification or reduction of temporal and spatial dimensions, and flexible, rapid, and dynamic data displays. However, there are claims that “virtual” materials are detrimental not only to the achievement of specific instructional objectives, but also to broader goals ranging from brain development to social development (Alliance for Childhood, 2000; Armstrong & Casement, 1998; Healy, 1999). Critics argue that virtual materials are ineffective for instruction because “hands on” manipulation of physical materials is essential for learning (Berk, 1999; Deboer, 1991; Diem, 2000). Advocates of hands-on manipulation of physical materials in science instruction argue that it promotes learning because (a) it is consistent with the concrete-to-abstract nature of cognitive development, (b) it provides additional sources of brain activation, and (c) it increases motivation and engagement (Flick, 1993; Haury & Rillero, 1994). Critics argue that physical materials (a) produce confusing and inconsistent feedback, (b) provide inadequate mappings between the behavior of physical

materials and their abstract representation in diagrams and equations and (c) tend to have higher logistical, financial, and temporal costs (Hodson, 1996).

In this paper we examine several aspects of this physical-virtual debate. First, we lay out some important dimensions of different types of studies in this area. Then we summarize two studies that compare the effectiveness of physical

Table 1: Instructional Space in Physical – Virtual Comparisons

		Knowledge Type in Learning Goal			
		Domain General		Domain Specific	
Instruction	Materials	<u>Direct</u>	<u>Discovery</u>	<u>Direct</u>	<u>Discovery</u>
“Physical” (Live instructor)	Physical	A	B	C	D
	Virtual	E	F	G	H
(Computer Tutor)	Physical	A'	B'	C'	D'
	Virtual	E'	F'	G'	H'

and virtual materials using two different approaches to instruction. Finally, we describe the status of an on-going project to develop a virtual tutor for an important topic in middle school science. The tutor will facilitate the comparison of the effectiveness of physical and virtual instruction.

Key attributes of the Design Space for Computer-Based Education (Table 1).

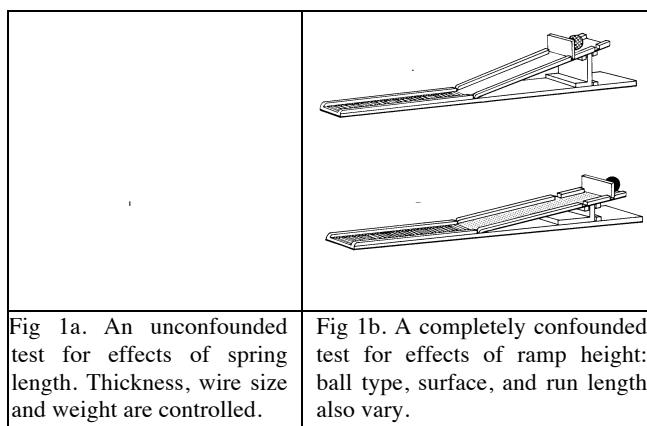
Instructional Materials: Physical or Virtual? “Physical Materials” include materials such as ramps, test tubes, plants, mechanical devices, chemicals, instruments, and electrical components typically found in science kits. “Virtual Materials” consist of computer programs under control of mouse and keyboard that display and enact animations or videos that depict the same range of actions that occur when the physical materials are used.

- *Instruction: Live or Virtual?* Instructional delivery itself can be either “physical” (a live teacher) or virtual (e.g., a computer tutor).
- *Learning Goal: Domain-General or Domain-Specific Knowledge?* Domain-general knowledge includes knowledge that transcends any particular branch of science, such as knowledge about the relation between theory and evidence (Kuhn, 2002). Domain-specific knowledge pertains to particular domains, such as physics, chemistry, or ecology (cf. Lehrer & Schauble, 2006).
- *Type of Instruction: Direct or Discovery?* Another contrast depicted in Table 1 involves the instructional context in which the physical or virtual materials are being presented: either as part of direct instruction, or in a discovery mode in which little explicit instruction is provided.

