

1. Introduction

Developments from the learning sciences move slowly, if at all, into educational practice. Consider the low adoption of spaced practice of learned material, a robust finding first observed by Hermann Ebbinghaus in 1885 and validated at scale in classrooms with thousands of students in 1939 (Ebbinghaus, 1913; Spitzer, 1939; 1939; Whitehurst, 2003). Students still cram. Teachers still march linearly through curriculum and rarely repeat assessments. Why is it so difficult to implement research findings into practice? Burkhardt and Schoenfeld argue that “part of the reason is that the traditions of educational research are not themselves strongly aligned with effective models linking research and practice” (2003).

The field of human computer interaction research, like many applied sciences, has also grappled with the gap between research and practice (Buie et al., 2010). In this dissertation I draw on methods common in HCI to describe an emerging model of linking research to practice in education: the operant probe. I adapt the HCI methods user experience design to an education-focused *learner experience design*, first exploring learner experiences through sketching to map out design opportunities. Then from this map I designed two operant probe systems to support the practice of studying: Nudge for allocating study time and Exemplify for better learning from example solutions. In the following chapters, I describe the potential benefits of operant probe development, the fieldwork that inspired the systems, the design iterations to create them, and the formal evaluation to rigorously validate the design decisions. I then return to examining the process. But first, I elaborate how research in education can benefit from new design processes.

1.1 Motivation

1.1.1 Education is important to improve

Quality education is critical to modern society. The cognitive skills of a population are powerfully related to their individual earnings, distribution of income, and economic growth (“The Role of School Improvement in Economic Development,” 2007). While continued economic growth requires growth in cognitive skills, the United States educational system is in decline. In 1983, the Reagan administration published *A Nation at Risk* (United States National Commission on Excellence in Education, 1983) and in 2012 the Council on Foreign Relations reports that the poor state of American education threatens not just US prosperity but its national security (Klein & Rice, 2012). The comparative decline of US education is most acute in higher education. While America is ranked first in the world in college degrees per capita for people aged 55 to 64, for ages 45 to 54, it is third and in ages 25 to 34, it has fallen to 10th place globally. While this is a decline in rank, it is mostly due to other countries improving their education while the US has been stuck at 40% college completion for decades, despite large scale reforms such as the No Child Left Behind act (*No Child Left Behind Act*, 2002).

1.1.2 Education is difficult to improve reliably

Part of the challenge is how difficult it is to know what proposed changes work in practice. This is due in part to the history of education research in the US. Quoting from the U.S. Department of Education's Strategic Plan for 2002–2007 (2002, cf. Burkhardt & Schoenfeld 2003):

Unlike medicine, agriculture and industrial production, the field of education operates largely based on ideology and professional consensus. As such, it is subject to fads and is incapable of the cumulative progress that follows from the application of the scientific method and from the systematic collection and use of objective information in policy making. We will change education to make it an evidence-based field. (p. 48)

Since then the federal Institute for Educational Science, established by the No Child Left Behind Act, has fostered an emphasis on determining “what works”. They have defined a gold standard for research, randomized controlled trials. Yet, many of the studies they fund fail with “no effects”. Researchers and other experts question the design of these studies, running up to \$14.4 million. Scholars worry that even when study results are positive, they do not carry over into other educational settings. Policy makers want to know, “What will work in my school?” (Viadero, 2009).

Part of the difficulty in reforming education research is the tension between understanding how learning works, understanding how to improve learning, and actually improving learning. Let us return to the example of spaced repetition. Why isn't it used more in practice? We understand learning enough to know that distributed (spaced) practice has better long term learn effects than massed practice (Committee on Developments in the Science of Learning, 2000; Karpicke & Blunt, 2011). We also know that we could improve learning by increasing the amount of distributed practice by students. However, the next step is the hard part. Students know massed practice by another name, cramming, and when they are told it is not the best strategy for studying, they continue to use it. The classroom environment makes that the easiest option for most students. It doesn't require planning or self-regulation; it let's them avoid confronting the limits of their knowledge; it isn't facilitated by the instructor; and it doesn't affect their grades much as they still do as well on the assessments they're given.

