RUNNING HEAD: A Video Methodology for Experimental Research

Bridging Cognitive Science and Real Classrooms:

A Video Methodology for Experimental Research in Education

Kreshnik Nasi Begolli

Temple University

Lindsey Engle Richland

University of Chicago

The Journal of Experimental Education (accepted 2017)

Abstract

We describe a new approach to the use of video-based technology for conducting controlled experiments in classroom contexts. Specifically, we describe a process for editing video recordings of live classroom lessons to create multiple versions, such that only one aspect of the lesson is systematically varied. Other aspects of the instruction are all held constant, including audio, curricular content, student participation, and other notoriously hard to control details of the interactional context that nevertheless impact learning (e.g. gestures, affect). These edited lesson versions can be randomly assigned to students by condition within classrooms to meet a high standard for random assignment. This technology provides opportunities for deriving causal data on the efficacy of teaching practices through stimuli approximating a typical everyday classroom context.

Keywords: Technology, Video, Classroom Experiments, Video stimulus, Video lessons.

Bridging Cognitive Science and Real Classrooms: A Video Methodology for Experimental Research in Education

Bridging experimental psychological studies with classroom needs, interests, and contextual dynamics is challenging. In the past decade, there has been a surge of important experimental work in cognition that has shifted from the laboratory to the classroom (Carpenter, Pashler, & Cepeda, 2009; Klahr & Li, 2005; Koedinger, Aleven, Roll, & Baker, 2009; Rittle-Johnson & Star, 2009; Schwartz & Bransford, 2005; for review see Mayer, 2008; Pashler et al., 2007; Richland, Linn, & Bjork, 2007; Roediger & Pyc, 2012), but controlled examinations of instructional manipulations executed by a teacher are less frequent (for critiques of existing methodologies, see: Design-Based Research Collective, 2003; McCandliss, Kalchman, & Bryant, 2003) and are often misaligned with laboratory-based studies due to differences in operational definitions (Klahr, 2013). While teachers continue to play the most central role in everyday classroom instruction, many psychological experiments seek to minimize these complications and thus aim to derive abstractions by manipulating software or paper executed interventions, due to concerns about control and reliability. However, this often minimizes the role of the teacher in the process (Daniel, 2012). Methodologically, conducting carefully controlled work in dynamic classrooms in ways that can inform theoretical questions of both teaching and cognitive mechanisms poses challenges in terms of preserving high internal and external validity.

We describe a methodology for using videotaped classroom instruction as experimental stimuli, which has the potential to make great strides toward this aim. Edited videotapes are used as a basis for experiments in which new students learn from matched versions of the recorded, and thus well-controlled, instruction. Video provides an efficient, reliable, and relatively inexpensive technological tool for creating experimental stimuli that approximate the real world. Video stimuli can serve to answer both applied and theoretical questions since it provides a mode for standardizing classroom-relevant instruction across conditions.

Further, the methodology can provide both high internal and external validity. When conducting research to both inform theory and practice, there is often a trade-off made between external validity and internal reliability. We posit that video technologies provide an opportunity to create controlled stimuli embedded within the contextual variability that is integral to real classroom lessons, making this tradeoff less severe. Many of these hard-to-control features of the classroom context have been shown to impact learning, such as teacher and student gesture, affect, or discipline (e.g., Alibali, Flevares, & Goldin-Meadow, 1997; Pardos, Baker, San Pedro, Gowda, & Gowda, 2013). Thus, making experimentally-derived causal claims about classroom learning principles without incorporating these multi-faceted aspects of classroom contexts in the stimuli may be failing to account for the complexity of the interrelationships between cognitive principles. This excluded variability may also in part explain regular failures for laboratory-based principles to generalize with large effect sizes to classrooms (Donovan, 2013; Dunlosky & Rawson, 2012). These challenges in part echo the concerns of many cognitive scientists whose ultimate goal is to understand the workings of the brain in the real world, but who are unable to account for the interrelated, highly complex brain activation that occurs in response to dynamic, real-world contexts (e.g. for discussions on the topic in memory see Kvavilashvili & Ellis, 2004; in attention see Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003).

An alternative model of studying learning in classroom contexts is to use iterative designs to formulate, test, and refine optimal curriculum and instructional stimuli. These projects embrace the complexity of the classroom context and are much better able to describe the complex interplay between an instructional manipulation and student thinking (Sawyer, 2014). At the same time, this work is not designed to maintain internal reliability, and thus causal claims about the efficacy of instructional choices are not part of the theory building enterprise.

We describe a mode through which video may be used to bridge these endeavors, such that it enables the researcher to design experimental materials that are well controlled but that also include more of the ecologically valid variability that is part of everyday classroom learning contexts. Our goal is to more fully meet Brown's (1992) vision to "traverse between the real world and the laboratory" and better test the causal questions arising from cognitive scientific theory as well as classroom-based qualitative designs (Design-Based Research Collective, 2003; McCandliss et al., 2003).

INSERT TABLE 1 HERE

The manuscript will describe this model for creating video-based stimuli for instructional experiments. We explain the rationale for studying cognition in the classroom context and provide a brief overview on the usage of video to capture realistic settings. Afterwards we walk the reader through the process of the methodology, present a case study (Begolli & Richland, 2016) to provide the technical details involved in video-recording and video-editing and conclude this section with some of the limitations of utilizing video-stimuli. Finally, we compare the video methodology with current trends in classroom experimentation and present the benefits of the video methodology as an alternative with respect to current research approaches to suggest how this methodology provides opportunities that overcome the limitations in either internal or external validity with implications for informing teaching practice.

The Importance of Capturing the Environment in Stimuli

While experimental control means that all aspects of (two or more) conditions are held constant except for the factor being intentionally varied, many psychological studies that aim to reveal basic cognitive or learning processes tend to also minimize the complexity of the tasks that are encapsulated within these conditions. These studies often involve learning materials that are isolated from their learning context (e.g., the broader curriculum), as well as from other interactional factors such as talk, or even perceptual complexity such as visual complexity. These types of studies, when conducted to examine cognitive underpinnings of learning or memory, are often used to provide evidence for how to improve educational materials or classroom learning, and do provide important insights, but must be replicated in everyday educational settings (Pashler et al., 2007).

The need for conducting research in ecologically valid settings before making broad generalizations about best educational practices is well recognized. At the same time, there are a growing number of cognitive studies that go even further, suggesting that even understanding basic mechanisms of cognition may be enhanced through experimental materials that more closely simulate the settings in which those cognitive resources will be deployed. Cognitive processes may engage differently depending on context (e.g. see Duncan & Owen, 2000; and Monin et al., 2014, for review on brain imagining studies tying stability of cognitive processes with context).