Inattention to the dimensions depicted in Table 1 makes it easy to conduct a confounded comparison between physical and virtual instruction. For example, a comparison between Cell C (live teacher, physical materials, domain-specific knowledge, and direct instruction), and Cell E (live teacher, virtual materials, domain-general knowledge, and direct instruction) would make it impossible to attribute any learning

differences only to the “physical vs. virtual materials” distinction because type of knowledge had also been changed. Similarly, comparing G to C’ would be uninformative because cell G corresponds to a live teacher using direct instruction with virtual materials (e.g., demonstrating a simulated chemical process), and cell C’ to a student receiving direct instruction from a virtual teacher, but handling physical materials (e.g., an instructional video for a hands-on science kit about plant growth).

STUDY 1: DIRECT INSTRUCTION ON EXPERIMENTAL DESIGN



Study 1 compared physical with virtual materials in the context of direct instruction -- contrasting Cells A and E (Triona & Klahr, 2003). Instruction for one group of 4th and 5th grade children used physical materials, and instruction for the other group used virtual materials that were otherwise identical to the physical materials. All other variables were the same in both conditions.

ables-were the same in both conditions.

The topic was how to design simple unconfounded experiments, also known as the control of variables strategy (CVS). CVS includes both the rationale and the procedure for creating unconfounded experimental contrasts. Figure 1a shows an example of an unconfounded experiment for the target variable of spring length in which all other variables --spring width, wire thickness, weights -- are set to the same level. Figure 1b shows a confounded experiment with ramps. The procedure was based on a highly effective CVS training study (Chen & Klahr, 1999). In one condition children manipulated physical materials (Springs) while setting up simple experiments and in the other they ran their experiments by pointing, and clicking on a computer simulation that

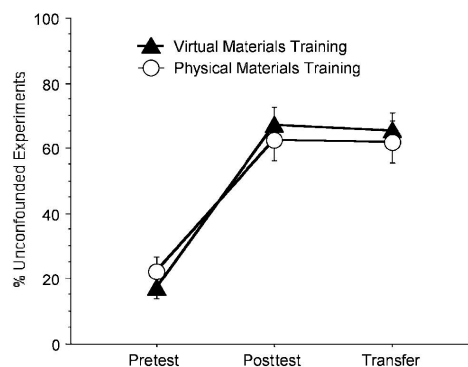


Figure 2. Mean proportion of unconfounded experiments for each phase separated by training condition with standard error bars.

presented video depictions of the physical springs and weights. Children were assessed at pre-test and immediately after Direct Instruction on Springs. A week later they were tested in another domain (Ramps). For this transfer assessment, children in *both* conditions used physical materials.

The two types of training were equally effective (Fig. 2). Moreover, the children trained on *virtual* materials performed as well with the *physical* materials in the transfer phase as children who used physical materials for all phases.

STUDY 2: ENGINEERING DESIGN WITH “MOUSETRAP CARS”

In this study with 7th and 8th grade children, we compared physical vs. virtual materials in an engineering design challenge using discovery learning -- thus contrasting Cells D and H (Klahr, Triona, & Williams, 2007). The task involved “mousetrap cars”: small mobile cars -- powered by an ordinary mousetrap. Children had to discover the combination of features that produced a “maximum distance” car. No instruction was provided about how to approach the comparisons between one design and another beyond

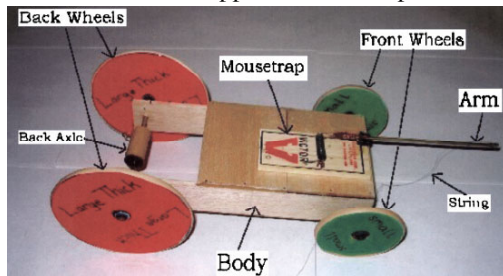


Figure 3. An assembled mouse-trap car.

explaining how to construct specific instances and the intended goal of the investigation. One group of children worked with physical cars (Fig. 3). They selected various components, assembled cars from them, and then ran the cars to see how far they would go. The other group constructed virtual cars (Fig. 4) by “pointing and clicking” to

select components, assemble cars, and then “run” them in a virtual window.