1.2 Scientific Research through Interaction Design

How do you replace cramming with more effective study strategies? The answer is not a matter of just science but also of design. Science is powerful because of its ability to generalize, through nomothetic descriptions. In this case, the science falls short of solving the problem of shaping student learning. To change a student' study strategies requires recognizing and fitting with the fuzzy factors that influence that student's behavior in her specific environment. Design is an idiographic tradition, which tends to specify and understand the meaning of contingent phenomena in order to change a current state of the world to a preferred state. To use the power of scientific theories of learning to improve actual learning requires creating working solutions that operationalize basic research into practice. To evaluate the operationalization requires then evaluating those artifacts and their features for

Introduction

their ability to effect the desired changes. Through many specific designs, generalizations can be developed for normalized solutions and theories of implementation. I describe this approach as Scientific Research through Interaction Design, building on the concept of Research through Design in the HCI literature (J. Zimmerman, Forlizzi, & Evenson, 2007).

1.3 Operant Probes

Further, we as a field need these artifacts to be designed to operate in the “real world”, as Fishman et al. contend:

Why are cognitively oriented technology innovations not widely used in schools? Why aren't they scaleable or sustainable? We believe an underlying explanation to be that we, as a scholarly community, have not focused our research on the development and use of cognitively oriented technologies in a way that addresses the fundamental needs of school systems. Instead, research on cognitively oriented learning technologies has focused primarily on students, teachers, and classrooms as the primary unit(s) of analysis. Though we recognize the need to link technology and reform, the field lacks a bridge between focused research and development of learning technologies and the broad-based systemic use of these innovations in schools. Shepard (2000) recognized this as problem for the broader educational research community in her AERA Presidential Address, when she advised researchers to develop methodologies that embrace “dilemmas of practice.” Such work “would advance fundamental understandings at the same time that they would work to solve practical problems in real-world settings” (p. 13). This focus would lead to the production of more readily “usable knowledge” (Lagemann, 2002). As researchers, we have developed rich understandings of how technology can foster learning in specialized situations; we now need to develop knowledge about widespread appropriation and use of cognitively oriented technologies by schools and school systems as part of real-world reform efforts. (Fishman, Marx, Blumenfeld, Krajcik, & Soloway, 2004)

In this dissertation I developed two cognitively oriented learning technologies that operationalize theory into real-world contexts. The resulting designs are not just “usable knowledge” but “usable systems”. Moreover, their use serves to help inform scientific theory. I classify them as “operant probes”, a term and type of research contribution that I motivate and define in Chapter 2. Operant probes form part of an emerging paradigm of research, using web technologies and scale to design research artifacts that operate *in vivo* and provide the controls and data collection needed for rigorous quantitative research.

1.4 Process

This work explores methods for designing operant probes in the approach of Scientific Research through Interaction Design, adapting HCI methods to education research. The process can be organized by the stages of Bannan-Ritland’s Integrative Learning Design Framework (Bannan-Ritland, 2003). In this framework, the design process begins with Informed Exploration of the design context, followed

Introduction

by Enactment and then Evaluation of Local Impact. Finally, the design may be evaluated for Broader Impact. The designs in this work have not yet been evaluated for Broader Impact but I will speak to the aspects that support confidence in their suitability to transfer into other contexts.

1.4.1 Informed Exploration

To develop systems that fit into real-world contexts requires a rich understanding of those contexts. Because my goal is to develop new types of probes to open new opportunities for research, I begin with an exploration of opportunities in a specific context. Finding opportunities does not require a formal method, but here I offer a reliable method, adapted from the validated best practices of HCI, to rapidly identify new opportunities. I extend HCI user experience design to the unique challenges of *learner experience design*, evaluating these opportunities by both learner impact and contributions to accretive education research.

I chose large college lecture courses as the context of inquiry. Lecture courses are presently the dominant way that the 20 million college students in the US are taught. As a centuries-old mode of instruction they are ripe for innovation. The goal here was to find opportunities to directly improve learning for a large number of students. This opportunity mapping approach to the Informed Exploration is fully described in Chapter 3. In this phase I identified two opportunities for which to design: 1) helping students to better allocate their study time and 2) provide students with more immediate feedback on their learning.