Increasing the ecological validity of stimuli could include imbuing even the controlled aspects of stimuli with more of the complexity that is part of typical everyday contexts. Ecological psychology studies have long argued that reasoning processes are grounded in the modality, action, and perspective of the thinker and the environment (Carraher, Schliemann, & Carraher, 1988; Lave, 1988; Saxe, 1991; Suchman, 1987). Even basic processes such as perception emerge as a result of the interaction between environmental factors and the thinker (e.g. Gibson, 1979). This notion has been framed in several lines of theories of cognition including embodied (e.g. Barsalou, 1999), distributed (e.g. Clark & Chalmers, 1998; Hutchins, 2000), situated (e.g. Greeno, 1998), and socio-cultural cognition (e.g. Correa-Chavez & Rogoff, 2005). Inherent in these ideas is that the environment-thinker interactions include events occurring in both the thinker and the environment. As such, there is mounting evidence that cognitive processes of basic perception (e.g. Gibson, 1979), attention (e.g. Simons & Chabris, 1999), and memory (Barnier, 2012; Kvavilashvili & Ellis, 2004) operate differently in response to an event (in our case a classroom lesson) than in response to stationary stimuli. These close relationships suggest that even minor changes within a laboratory context could impact the replicability of an effect (e.g. Atchley & Kramer, 2001; Berry & Klein, 1993; Bindemann, Mike Burton, & Langton, 2008)).

The difficulty of replicating rigorously tested, published effects is becoming a growing challenge to psychological science, but part of the challenge may be that the dynamic role of context in cognition is currently under-theorized. This might include the participant sample (see Henrich, Heine, & Norenzayan, 2010), timing in participants' broader lives (e.g. participating in experiments occurring during midterm week), or other features of the interactional context. The Open Science Collaboration's (2012, 2015) examination of 100 experimental and correlational studies published in top journals within psychological science revealed a low replication rate in regards to significance testing (a fall from 97% of studies showing significant results to 36% in the replication sample), and a fall in the magnitude of the mean effect size. Most predictive of a successful replication was the effect size of the original result, suggesting that effect sizes are an important measure. Effect sizes may also reveal which psychological phenomena are likely to occur regardless of contextual changes (larger effect sizes) and those that may be dwarfed by contextual factors (smaller effect sizes). Even if the specific features of the experiments were held constant, features of the setting typically not reported but that could be important include time of year, concurrence between study tests and college testing, city traffic, etc. These concerns suggest that using methodological approaches to increase the ecological validity of experiments on cognition may improve the likelihood of capturing cognitive processes in the way that they are executed on an everyday basis, possibly improving replicability in the contexts in which the research designs are aiming to simulate.

Video-Methodology to Improve Ecological Validity

Video methodology provides one mode for increasing the ecologically valid “factors”, which are held constant across experimental studies of cognition, enabling a more realistic person-environment interaction. While not precisely identical to the interaction of everyday lives, the use of video-based stimuli also maximizes the comparability of the person-environment interaction across conditions.

This methodology also allows for the use of highly active, motion-based stimuli. One common difference between laboratory and real world contexts for cognition is the static versus active nature of the stimuli used to prompt a studied cognitive change. Laboratory, and controlled classroom, experiments frequently use static images or written text to provide a highly controlled stimulus prompt in order to more carefully examine the cognitive responses to variations in the stimulus. At the same time, this raises several problems. First, in a laboratory, one may find that processing for static materials is quite different and not generalizable to the dynamic, more complex processing that emerges when learning stimuli are *active lessons* unfolding in a context such as a classroom. Secondly, even when static materials are presented in a controlled way, providing them in a dynamic setting such as a classroom means that the experimenter then cannot control what the reasoner does in response to these materials, and how they talk or act to engage with them.

We highlight therefore that video-lessons: a) involve complex linguistic and interactional information in addition to the cognitive prompts, b) carry normative classroom cultural information, and c) carry dynamic social and perceptually rich information.

Thus, video-lessons are more likely to stimulate the complex cognitive work engaged when students are learning in an everyday classroom to a greater extent. Further, video-lessons have the following benefits. 1) They can be used to manipulate specific cognitive factors (described in the following section). 2) They approximate everyday classroom learning, requiring a smaller leap for generalizing from research findings to everyday instruction. 3) They can be administered in one-on-one computer delivery, meaning that research subjects' participation can be randomized within classroom or teacher and delivery remain highly controlled.

For many reasons, therefore, videotaped stimuli provide a rich methodology for bridging classroom studies and controlled experimentation. The idea of using video to capture the real world is not novel in behavioral experiments (Gibson, 1947) and education research (e.g. TIMSS 1999), however. Thus, before presenting the details of the proposed methodological approach, we present a brief review of video usage in observational and experimental studies.

From Video-Based Observations to Video-Based Experimentation

Video-based experimental stimuli have had widespread use for examining human and animal behavior, most pervasively in studies of visual perception where video has been deemed a favorable mode of stimulus for studying visual processing of motion, shape, texture, size, and brightness (see Gibson, 1947; Oliveira et al., 2000; Webb, Knott, & MacAskill, 2010). Other scientists have used it to investigate, for example, social behavior (Soble, Spanierman, & Liao, 2011), gestures (Alibali et al., 1997; Cook, Duffy, & Fenn, 2013; Valenzeno, Alibali, & Klatzky, 2003), health and clinical training (Kilduff, Hopp, Cook, Crewther, & Manning, 2013), and autism (Marcus & Wilder, 2009).

In education, however, video has been mostly employed as a data *collection* tool for the purposes of observing teacher or student behaviors (e.g. TIMSS, 1999), or as an artifact to be used in teacher professional development (see Rossi & Fedeli, 2016; Star & Strickland, 2008). In fact, most published guides on the use of video for research in education explain methods for conducting classroom observations, such as selecting appropriate video segments, conducting video analysis, and developing descriptive video coding schemes (e.g., Derry, 2007; Derry et al., 2010; Goldman, 2007; Santagata, Gallimore, & Stigler, 2005). Such observational videos, for example, have been used to create video clubs where teachers reflect on their teaching (van Es & Sherin, 2008) to investigate reflections between expert and novice teachers (Rich & Hannafin, 2008), and to assess teacher knowledge (Kersting, 2008) and efficacy (Hill et al., 2008).

Observational studies have revealed that many features of classroom interactions may affect student learning in subtle yet meaningful ways that cannot be well constrained in a teacher-led experimental lesson, such as teacher responses to errors (Santagata, 2005), or gestures (Valenzeno et al., 2003).The role of teacher and student gestures on learning is one such line of study. Observational studies have identified the powerful effects of gesture on learning, both teacher and student gesture, and experimental studies have clarified how these effects are related to cognitive change. Video studies in this research area have used videotaped instruction in which specific aspects of the stimulus are varied (Alibali et al., 1997; Cook et al., 2013; Sueyoshi & Hardison, 2005; Valenzeno et al., 2003).

Generally, the stimuli are recorded at two different time points, but share the same audio (Cook, et al., 2013) or multiple cameras record the same event from different perspectives either capturing gestures or not (Sueyoshi & Hardison, 2005). These approaches exploit the main advantage of video: capturing realistic details of teacher-student interactions (e.g. gestures, affect) while keeping this stimulus constant across participants. This way conditions can be randomly assigned by administering participants' different versions of the otherwise same videotaped lessons. However, this approach has largely been overlooked as a method for testing the effect of a particular teaching practice on student learning within a classroom setting. This could be due to the traditional difficulties of conducting research in schools and maintaining rigor when embedding research questions within a full classroom lesson. Also, these videos tend to be highly constrained and short, quite different from an everyday lesson (for exceptions see Begolli & Richland, 2016; Endres, Carpenter, Martin, & Renkl, 2017; Glogger-Frey, Fleischer, Grüny, Kappich, & Renkl, 2015; Richland & Hansen, 2013).

To promote the discovery of learning principles that optimize teaching practices, we next present a methodology for creating video-based stimuli that could contribute to the experimental study of scientific hypotheses with high internal and external validity.