Physical mousetrap cars. Each car was assembled by choosing from two different bodies, two different back axles, three different back wheels and front wheels. The car was energized by winding the string around the back axle. As the mousetrap spring returned the arm to its initial position, the rotating axle propelled the car.

Virtual mousetrap cars. The virtual display for the assembly and testing of mouse-trap cars was designed to be as sparse as possible (see Fig. 4)..

Knowledge assessment. Before and after the assembly and test phase, children completed a questionnaire that assessed their knowledge about the features that contributed to a car’s travel distance. We used a 2 (material: physical or virtual) X 2 (constraint: 20 min or 6 cars) X 2 (test phase: pretest vs. posttest) factorial design.

Our primary question was whether gains in children's knowledge about the "best" value of each factor (body length, axle width, front and back wheel diameters) for maximizing travel distance would be different in the physical or virtual conditions. Children's initial knowledge was better than guessing and it increased

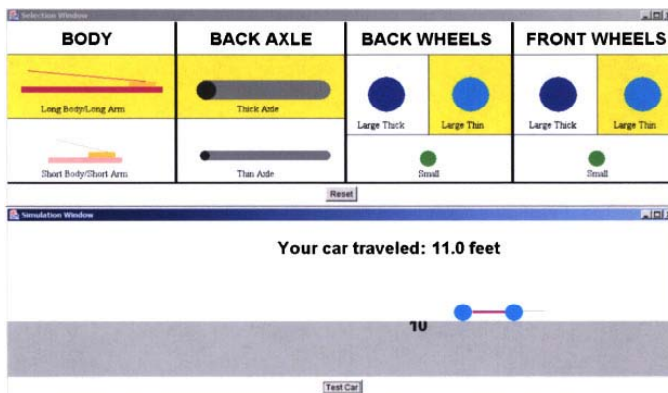


Figure 4. The virtual mousetrap car display. The highlighted panels at the top indicate that the student has constructed a car with a long body, thick back axle, large thin back wheels, and large thin front wheels. The bottom panel displays an animation of the car moving and its final distance.

significantly from pretest to posttest (Fig 5). A repeated-measures ANOVA on children's knowledge, with phase as the repeated measure, showed a main effect for phase and no other main effects or interactions. That is, physical and virtual materials were equally effective in all conditions and phases.

STUDY 3: A VIRTUAL TEACHER FOR EXPERIMENT DESIGN

Studies 1 and 2 showed that physical and virtual *materials* were equally effective, but both studies used "physical instruction": that is – humans. In Study 3 we are designing a system in which direct instruction on a domain general procedure will be provided in a context where both the instructor and the materials are virtual (thus situating it in cell E'). Our goal is to increase elementary school children's understanding of CVS and close the gap between high- and low- SES students on this crucial component of science education. It is an

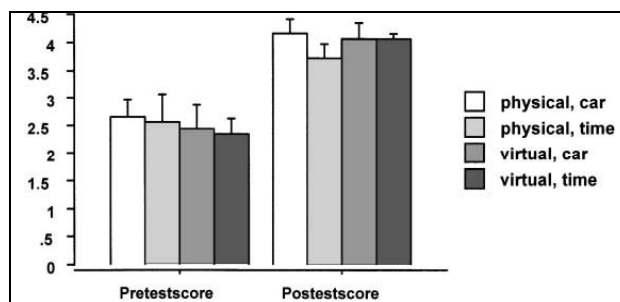


Fig. 5 Mean number of correct answers in all conditions

ongoing development project – a design experiment -- in which we are building an intelligent tutor for teaching experimental design procedures and concepts to middle school students.

Once “TED”, (for “Training in Experimental Design”), is operational, we will be able to compare it with a live teacher using the same virtual materials (thus contrasting cells E and E’). TED will provide adaptive instruction based on individual learners’ knowledge and mastery in real time across a variety of tasks and science content domains. Each student will receive instruction adapted to TED’s assessment of that student’s specific needs. Such individualization will allow students who already know the basics of experimental design to review, fill gaps in their knowledge, apply their knowledge to more challenging contexts and, ideally gain familiarity with new, advanced concepts (e.g. interactions, reliability, validity, bias, sample selection). Students who show little knowledge of experimental design will receive instruction that, from the outset, addresses their particular misconceptions, biases, and lack of knowledge in order to most effectively move them through the basic lesson.

