Journal of Science Education and Technology (2018) 27:369–384
https://doi.org/10.1007/s10956-018-9730-8

Tangible User Interfaces and Contrasting Cases as a Preparation for Future Learning

Bertrand Schneider[1] · Paulo Blikstein[2]

Published online: 20 April 2018

Abstract
In this paper, we describe an experiment that compared the use of a Tangible User Interface (physical objects augmented with digital information) and a set of Contrasting Cases as a preparation for future learning. We carried out an experiment ($N = 40$) with a 2×2 design: the first factor compared traditional instruction ("Tell & Practice") with a constructivist activity designed using the Preparation for Future Learning framework (PFL). The second factor contrasted state-of-the-art PFL learning activity (i.e., students studying Contrasting Cases) with an interactive tabletop featuring digitally enhanced manipulatives. In agreement with prior work, we found that dyads of students who followed the PFL activity achieved significantly higher learning gains compared to their peers who followed a traditional "Tell & Practice" instruction (large effect size). A similar effect was found in favor of the interactive tabletop compared to the Contrasting Cases (small-to-moderate effect size). We discuss implications for designing socio-constructivist activities using new computer interfaces.

Keywords Learning · Collaboration · Tangible User Interfaces · Contrasting Cases · Preparing for Future Learning

Introduction

Over the past decades, mainstream educational research has cyclically swung between constructivist and behaviorist approaches. One of the main dimensions along which Behaviorist approaches ("programmed instruction," Skinner 1986) differ from Constructivist approaches (Piaget 1928) is the space students have for discovery of the target content. Despite the fact that discovery (in particular, before formal instruction) has been shown to be a powerful enhancer of learning (Schwartz and Bransford 1998), the controversy is very much alive (Kirschner et al. 2006), especially with the popularity of lecture-based innovations, such as "flipped classrooms" and MOOCs (Thille et al. 2015)[1], and empirical studies finding that there is no benefit in recall for having learned through discovery methods (Klahr and Nigam 2004).

But new technologies also brought new possibilities to Constructivist researchers and designers, who incorporated into their toolbox tangible user interfaces, mobile devices, and physical computing tools. These tools, while opening up greatly the design space, also pose the challenge of how to engineer good constructivist activities within a much more complex design space and how to collect data and measure students' progress in these new systems. The goal of this paper is to present such a case of a successful application of constructivist theories informing the design of discovery-based activities using new technologies, as well as the analysis of the mechanisms that led students to achieve higher learning gains compared to traditional instruction. In this paper, we refer to traditional instruction as a specific sequence ("Tell and Practice") where students are *first* told what they need to know and *then* asked to practice their understanding of

✉ Bertrand Schneider
bertrand_schneider@gse.harvard.edu

Paulo Blikstein
paulob@stanford.edu

1 Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA

2 Graduate School of Education, Stanford University, Stanford, CA, USA

[1] MOOCs and flipped classroom models can be used in a constructivist fashion, but in the vast majority of cases, they still follow behaviorist-inspired instructional approaches (i.e., "Tell & Practice," where students are first told what they need to know and then asked to practice their understanding of various concepts on a series of quizzes or exercises).

the concepts taught with a subsequent activity. More specifically, our goal is to investigate whether Tangible User Interfaces (TUIs; described subsequently) have the potential to efficiently support constructivist activities in collaborative settings and promote deep understanding of STEM concepts by following an alternative instructional sequence.

Theoretical Framework

The general theoretical framework of this paper is the idea that people learn best by using their prior knowledge to make sense of new knowledge (Piaget 1928). In particular, it has been shown that building digital or physical artifacts and going through a process of debugging mental models by externalizing them are especially powerful (Papert 1980). Even though this constructivist view of learning is one of the most commonly used frameworks in educational research, there are two main limitations when implementing this approach.

First, students need to have some pre-existing knowledge or experiences that they can use to make sense of new concepts; this approach falls short if students do not have any prior experiences in the domain taught or if they do not have the opportunity to build some foundations or intuitions prior to instruction. Motivated students in this situation would be required to take plenty of notes and memorize as much of the teacher's lesson as possible with the hope that they will understand the content later. This scenario could, in fact, favor rote memorization and hinder transfer (Bransford and Schwartz 1999). One approach that attempts to mitigate this problem is the Preparing for Future Learning framework (PFL; Schwartz and Bransford 1998). The idea is to provide students with an open-ended activity prior to the lecture to allow them to construct some intuition about the concepts taught. Schwartz and Bransford (1998) argue that Contrasting Cases (CCS) are ideal candidates for this task. Contrasting cases are instructional materials designed to help students notice information they might otherwise overlook, in the same way that "tasting wines side-by-side can improve discernment" (Bransford and Schwartz 1999). They allow students to separate surface and deep features of a concept: for example, Schwartz and Martin (2004) demonstrated how having students inventing a reliability index for baseball pitching machines promoted learning (and the transfer) of concepts in statistics such as standard deviation. The CCS included performance diagrams of each machine to draw attention to features like precision, variability, and sample size. This framework motivated the first comparison of our study between a PFL-style constructivist activity and a more standard T&P instruction.