INSERT FIGURE 1 HERE

A Video Methodology for Bridging Cognitive Research with Teaching Practices

Video raises the external validity of experimental stimuli from simpler materials such as paper or interactive text-based computer programs because it can capture more of the human context of real world settings, such as a classroom. For educational experimentation, video-lessons are also beneficial because they are able to facilitate high internal validity between experimental trials. They are potent tools for supporting high internal validity while including more of the complexity of real-world discourse because video remains stable every time it is played. Additionally, video-recordings of a single lesson can be used to create well matched control group stimuli through systematic video editing. Edited versions of a primary lesson could be used to create two versions of a lesson stimulus, for example two versions of the same lesson with the same unscripted, meandering classroom discussion, but in which one version of the video shows the board throughout the discussion while the other one does not. A third benefit of using such videos as controlled stimuli is that two or more versions of a video-lesson can be randomly assigned to students within a single classroom when watched on individual computers, reducing the potency of teacher variables on outcome data.

We next provide details about a methodology for how one might use and implement video-lessons as stimuli in experimental designs. The process is illustrated in Figure 1, and described in detail here. Of course, one may modify based on one's research needs, but this provides a technique for creating a classic experimental design using classroom-based video in an experimental design.

Because we are highlighting experimentation, the methodological process begins with a hypothesis, which may draw from the experimental, basic research literature, from video or field observations of classroom practice, from shared professional knowledge of teaching practices, or a combination of these. For instance, a specific hypothesis could be: should the teacher keep multiple solutions to a single problem visible on the board throughout the lesson?

Next, the researcher collaborates with a classroom teacher to co-design a lesson which incorporates aspects of the lesson that will allow the researcher to test the hypothesis of interest. For that example hypothesis, it would mean that the teacher conducts the lesson while leaving all solutions visible on the board, but also explains them verbally, so that a student who did not see the board would be able to understand the lesson. This lays the foundation for video-editing that lesson to produce a version in which the board is visible, and in which it is not. To avoid confusion, the classroom in which the teacher collaborator actually teachers the lesson will be called the stimulus-creation classroom, because it is only used for creating a base video-lesson.

The teacher conducts the designed lesson in her stimulus-creation classroom while the researchers videotape it, capturing the complexity of a real classroom including spontaneous student-teacher interactions. This video-lesson maximizes the representation of a natural teaching and learning environment, and maximizes external validity of the stimuli. The recording itself is done in a manner that allows for later video-edits to create stimuli that will become the basis of experimental manipulations.

Next, the recorded video-lesson is systematically video-edited to create two or more *identical* instructional-videos with a single systematic difference – the manipulation of interest. The result is two or more versions of the same instructional-video that approximate a real classroom experience (e.g. Version I - solutions are visible on the board and Version II - solutions are not visible on the board).

Interactive details may then be added to the videotaped lesson. Typical school periods generally last between 40 – 60 minutes, which has been shown to provide difficulties for students to sustain attention whether in real classroom lectures or in flipped video-lectures in which students watch a videotaped lecture outside of class and discuss homework during a shared class period (Bunce, Flens, & Neiles, 2010; Risko, Buchanan, Medimorec, & Kingstone, 2013; Wilson & Korn, 2007). The lessons may be edited to include interactive prompts which require short answers throughout the lesson. One option is to allow students watching the video to respond to questions asked by the teacher in the video. Another option is to pause the video and embed question prompts that probe students to analyze, interpret, practice, or make inferences based on aspects of the lesson. While this is not an essential component of the methodology, it helps overcome the challenge of maintaining student attention (e.g. Kauffman, Ge, Xie, & Chen, 2008). This phase completes the stimulus creation of the video-lesson (technical video-editing details are provided in the following section).

In a different school, new teachers are recruited and the new teachers' students are administered a pretest. These are intervention-classrooms. To minimize testing effects, we recommend the pretest to take place 1-week prior to the intervention and posttest. Using intervention-schools' computer labs, or small netbooks, each student or students with partners, can be randomized within intervention-classrooms to either watch video-lesson Version I (e.g. solutions visible) or Version II (e.g. solutions not visible), while remaining within the natural classroom environment. This is followed by an immediate posttest then a 1-week delayed posttest of students' learning outcomes. Finally, the outcomes of each condition are analyzed to answer the hypothesis of interest.

The video-methodology requires that students are comfortable sitting in front of a computer screen and following prompts through a video-lesson. Thus, the methodology may be most appropriate for grades 3 and up, and especially from middle grades through university levels, where it is typical in many countries for a teacher to stand in front of students sitting at their desks. However, researchers should use their best judgment when deciding the most appropriate setting and participant age in relation to their research questions and the generalizability of their findings when utilizing this methodology.

The following section describes a case study utilizing the video methodology within a mathematics classroom with middle school students in a U.S. public school. In this case study, we provide more detail about the specific technical steps involved in video-taping and video-editing for these classroom lesson stimuli.

Video-Editing for Creating Experimental Conditions that Test Efficient Classroom Techniques: Case Study

A recent study by Begolli and Richland (2016) illustrates the potential of video editing for examining questions that lead to causal inferences about specific teaching practices. They examined the benefit of making student responses visible during a math lesson. This section will illustrate how to create different experimental conditions from a single lesson through systematic video-editing techniques, in Final Cut Pro (FCP), to vary only one teaching practice. Technological advances will make these techniques easier, so we urge the readers that are using alternative video-editing software to gloss over the technical detail while attending to the process.

Begolli and Richland (2016) built on the reasoning and classroom mathematics literatures to test whether problem solutions on the board enhanced students' ability to draw connections when compared with hearing them verbally. In particular, the study aimed to learn whether making multiple solution strategies visible on the board enhanced students' ability to draw connections when compared to only having the most recent solution visible, as evidenced by better procedural and conceptual transfer following the lesson. As an applied goal, these questions targeted a specific teaching practice, gaining insight on whether teachers should make student solutions visible or simply discuss them verbally in a classroom mathematics discussion.

The study aimed to compare the efficacy of three pedagogical strategies for leading classroom mathematics discussions – to write three solutions to the same problem on the board for discussions (All Visible condition), write one solution on the board at the time (Sequentially Visible), or to have only a verbal discussion of three different ways that students solved the problem (Not Visible). We next illustrate how the filming and editing was conducted to create these three matched conditions, such that all of the teacher discourse, meandering student explanations, and unscripted comments/ questions, expressions, and such were the same.

INSERT FIGURE 2 HERE

The use of zooming, cropping, and different camera angles. The most important aspect of creating a video-lesson stimulus is planning ahead. The lesson script should be well thought out to embedded the hypothesis of interest without sacrificing natural delivery of the lesson, which is why it is critical to cooperate with the stimulus-creation teacher. The teacher and researcher should coordinate practice trials before recording the lesson to anticipate teaching or equipment errors, because there will only be a single take that will capture students' spontaneous responses. Certain aspects of the manipulation may be challenging to ensure in a single take, thus, researchers may need to schedule with the teacher to reshoot those parts without the students. This footage can be later incorporated using video-editing to simulate a true classroom experience. Also, careful consideration needs to be given to acquire recording equipment and to place the equipment appropriately in the classroom. The hypothesis informs lesson design as well as camera and microphone placement such that the recorded footage can be utilized to create the manipulations of interest during postproduction.