1.4.2 Enactment

With the two design goals, I iteratively developed two software systems, Nudge and Exemplify. Nudge (attempts to improve how student allocate their study time by decomposing the course syllabus and adding explicit tasks with due dates (such as “Study for the upcoming exam” one week before it takes place.) The small tasks are sent to the student by email when they are due and students can indicate their progress on the task. Exemplify enhances traditional answer keys with an interactive activity to scaffold how students learn from them.

From storyboards evaluated through user interviews to working prototypes evaluated in pilot classroom trials, each design decision was weighed between its ability to operate on the environment to achieve the desired outcome and to probe the environment and its use to advance the science of learning. As operant probes, part of the design process was to refine an understanding of the implications and limits of the evidence in general theory and local observations for each design decision. When design questions could not be satisfactory resolved by general theory or local observations, they were identified as candidates for resolving empirically using probe variants. In this work, a tension was identified in the practical recommendations of the literature on the “worked example effect” in cognitive load theory versus the “testing effect” in memory theory. To help resolve this both for the design of this system and future similar systems, a variant of Exemplify was produced for each of the two competing predictions of the theories in

Introduction

practice. In one variant of the tool, the problem-solving prompt was removed so that it was strictly a worked example. The full iterative development process of each is described in the chapters on Nudge and Exemplify (Chapters 4 and 5).

1.4.3 Evaluation

With the two operant probes fully developed, the next phase was to evaluate their efficacy through an experiment. The experiment took place in a large college introductory chemistry course consisting of two similar sections. The larger section was chosen as the Experimental section and received both Nudge and Exemplify while the other, control section, had neither. Within the experimental section, students who opted into the study were randomly assigned to a Nudge condition (Nudge required or unavailable) and an Exemplify condition (including prompts to solve or not). Nudge and Exemplify both fit well to the context. Students used both systems voluntarily through the whole semester, including students not in the study.

Both Nudge and Exemplify affected student learning measures. Nudge interacted with students' time management skills to better aid students with worse time management. Exemplify provided big gains on robust learning, supporting the testing and proceduralization over worked example effects in practice. On immediate measures, students with the variant that prompts to solve performed better than students both with the nonsolving control variant ($d=.35$) and business-as-usual control section ($d=.26$). On delayed measures, the effect was roughly a full letter grade over the nonsolving control variant ($d=.77$) and business-as-usual control section ($d=.40$).

On the data production measures, Nudge provided data on student activities that could be used to model student study practices. Exemplify logs provided data that helped explain the mechanisms of its effects. The full analyses of each system are detailed in their respective chapters (4 and 5).

1.4.4 Wrap-up

I do not set out in this work to answer a specific theoretical question. Instead, this work is to improve the practice of education by designing 1) technological artifacts that enact learning science principles to effect learning objectives in a specific natural context and 2) evaluations of the artifacts to inform future applications of said principles and the principles themselves.

In doing so, this work contributes to processes of design research in education and to specific design implications for two classes of technologies for education. The primary contribution to design research broadly is the articulation of the operant probe as a productive research artifact and the scholarship to situate it within existing design practices and research issues (Chapter 2). A related contribution is the reflection on the learner experience design methods suitable to designing operant probes, chiefly mapping of opportunities (detailed in Chapter 3). With these processes, I developed two systems that operationalize and thus inform the real-world application of theory. Nudge informs the potential for supporting student

Introduction

time use and the contextual utility of theoretical principles of motivation (Chapter 4). Exemplify informs the potential for supporting students learning from worked solutions and the contextual utility of theoretical principles of cognitive load, practice and proceduralization. Further, the evaluation of Exemplify experimentally measures the relative utility of competing theoretical principles when put to use (Chapter 5).

I argue that through an analysis of existing methods and reflection on my design processes and artifacts, I will demonstrate a new and effective approach to design research in education. Stated explicitly:

The Scientific Research through Interaction Design approach can enact preferred states in a manner that explains outcomes, informs the conditions for applying scientific theory, and generates new experimental hypotheses.

In Chapter 6, I evaluate the success and limitations of this work in supporting this thesis statement by reflecting on the cases of Nudge and Exemplify.