The capacity to accurately and efficiently provide fine-tuned, student-specific, “proven” instruction and remediation in a full-class setting is what largely sets the TED tutor apart from what teachers are able to do. Its ability to choose the “next step” at any given moment will depend, in fact, on having an explicit “understanding” of the student’s current mental state. By contrast, teachers simply cannot “be in the heads” of every one of their students simultaneously. A human teacher must choose some level or style of instruction that is expected to meet the majority of a class’s needs, whereas TED will be designed to respond to individual misconceptions and procedural bugs in each student’s understanding of CVS (see Study 1).

At first glance, it might seem that a computer tutor is not necessary to teach such a basic skill as CVS. Procedurally, all that CVS requires is for the student to identify and vary the focal variable and to control the others. One might think a short lecture would be enough to convey this to students. However, our prior studies show that, even in upper-middle class U.S. schools, there is always a non-trivial proportion of children who do not learn from our instruction. Moreover, in our work with schools with less advantaged children, we find much higher rates of non-learning and fragile learning (Klahr & Li, 2005). Many students find it difficult to master CVS in normal classroom settings and one-on-one tutoring is required to overcome their difficulties with such factors as remembering the basic premise/goal of an experimental contrast, identifying and maintaining focus on the tested variable, and understanding the reasons why such procedures are necessary in order to make valid causal inferences. Even students who seem to understand CVS in relatively “abstract” or “remote” contexts often revert to ineffective approaches when tested in a familiar domain in which they likely have causal biases.

During the first 18 months of this project, we “virtualized” our materials and used them along with “physical” instruction in either full-class or small group settings – in effect, working in cell E of Table 1. Then we used the WOZ process -- simulating a computer tutor by having researchers provide instruction primarily using the interface (comprised of lesson goals, vocabulary instruction, an interactive ramps simulation, and explicit presentation of experimental procedures and concepts). The virtual instructional components and simulations were supplemented with a) discussion or questioning tailored to each student’s current knowledge level and struggles and b) researchers’ selections of problems from a pre-determined set of domain-specific paper/pencil problems (comprised of different reading requirements, domains, foci, and difficulty levels).

Next, we adapted the interface for use directly by individual students. This Wizard of Oz (Molin, 2004) version required a student to work at one computer and a researcher at another. Based on student interactions with the current pre-programmed instruction, feedback, scaffolding, help, or practice problem, the tutor presented to the researcher its best guess about appropriate instruction to offer the student next. In turn, the researcher (who had been monitoring student actions on her computer screen) either confirmed the tutor’s choice or chose from other available options. Periodically, the researcher also rated the student’s current level of understanding. We anticipate this process will enable us to fine-tune our ultimate decision rules about the next instructional event, in order to tailor the level of feedback that is appropriate throughout the tutoring process.

Repeated iterations of this WOZ will enable us to assess our hypotheses about good instructional strategies, ensure that the majority of student actions, misunderstandings, etc. can be addressed by the tutor, and work out the “bugs” of the system. Our final year of tutor development will be devoted to ensuring that TED can fully and independently guide instruction for students with widely varying levels of understanding.

Though much of our early development work required small-scale piloting, eventually the tutor will be used in full-class sessions where students work at their own pace, receiving individualized instruction. That is, we aim to produce a system that fits into cell E’ in Table 1: a virtual instructor, using virtual materials, to provide individualized direct instruction and practice on domain general knowledge. The teacher will thus be freed to assist those students struggling for reasons that could not be addressed by TED or to provide even more advanced, individualized instruction to those students who learn the content of the tutor very quickly.