To examine the hypothesis in Begolli and Richland (2016) and create three versions of the same lesson, the planned camera placement was: 1) to place one camera that would capture the whole board (board-cam), 2) a camera that would capture the teacher and only the most recent solution visible (teacher-cam), and 3) a camera that captured the teacher and students but not the board (student-cam). Omnidirectional microphones were planted around the room and the teacher to adequately capture all student speech and ambient sounds (including distractions experienced by students such as a pen dropping on a table) and one lapel microphone on the teacher (Figure 2). Each microphone was connected via a splitter jack to a camera audio input or audio recorder to bypass connecting through an audio mixer. Once all of the recordings are captured, each is imported into FCP (see Pistone, 2011 at for tutorial) and the multiple video and audio streams are synced (see Randle, 2011 for tutorial).

The researcher's objective while editing the video footage was to maintain all information contained in the lesson constant across conditions, while manipulating whether the writing on the board was visible or not. This raised challenges for the video-editor seeking to create the manipulations while maintaining a coherent video-lesson with comparably matched content. For instance, a simple, but inadequate way of creating a Not Visible condition (in the NV condition students do not see anything written on the board) from the same footage could be to simply show a camera always focused on the students sitting at their desks. But this would create a confound, because it would also test whether viewing the teacher influenced student learning. Thus, other cameras that captured the teacher were used and the writing on the board was obstructed by zooming in on footage from the teacher-cam. This was accomplished by dragging the corner of the "canvas" in FCP (see Jackson, 2010, for tutorial), thus, cropping out the writing on the board from the visual canvas. Another method for obscuring the writing on the board was to select a perspective from the student-cam focused on the profile of the teacher, which due to its angle did not capture the board (Figure 2). Segments of a video source can be selected by "slicing" them and these segments can be superimposed/connected onto the main video stream, using the blade tool of FCP (see SohoEditorsUK, 2011, for tutorial). The outcome makes for a video footage with seamless transitions between video camera angles and achieves the goal of occluding the writing on the board.

INSERT FIGURE 3 HERE

The video-editing resulted in the creation of three versions of the same lesson with equal length (see Figure 3). In order to make the lesson interactive, each version was strategically divided into 9 independent clips ranging from 1-min to 5-min using the blade tool. The endpoints of each clip were chosen based on when the teacher asked questions to the class. The clips were then exported into the Compressor software so that they can be used as regular video files. Ideal video settings are contingent on the resolution of monitors that will be used to display the video. This process takes a significant amount of time and it is recommended to test and try multiple versions to achieve a resolution that matches the student-participants' monitors. The finished clips were embedded in a computer program that, at the end of each clip, prompted students with questions that were asked by the teacher in the videotaped classroom. Students either wrote their answers on paper, or selected multiple choice questions that the computer program collected as assessment data. To avoid writing lengthy computer programs, questions may also be inserted into the video stream itself using the text tool in FCP (see Burns 2010, for tutorial) and data may be collected on paper & pencil format. A great resource for troubleshooting, is the peer-to-peer support community for media production professionals called creative communities of the world (Creative COW).

Video editing techniques enabled the creation of controlled stimuli that approximate a *live* lesson and learning in a dynamic, realistic context, which can be randomized within multiple student populations.

Example-study design. Begolli & Richland (2016) recruited 88 sixth grade students from a public school. Students in their classrooms completed a pretest, then 1-week later were administered the intervention in their computer laboratories followed by an immediate posttest, and a delayed posttest 1-week after the intervention. The pretest was administered 1-week prior to the intervention to minimize any testing effects (Roediger & Karpicke, 2006).

During the intervention, students were randomly assigned to one of the video-edited conditions: All solutions visible, no solutions visible, or visible one at a time. Each student worked independently with head phones at a computer, so they could be randomly assigned to condition on a within-classroom basis. These lessons were made interactive, so that all students wrote solutions and interpretations about the other students' discussed solutions, with no more than five minutes between each interactive opportunity.

Begolli and Richland (2016) found that students who watched a video where all solutions were visible on the board outperformed both groups in conceptual knowledge at posttest when controlling for pretest. Surprisingly students who saw solutions presented one-at-a-time performed worse in procedural knowledge than students, who could not see the board. The authors suggest those students may have retained the first solution presented, which was a misconception, and not being able to see the solution throughout the discussion meant that they couldn't go back and revisit their initial retention of that solution strategy. This may be related to working memory (WM) resources necessary to learn new concepts through instructional comparisons (Begolli, Richland, & Jaeggi, 2015).

These results open up new and broader questions regarding the role of individual differences in individual students, such as WM capacity, which could be easily explored through interactions with these videotaped stimuli as well. As such, Begolli & Richland (2016) provide one insight into ways that video-lessons can be used to answer questions with implications for advancing both theory and practice.

Video editing also can simulate interaction necessary in contexts such as medical education, where ethical restraints restrict medical students from interacting with patients for purely pedagogical goals, but teaching medical science is much facilitated through observation of, and engagement with, real patient scenarios that may be more messy than textbook examples (e.g., Balslev, De Grave, Muijtjens, & Scherpbier, 2005). The use of video may provide a controllable, and less ethically challenging, avenue for allowing students to engage with patients. Balsley and colleagues compared video to written text examples of patients experiencing periodic or episodic symptoms (e.g. caused by epilepsy or disorders mimicking epilepsy) and reported that medical students showed improvements in theory building, theory exploration, and diagnosis accuracy after watching a video. Thus, prerecorded video is useful in cases when others could be harmed. Despite video not including the full complexity of interpersonal interaction, it may provide greater educational benefits than more common textbook scenarios by providing a closer approximation of real world expressions of symptomology and the interpersonal dynamics of doctor-patient visits.

Limitations of Video-Stimuli. While the video stimuli methodology provides many potent opportunities for overcoming the shortcomings of decontextualized stimuli while maintaining experimental control, there remain limitations, and this methodology cannot be the final step in assessing the educational relevance of instructional principles. Video is an *approximation* of a classroom experience, not a true classroom experience, so an effective practice identified through a videotaped lesson experiment must be tested in classroom practice as well to ensure its efficacy. At the same time, as noted above, it offers a more realistic medium than text-based or computerized materials. Video can convey emotion, body language and other non-verbal cues, which can be filtered by the student based on their own individual differences.

Another consideration is that there are limitations in the range of questions that can be answered by systematically video-editing a single lesson or set of lessons. For instance, it may be difficult to edit the same lesson to introduce new or different information into each version. In a possible extension to Begolli and Richland (2016), it would be challenging to create from a single lesson, one version of a lesson in which a teacher compares solutions to a single problem, and a second version in which a teacher compares solutions to multiple problems. Yet, this could be achieved by creating two lessons.

A related point is that while one is comparing two versions of the same lesson to test some particular element that is being manipulated, there may be crucial problems in the lesson itself that limit student learning at all. For example, impromptu actions or instructional errors that arise during key instructional moments could obscure any variability in learning that would have been caused by the manipulated instructional practice. Reshooting the lesson with the same teacher and students would be unnatural and unlikely to be seamlessly incorporated into the main lesson video, and practically, would disrupt the teacher's curriculum. One technique for overcoming this is by doing short reshoots of the problematic parts of the lesson with the same teacher in the same classroom, outside of classroom time, and inserting them in the video-lesson during postproduction. Despite these limitations, there are vast opportunities for answering questions within these methodological constraints. Many of these questions could stem from surveying current experimental work in classroom settings.

Lastly, we note that there are ethical challenges in the use of video for experimentation. One is that videotaping teachers and students must be conducted responsibly and with the approval of not only the researcher's Institutional Review Board (IRB), but also with the approval of the institution in which the videotaping will be conducted (e.g. school and/or district) as well as the participants, such as a teacher and classroom students. Advances in editing can enable students to be obscured within the video record, and levels of requirements for parental consent will vary based on the level of risk inherent in the work. But, this must be a subject of clear consideration. In schools, while all studies will involve different considerations, one recommendation from our own work is that the classroom that was videotaped for a study should not be within the same school as the classrooms where the experiment will take place. Thus the researchers should minimize the likelihood that experimental participants can recognize teachers or students in the videotaped lesson. Finally, as in any other experiment with student populations, we caution researchers to ensure that experimental manipulations are not harmful or deleterious to students' long-term knowledge.

Similar ethical considerations are present in a medical education context. It is essential that any patient who is videotaped be informed and allowed to give informed consent to allow the videotaped content to be used in future experimentation, and the type of consent procedure should mirror the level of risk inherent in the videotaped information for violating the patient's confidentiality or other concerns.

Practically, therefore, some of the major challenges in this process relate to gaining appropriate permissions from parents and the Institutional Review Board (IRB) to record teachers and students, finding a stimulus-creation teacher with whom to collaborate, and then finding different intervention-classrooms in schools to administer the study. This can take considerable amounts of time, depending on University and school district policies; thus, we recommend that researchers begin the process early.

Current Trends of Controlled Experiments in Classroom Settings and their Limitations for Informing Teaching Practices: A Comparison to the Video Methodology

Finally, we discuss the relationships between this video methodology and other techniques that have been used to bridge classroom and laboratory-based psychological experiments. Three common methods have been used most frequently in these studies. First, *instructional materials* might be presented as text on paper or as computerized graphics requiring little or no intervention on the part of the classroom teacher. Second, an experimenter or teacher is trained to conduct a lesson by enacting a specific *teaching practice*. In both cases, the learning principle is embedded either within the *instructional materials* or *teaching practice,* and varied between two or more groups. Third, interactive technologies lead to tutoring or scaffolding interactions that allow for testing principles of best practices for supporting student learning in specific content areas. We briefly discuss an example of these common strategies in turn, and compare them to the use of a video-based lesson.

Worksheet-based Experimentation. Experimental designs in which learning principles are tested by experimenters crafting, or working with educators to design, stable instructional materials such as worksheets can be carefully controlled to maximize internal validity and make important practical contributions for complementing teaching practice. In one such set of studies, Booth and colleagues (e.g. Booth, Cooper, et al., 2015; Booth, Oyer, et al., 2015) examined the effects of embedding worked-out examples with self-explanation prompts in worksheets on students' learning of algebra concepts. In their studies, teachers within classrooms randomly assigned half of their students in their classrooms to either complete a worksheet which included worked-examples and self-explanation prompts or a worksheet with the same problems without worked-examples or self-explanation prompts. Students completed the worksheets individually as an in-class activity. Otherwise, all aspects of the lesson remained identical, since the students in the treatment and control condition were in the same classroom. A similar study was conducted by randomly assigning classrooms to either be in the treatment or control conditions at the classroom level while ensuring that teachers taught the same topics in both control and treatment classrooms.

A tremendous advantage of this methodology is that it does not require teachers to make changes to their curriculum or teaching style, and teachers report enriched discussions with treatment students or in treatment classrooms as a result of utilizing the worksheets. Teachers are involved in the design of the worksheets, which gives them agency over the materials and improves adoption rates, especially for collaborating teachers. Further, worksheets can be implemented without extensive professional development or changes to curriculum. In terms of teaching enactments, however, broad principles derived from these materials may be challenging to translate into *teaching practice.* This limits their generalizability, and while paper materials and learning software may prompt teachers to have richer discussions, they do not directly inform the nature of teacher actions that are effective when orchestrating classroom discussions.

Teacher- or Confederate-led Experiments. Teacher or confederate lessons are much more like what real-world teachers will do, and thus insights from such practices can more directly inform everyday teachers' practice. At the same time, experimental control is much more difficult. In fact, it is likely impossible with a live teacher or experimenter to repeat the same lesson introducing external or confounding variables (e.g. changes in teacher's affect, gestures, discourse timing, pauses, etc.). One may consider this in a recent study testing the efficacy of new curricular materials on epigenetics through a randomized controlled trial design with high school biology students (Drits-Esser, Bass, & Stark, 2014). Students were randomly assigned to be taught epigenetics using interactive, online materials - the treatment condition - or through videos and a brainstorming activity – the control condition. Both interventions were delivered by an educational specialist from the university. The education specialist taught first half of the class as the treatment condition, and then the second half of the class as the control condition. Students were randomly selected to spend half of the class time with their teacher in their classrooms and the other half of the time with the education specialist in the computer laboratory. The teacher taught all students using a worksheet that reviewed genetics concepts that were unrelated to the experimental materials.

Drits-Esser et al.'s (2014) measures of fidelity suggest lessons between classrooms contained the same activities without major adaptions and were comparable in affect. However, there is no manner to guarantee that details of the interactions, including student-educator discourse, clarity of explanations, and enthusiasm were identical between conditions and classrooms. A video-lesson would ensure that all aspects of the lesson are identical between conditions and classrooms. Further, the critical parts tested were *instructional materials*, as is common with educational interventions, which do not address how a teacher should enact a lesson. Lastly, the treatment curricula were always used first, which introduces teacher practice effects, and time of day differences that could lead to differential learning. As any educator who teaches multiple sections of the same lesson materials knows, even exactly the same materials can "go well" in one classroom period and less well another. These seemingly intangible differences can be of crucial importance in a lesson-based experiment comparing different versions of the materials.

In another, similar design, Schwartz, Chase, Oppezzo, and Chin (2011) tested the role of inventing solutions to physics problems before giving students the formula (inventing condition) versus “telling” the formulas to students before prompting them to solve problems (tell & practice condition). They placed students in each condition through stratified random assignment based on students’ cumulative science scores and similarly to Drits-Esser et al. (2014), they used separate classrooms taught by the same instructor when administering the manipulations to minimize teacher effects. Further, students in the inventing condition received a lecture after the treatment and students in the tell & practice condition received a lecture before the treatment.

While as carefully controlled as possible, one could imagine that this design could have been productively implemented using a video-based design, in which students were prompted through interactive prompts to either invent or learn the procedure first, with the same lecture simply varied in when it was edited to be sequenced.

Computer-led Tutoring or Other Interactive Environments. A third model for designing controlled studies in which lesson materials were delivered through an interactive system is the use of interactive technologies. These could include touch-screen or standard computer technologies, and many systems have been developed for one-on-one standardized delivery of content. These include cognitive tutors or interactive avatars (see Aleven & Koedinger, 2002; Sottilare et al., 2016), all of which can be highly scripted or made to interact in systematic ways. These allow for careful testing of design principles, but their application to real-world teachers are not at all clear. For example, in an experiment which attempted to understand the effects of self-explaining, Aleven and Koedinger (2002) randomly assigned students to learn from either a computerized tutor (Cognitive Tutor) which either prompted them to explain or not explain their steps during problem-solving. In contrast to utilizing live teachers, the computerized environment maintains the learning environment identical within and between conditions (apart from the manipulation), which maximized the study's internal validity and provide important insights for the improvement of intelligent tutoring systems, but they are limited in generalizability for informing teacher enactments.