DISCUSSION

We found that physical and virtual materials are equally effective when 4th and 5th graders are learning a complex procedure (control of variables) under conditions of direct instruction (Study 1) and when 7th and 8th graders are learning about a simple me-

chanical device in a discovery learning content (Study 2). However, in both studies, instruction about the task itself was provided by a human. In Study 1 this was extensive direct instruction, and in Study 2, the context and general procedure was delivered by a human instructor. Study 3 is now attempting to determine the extent to which instruction can be completely virtualized in the context of teaching a domain-general procedure, across a broad range of specific contexts.

This work has implications for the extensive debates in the education literature about “hands on science”. All of our studies have a condition in which the young science learners’ hands are on virtual rather than physical materials. This is an important contrast because most recommendations about hands-on science *exclude* computer simulations and virtual labs from their definition of “real” hands-on activities. For example, in the USA, the National Science Teachers’ Association recommends that “computers should enhance, but not replace, essential ‘hands-on’ laboratory activities” (NSTA, 1999). However, we have not yet found any difference in learning or transfer whether the children’s hands are on virtual rather than physical materials.

REFERENCES

- Alliance for Childhood. (2000). Fool’s Gold: A critical look at computers in childhood. Retrieved from http://www.allianceforchildhood.net/projects/computers/computers_reports.htm
- Akpan, J.P. (2002). Which comes first: Computer simulation of dissection or a traditional laboratory practical method of dissection. *Electronic Journal of Science Education*, 6. Retrieved July 6, 2005 from: <http://unr.edu/homepage/crowther/ejse/akpan2.pdf>
- Armstrong, A., & Casement, C. (1998). *The child and the machine: Why computers may put our children’s education at risk*. Toronto, Ontario: Key Porter Books.
- Berk, E. G. (1999). Hands-on science: Using manipulatives in the classroom. *Principal*, 78, 52–55, 57.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70, 1098–1120.
- Deboer, G. (1991). *A history of ideas in science education*. New York: Teachers College Press.
- Diem, K. G. (2001). Turn any science topic into hands-on fun. *Science Scope*, Apr-01, 46–49.
- Flick, L.B. (1993). The meanings of hands-on science. *Journal of Science Teacher Education*, 4, 1–8.
- Haurly, D.L., & Rillero, P. (1994). Perspectives of hands-on science teaching. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.
- Healy, J. M. (1999). *Failure to connect: How computers affect our children’s minds—and what we can do about it*. New York: Simon & Schuster.
- Hodson, D. (1996). Laboratory work as scientific method: Three decades of confusions and distortion. *Journal of Curriculum Studies*, 28, 115–135.
- Klahr, D. & Li, J. (2005) Cognitive Research and Elementary Science Instruction: From the Laboratory, to the Classroom, and Back. . *Journal of Science Education and Technology*, 4, 217-238.
- Koedinger, K. R., Anderson, J. A., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to the big city. *International Journal of Artificial Intelligence in Education*, 8, 30–43.

- Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.), *Blackwell Handbook of childhood cognitive development* (pp. 371–393). Malden, MA: Blackwell.
- Lehrer, R., & Schauble, L. (2006). Scientific thinking and science literacy. In W. Damon, R. Lerner, K.A. Renninger, & I.E. Sigel (Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice (6th ed.)* Hoboken, NJ: John Wiley & Sons.
- Molin, L. 2004. Wizard-of-Oz prototyping for co-operative interaction design of graphical user interfaces. In *Proceedings of the Third Nordic Conference on Human-Computer interaction* (Tampere, Finland, October 23 - 27, 2004). NordiCHI '04, vol. 82. ACM Press, New York, NY, 425-428
- National Science Teachers Association. (1999). Position statement on the use of computers in science education. Retrieved July 6, 2005 from: <http://www.nsta.org/positionstatement&psid¼4>
- Triona, L.M., & Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition & Instruction*, 21, 149–173.

Acknowledgments This work reported here was supported in part by grants to David Klahr from NICHD (HD25211), the James S. McDonnell Foundation (CSEP 96-37), the National Science Foundation (BCS-0132315) and the Institute of Education Sciences (R305H060034).