Discussion

Decades of research in cognitive science on learning behavior have not made serious inroads into our educational discussions, and less so into teaching practices (Dunlovsky & Rawson, 2012). This may be in part explained by the fact that much of this research has been conducted in laboratory settings using stimuli that are quite different from those of everyday classroom instruction (Richland, Linn, & Bjork, 2009). Recently, there has been a rise in the number of classroom experiments conducted to examine the generalizability of learning processes derived from laboratory work by cognitive and educational psychologists to classroom contexts (e.g., see Mayer, 2012; Roediger & Pyc, 2012). This body of work has been able to make precise, testable claims about cognitive learning mechanisms, with many lines of work summarized in Table 2 by Koedinger, Booth, & Klahr (2013). The overarching commitment to maintain experimental control, however, has driven many of the studies bridging laboratory and classroom settings to use stimuli or methodologies that are more careful, constrained, and shorter term learning interventions than even a single everyday lesson. This may contribute to a lack of specificity when making recommendations about teaching practices (e.g., see Daniel, 2012; Kornell, Rabelo, & Klein, 2012; Mayer, 2012).

INSERT TABLE 2 HERE

Designing manipulations that maintain experimental control but that involve teacher-led instruction are often difficult to achieve (Brown, 1992). Recommendations stemming from laboratory work in cognitive and educational psychology (e.g. immediate versus delayed feedback, use of concrete versus abstract materials), often are in contradiction, and their effectiveness varies across contents and contexts (Koedinger, Booth, & Klahr, 2013). This issue is further made challenging for teachers looking to adopt these recommendations, because many experiments derive abstractions through manipulations and/or stimuli that do not aim to mimic an actual teacher's behavior (for exceptions see Begolli & Richland, 2016; Cook et al., 2013; Endres et al., 2017; Glogger-Frey et al., 2015; Richland & McDonough, 2010). These factors may contribute to the challenges of translating research into practice (e.g., see Daniel, 2012; Kornell, Rabelo, & Klein, 2012; Mayer, 2012). While promising work in emerging fields is attempting to reduce the complexity of instructional recommendations (Society for Research on Educational Effectiveness; International Education Data Mining Society), the proposed methodology focuses on the latter issue, that of discovering findings that better translate to *teaching practices,* by using video-based stimuli.

Thus, video methodology and video editing techniques presented in this paper provide a model that can be adopted by many researchers interested in making causal claims regarding teaching practices. This model also provides opportunities for creating a bi-directional exchange between the classroom and the laboratory, as outlined in Figure 1.

Testing Teaching Practices based on Cognitive and Education Theories

Finally, we suggest some directions in which this methodology could be productive to explore current research subjects in new ways. To illustrate the breadth of the potential within this methodology, one might consider a theory-based question such as: how do invention activities (Schwartz et al., 2011) guided by a teacher affect student outcomes when presented before versus after direct instruction? This question could be answered by recording a lesson that begins with (a) invention activities that are followed by (b) direct instruction and practice. In video post-production, researchers could create two versions of the lesson in which the order of these two activities (a) and (b) is counterbalanced. The video-lesson could be split into two (or more) clips, one clip contains (a) the invention activity and the other clip (b) direct instruction with practice. This would enable the researcher to create two versions of the lesson, which maintains constancy, apart from manipulating the sequence of instructional activities (invention vs. direct instruction). To make the link seamless, we recommend including a question and answer period in between the two clips. This would provide a more rigorous test of the order effects than having teachers teach different students in the two counterbalanced versions, or different teachers teach the two versions. Each of those has the potential for confounding the order effects with contextual features of the teacher-student interactions.

Another example derives from Stein et al.'s (2008) recommendation for how to sequence multiple solutions to a single problem. While Stein et al. (2008) propose that teachers should sequence solutions from simpler to more complex, they admit that more work needs to be done to compare the effects of different sequencing methods on student learning. For example, prior work stemming from observations of Japanese lessons from the TIMSS 1999 study (Stigler & Hiebert, 1999) show that teachers in higher achieving countries present students a solution that represents a common misconception before moving on to the correct solution (Shimizu, 2003). But whether common misconceptions should be presented first or last, remains to be tested empirically. A video-based lesson could be split into independent clips (similarly to the example above) and the order of problem solutions could be manipulated through video-editing. One version orders the video clips so the common misconception is presented first in the sequence of multiple problem solutions, whereas in another version the common misconception is presented last.

Video-lessons could also include minor variations of key moments within a lesson, but maintain the rest of the lesson unchanged. Many educational recommendations are difficult to test through whole lesson comparisons that include many differences beyond the specific moments of interest, so researchers tend to either test small learning opportunities in controlled ways or more variable whole lesson comparisons with less control (for review see Mayer, 2008, 2012; Pashler et al., 2007; Richland et al., 2007; Roediger & Pyc, 2012).

A prime educational platform for examining theory-based questions in teaching and learning with video-editing are massive open online courses (MOOCs). The MOOC platform has spawned multiple new lines of research, however, these efforts have delivered few insights for teaching and learning (Reich, 2015). The video methodology within MOOC video-lessons may have a deep impact where video-editing can be used to highlight theory-driven instructional practices. Further, experimental stimuli could fully emulate the MOOC instructional environment affording discoveries in basic research and direct application to MOOC environments.

Summary

Altogether, video-editing provides a rich opportunity for bridging cognitive and educational research through experimental designs. Unlike other methodologies that test learning principles but may leave teachers with the challenge of knowing how or which recommendations to incorporate into a real lesson, video can be used as a tool to give specificity to teaching strategies. We described ways that videotaped-lessons can be scripted to incorporate learning principles, modified to create manipulations that test hypothesis in learning behavior, and used as stimulus for a randomized experiment. In particular, we described a demonstration study that tested whether making problem solutions visible in a videotaped-teacher guided lesson affected student learning. Begolli & Richland (2016) demonstrated one way that zooming and different camera angles could be used to create video-stimuli. These stimuli that approximate a real lesson maximize external validity while also maintaining consistency across conditions, enabling researchers to infer causality to manipulated instructional practices. Many teachers use effective instructional techniques embedded in their lessons and this tool provides us with ways that can test which of these techniques are most efficient when the rest of the lesson is held constant. The examples provided in the discussion of the manuscript were aimed to illustrate the opportunities for transforming and developing research in the context of testing educational strategies that have particular relevance for teachers looking to support student learning.

There is a growing necessity for controlled experiments to better incorporate contextual factors inherent in a classroom environment in their stimuli (e.g., Daniel, 2012; Kornell, Rabelo, & Klein, 2012; Mayer, 2012). We hope that future researchers will begin considering how to embed learning principles in real lessons, and in particular, will adopt video and video editing tools to create manipulations that test the efficiency of learning principles in a contextualized way that is close to everyday classroom practice. In this way, ideally they will make gains in developing a broader research base on cognition in everyday settings, as well as providing teachers with successful, specific, and relevant instructional techniques.

References

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26*(2), 147–179.

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology*, *89*(1), 183–193. http://doi.org/10.1037/0022-0663.89.1.183

Atchley, P., & Kramer, A. F. (2001). Object and space-based attentional selection in three-dimensional space. *Visual Cognition*. http://doi.org/10.1080/13506280042000009

Balslev, T., De Grave, W. S., Muijtjens, A. M. M., & Scherpbier, A. J. J. A. (2005). Comparison of text and video cases in a postgraduate problem-based learning format. *Medical Education*, *39*(11), 1086–1092. http://doi.org/10.1111/j.1365-2929.2005.02314.x

Barnier, a. J. (2012). Memory, ecological validity and a barking dog. *Memory Studies*, *5*(4), 351–359. http://doi.org/10.1177/1750698012461243

Barsalou, L. W. (1999). Perceptual symbol systems. *The Behavioral and Brain Sciences*, *22*(4), 577-609-660. http://doi.org/10.1017/S0140525X99252144

Begolli, K. N., & Richland, L. E. (2016). Teaching Mathematics by Comparison: Analog Visibility as a Double-Edged Sword. *Journal of Educational Psychology*, *108*(2), 194–213. http://doi.org/http://dx.doi.org/10.1037/edu0000056

Begolli, K. N., Richland, L. E., & Jaeggi, S. M. (2015). The role of executive functions for structure-mapping in mathematics. In R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, A. Warfaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (Vol. 1, pp. 1–6). Austin, TX: Cognitive Science Society.

Berry, G., & Klein, R. (1993). Does motion-induced grouping modulate the flanker compatibility effect?: A failure to replicate Driver & Baylis. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *47*(4), 714–729. http://doi.org/10.1037/h0078867

Bindemann, M., Mike Burton, A., & Langton, S. R. H. (2008). How do eye gaze and facial expression interact? *Visual Cognition*, *16*(6), 708–733.

Booth, J. L., Cooper, L. A., Donovan, M. S., Huyghe, A., Koedinger, K. R., & Paré-Blagoev, E. J. (2015). Design-based research within the constraints of practice: AlgebraByExample. *Journal of Education for Students Placed at Risk*, *20*(1–2), 79–100. http://doi.org/10.1080/10824669.2014.986674

Booth, J. L., Oyer, M. H., Paré-Blagoev, E. J., Elliot, A. J., Barbieri, C., Augustine, A., & Koedinger, K. R. (2015). Learning algebra by example in real-world classrooms. *Journal of Research on Educational Effectiveness*, *8*(4), 530–551. http://doi.org/10.1080/19345747.2015.1055636

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *Journal of the Learning Sciences*, *2*(2), 141–178. http://doi.org/10.1207/s15327809jls0202_2

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education*, *87*(12), 1438–1443. http://doi.org/10.1021/ed100409p

Burns, J. (2010). *Adding text in Final Cut Pro by John Burns.* Retrieved from https://www.youtube.com/watch?v=8DaqLlvzN70.

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students’ retention of US history facts. *Applied Cognitive Psychology*, *23*(6), 760–771.

Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1988). Mathematical concepts in everyday life. *New Directions for Child and Adolescent Development*, *1988*(41), 71–87.

Creative Communities of the World (n.d.). Retrieved from https://www.creativecow.net/

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 7–19.

Collins, A. (1999) The changing infrastructure of education research. In E. Lagemann & L. Shulman (Eds.), *Issues in education research* (pp. 289–298). San Francisco: Jossey-Bass.

Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84*(0), 1863–1871. http://doi.org/10.1111/cdev.12097

Correa-Chavez, M., & Rogoff, B. (2005). Cultural research has transformed our ideas of cognitive development.pdf. *International Journal of Behavioral Development*.

Daniel, D. B. (2012). Promising principles: Translating the science of learning to educational practice. *Journal of Applied Research in Memory and Cognition*, *1*(4), 251–253. http://doi.org/10.1016/j.jarmac.2012.10.004

Derry, S. J. (2007). *Guidelines for Video Research in Education: Recommendations from an Expert Panel*. *Report submitted to the National Science Foundation*. Retrieved from http://drdc.uchicago.edu/what/video-research-guidelines.pdf

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., … Sherin, B. L. (2010). Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics. *Journal of the Learning Sciences*. http://doi.org/10.1080/10508400903452884

Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for

Educational Inquiry. *Educational Researcher*, *32*(1), 5–8. http://doi.org/10.3102/0013189X032001005

Donovan, M. S. (2013). Generating improvement through research and development in education systems. *Science*, *340*(6130), 317–319.

Drits-Esser, D., Bass, K. M., & Stark, L. A. (2014). Using Small-Scale Randomized Controlled Trials to Evaluate the Efficacy of New Curricular Materials. *CBE Life Sciences Education*, *13*(4), 593–601. http://doi.org/10.1187/cbe.13-08-0164

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*. http://doi.org/10.1016/S0166-2236(00)01633-7

Dunlosky, J., & Rawson, K. A. (2012). Despite their promise, there's still a lot to learn about techniques that support durable learning. *Journal of Applied Research in Memory and Cognition*, *1*(4), 254–256. http://doi.org/10.1016/j.jarmac.2012.10.003

Endres, T., Carpenter, S., Martin, A., & Renkl, A. (2017). Enhancing learning by retrieval: Enriching free recall with elaborative prompting. *Learning and Instruction*, *49*, 13–20.

Gibson, J. J. (1947). *Motion picture testing and research*. *Washington, DC: US Government Printing Office*. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0651783%5Cnhttp://scholar.google.ch/scholar?q=motion+picture+testing&hl=en&btnG=Search&as_sdt=1,5&as_sdtp=on#4

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. *Book* (Vol. 40). http://doi.org/10.2307/989638

Glogger-Frey, I., Fleischer, C., Grüny, L., Kappich, J., & Renkl, A. (2015). Inventing a solution and studying a worked solution prepare differently for learning from direct instruction. *Learning and Instruction*, *39*, 72–87. http://doi.org/10.1016/j.learninstruc.2015.05.001

Goldman, D. R. (2007). A Framework for Video Annotation , Visualization , and Interaction. *Framework*.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*(1), 5–26. http://doi.org/10.1037/0003-066X.53.1.5

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61-83-135. http://doi.org/10.1017/S0140525X0999152X

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study. *Cognition and Instruction*, *26*(4), 430–511. http://doi.org/10.1080/07370000802177235

Hutchins, E. (2000). Distributed Cognition. *Representations*. http://doi.org/10.1007/s10111-004-0172-0

Kauffman, D. F., Ge, X., Xie, K., & Chen, C.-H. (2008). Prompting in web-based environments: Supporting self-monitoring and problem solving skills in college students. *Journal of Educational Computing Research*, *38*(2), 115–137.

Kersting, N. B. (2008). Using Video Clips of Mathematics Classroom Instruction as Item Prompts to Measure Teachers' Knowledge of Teaching Mathematics. *Educational and Psychological Measurement*, *68*(5), 845–861. http://doi.org/10.1177/0013164407313369

Kilduff, L. P., Hopp, R. N., Cook, C. J., Crewther, B. T., & Manning, J. T. (2013). Digit ratio (2D: 4D), aggression, and testosterone in men exposed to an aggressive video stimulus.

Evolutionary Psychology, *11*(5), 147470491301100500.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*(5), 176–180.

Klahr, D. (2013). What do we mean? On the importance of not abandoning scientific rigor when talking about science education. *Proceedings of the National Academy of Sciences of the United States of America*, *110 Suppl*(2003), 14075–80. http://doi.org/10.1073/pnas.1212738110

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*, *14*(2), 217–238. http://doi.org/10.1007/s10956-005-4423-5

Koedinger, K. R., Aleven, V., Roll, I., & Baker, R. (2009). In vivo experiments on whether supporting metacognition in intelligent tutoring systems yields robust learning. *Handbook of Metacognition in Education*, *42*(1983), 647–51. http://doi.org/10.1002/lsm.20954

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional Complexity and the Science to Constrain It. *Science*, *342*(22 november 2013), 935–937. http://doi.org/10.1126/science.1238056

Kornell, N., Rabelo, V. C., & Klein, P. J. (2012). Tests enhance learning-Compared to what? *Journal of Applied Research in Memory and Cognition*, *1*(4), 257–259. http://doi.org/10.1016/j.jarmac.2012.10.002

Kvavilashvili, L., & Ellis, J. a. (2004). Ecological validity and the real-life/laboratory controversy in memory research: a critical and historical review. Retrieved from http://centaur.reading.ac.uk/13968/

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.

Marcus, A., & Wilder, D. A. (2009). A comparison of peer video modeling and self video modeling to teach textual responses in children with autism. *Journal of Applied Behavior Analysis*, *42*(2), 335–341. http://doi.org/10.1901/jaba.2009.42-335

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *The American Psychologist*, *63*(November), 760–769. http://doi.org/10.1037/0003-066X.63.8.760

Mayer, R. E. (2012). Advances in applying the science of learning to education: An historical perspective. *Journal of Applied Research in Memory and Cognition*, *1*(4), 249–250. http://doi.org/10.1016/j.jarmac.2012.10.001

McCandliss, B. D., Kalchman, M., & Bryant, P. E. (2003). Design Experiments and Laboratory Approaches to Learning: Steps Toward Collaborative Exchange. *Educational Researcher*, *32*(1), 14–16. http://doi.org/10.3102/0013189X032001014

Monin, B., Oppenheimer, D. M., Ferguson, M. J., Carter, T. J., Hassin, R. R., Crisp, R. J., … Kahneman, D. (2014). Commentaries and Rejoinder on Klein et al. (2014). *Social Psychology*, *1*(1), 1–13. http://doi.org/10.1027/1864-9335/a000202

Oliveira, R. F., Rosenthal, G. G., Schlupp, I., McGregor, P. K., Cuthill, I. C., Endler, J. A., … Waas, J. R. (2000). Considerations on the use of video playbacks as visual stimuli: the Lisbon workshop consensus. *Acta Ethologica*. http://doi.org/10.1007/s102110000019

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, *7*(6), 657–660. http://doi.org/10.1177/1745691612462588

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Pardos, Z. a., Baker, R., San Pedro, M. O. C. Z., Gowda, S. M., & Gowda, S. M. (2013). Affective states and state tests. *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*, 117. http://doi.org/10.1145/2460296.2460320

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A. C., Koedinger, K. R., McDaniel, M. A., & Metcalfe, J. (2007). *Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004. National Center for Education Research*.

Pistone, B. (2011). *Log and Capture in Final Cut Pro.* Retrieved from https://www.youtube.com/watch?v=YXNA6eUUb_M&feature=youtu.be

Randle, B. Q. (2011). *Syncing Audio and Video Using the Timeline in Final Cut Pro.* Retrieved from https://www.youtube.com/watch?v=sHyCWDTYemg&feature=youtu.be

Reich, J. (2015). Rebooting MOOC research. *Science*, *347*(6217), 34–35.

Rich, P. J., & Hannafin, M. (2008). Video Annotation Tools: Technologies to Scaffold, Structure, and Transform Teacher Reflection. *Journal of Teacher Education*, *60*(1), 52–67. http://doi.org/10.1177/0022487108328486

Richland, L. E., & Hansen, J. (2013). Reducing Cognitive Load in Learning by Analogy. *International Journal of Psychological Studies*, *5*(4), 69–80. http://doi.org/10.5539/ijps.v5n4p69

Richland, L. E., Linn, M. C., & Bjork, R. A. (2007). Instruction. *Handbook of Applied Cognition*, 553–583.

Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology*, *35*(1), 28–43.

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers and Education*, *68*, 275–283. http://doi.org/10.1016/j.compedu.2013.05.001

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*(3), 529–544. http://doi.org/10.1037/a0014224

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. *Psychological Science*, *17*(3), 249–255. http://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, *1*(4), 242–248. http://doi.org/10.1016/j.jarmac.2012.09.002

Rossi, P. G., & Fedeli, L. (2016). The Use of Videos in the Triangulation Process among Professors, School Teachers, and Students: Promoting Permeability between Pre. *Integrating Video into Pre-Service and In-Service Teacher Training*, 24.

Santagata, R. (2005). Practices and beliefs in mistake-handling activities: A video study of Italian and US mathematics lessons. *Teaching and Teacher Education*, *21*(5), 491–508.

Santagata, R., Gallimore, R., & Stigler, J. W. (2005). The use of video for teacher education and professional development. Past experiences and future directions. *Current Perspectives on Applied Information Technologies (Volume 2): Preparing Teachers to Teach with Technology.*, 151–167.

Sawyer, R. K. (2014). *The Cambridge Handbook of the Learning Sciences* (Second). Cambridge University Press.

Saxe, G. B. (1991). Culture and cognitive development: Studies in mathematical development. Hillsdale NJ: Lawrence Erlbaum Associates.

Schwartz, D. L., & Bransford, J. D. (2005). Efficiency and Innovation in Transfer. *Transfer of Learning from a Modern Multidisciplinary Perspective*, 1–51. http://doi.org/10.1111/j.1365-2133.2005.06492.x

Schwartz, D. L., Chase, C. C., Oppezzo, M. a., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759–775. http://doi.org/10.1037/a0025140

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. http://doi.org/10.1068/p2952

Soble, J. R., Spanierman, L. B., & Liao, H.-Y. (2011). Effects of a brief video intervention on White university students' racial attitudes. *Journal of Counseling Psychology*, *58*(1), 151–157. http://doi.org/10.1037/a0021158

SohoEditorsUK (2011). *Final Cut Pro 7 Blade Tool.* Retrieved from https://youtu.be/JJFQZcM8o_M.

Sottilare, R. A., Graesser, A. C., Hu, X., Olney, A., Nye, B., & Sinatra, A. M. (2016). *Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling* (Vol. 4). US Army Research Laboratory.

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, *11*(2), 107–125.

Suchman, L. A. (1987). *Plans and Situated Actions*. *Cambridge University Press*. Retrieved from http://books.google.com/books?id=AJ_eBJtHxmsC

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(December), 661–699. http://doi.org/10.1111/j.0023-8333.2005.00320.x

Valenzeno, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187–204. http://doi.org/10.1016/S0361-476X(02)00007-3

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244–276. http://doi.org/10.1016/j.tate.2006.11.005

Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychologica*, *133*(1), 51–56. http://doi.org/10.1016/j.actpsy.2009.09.001

Wilson, K., & Korn, J. H. (2007). Attention During Lectures: Beyond Ten Minutes. *Teaching of Psychology*, *34*(2), 85–89. http://doi.org/10.1080/00986280701291291