The second limitation is that efficient discovery learning activities are notoriously difficult to design (e.g., De Jong and Van Joolingen 1998). It takes a considerable engineering and designing effort to create an activity that engages and motivates learners, targets specific learning goals, has affordances for conceptual reflection, works with both high-achieving students and less proficient ones, and allows for productive failure (Kapur 2008). CCs, for instance, are an especially difficult case: they require designers to carefully separate surface and deep features of a concept and highlight those differences in an engaging set of examples. Over the past years, many believed that computer simulations and programming environments would bring a solution to this problem, by providing engaging virtual environments where students could explore rich micro-worlds and experiment with scientific and mathematical phenomena in a hands-on fashion (Papert 1980; Edwards 1995; Kynigos 2007). De Jong and Van Joolingen (1998) conducted a review of the various empirical studies using computer simulations as discovery-learning tools and found some mixed (but mostly positive) effects on students' learning. To facilitate discovery learning, they suggest (based on prior work) to add ways to directly access domain knowledge, support hypothesis generation, the design of experiments, making predictions and regulating learning processes. More modern systems are incorporating these principles into software and curricula, showing positive results for a variety of age groups (e.g., Blikstein and Wilensky 2009; Wilkerson-Jerde et al. 2015; Olson and Horn 2011). We build on this prior work and extend this idea to new interfaces appropriate for collaborative hands-on learning.

Tangible User Interfaces as a Way to Support Collaborative Learning

TUIs are computer systems that replace the traditional mouse and keyboard with physical objects, detect their states (such as their location), and provide a feedback loop to replace the screen with an augmented reality system (for instance by using a projector and displaying additional information directly on the objects). TUIs have specific features that can support constructivist learning (e.g., by providing easy-to-use physical objects, which has been shown to support students' exploration of a complex domain; Schneider et al. 2011), students' enjoyment (Shaer et al. 2011), engagement (Schneider et al. 2012), and collaborative hands-on activities in co-located settings (Dillenbourg and Evans 2011). For instance, researchers (e.g., Falcão and Price 2011) have shown that tangible interfaces support balanced levels of participants among all group members due to their physicality and "present at hand" nature; in particular, findings suggest that interferences in those environments prompt verbal negotiations and synchronization of actions, which facilitates conflict resolution and more generally collaborative learning activities. Figure 1 presents some examples of educational interfaces built by (or in collaboration with) learning scientists over the past decade.

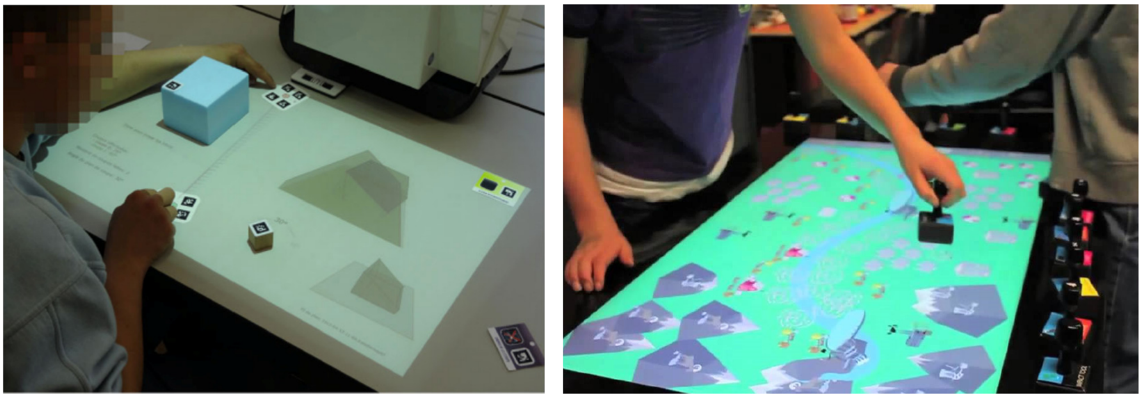


Fig. 1 On the left, the TapaCarp system, designed to help train carpenter apprentices; this system has been implemented in classrooms by Cuendet et al. (2015). On the right, the Youtopia system designed to help 5th graders to collaboratively learn about sustainability (Wise et al. 2015)

As mentioned by Dillenbourg and Evans (2011), interactive tabletops "convey a socio-constructivist flavor: they support small teams that solve problems by exploring multiple solutions". Several studies have reported higher levels of collaboration when small groups of students interacted with an interactive tabletop compared to a multi-touch table (e.g., Schneider et al. 2011), a pen and paper activity (e.g., Schneider et al. 2012), or paper handouts (Piper and Hollan 2009). The goal of our study is not to test whether tangibles have a positive impact on collaborative learning; rather, we leverage this positive effect of TUIs on collaboration by having students work in dyads (group of two students). By setting up this study as a collaborative task, we hope to facilitate knowledge building and potentially socio-cognitive conflicts during the learning activity.

Previous Work

This contribution is a direct continuation of several other studies that we previously conducted. In Schneider et al. (2013), we used a more basic version of the TUI described in this paper and a simpler experimental design (Fig. 2). The design of the system is described in Schneider et al (2013). We built on De Jong and Van Joolingen (1998)'s principles to offer students a space where they could iteratively formulate hypotheses about how the human brain processes visual information, create a lesion on a virtual brain, observe the effect on its visual field, and predict what would happen for other lesions.

In this study, 28 college-level students interacted with a TUI either *before* or *after* following a more standard kind of instruction (i.e., reading a textbook chapter). We found that students who used the TUI before reading the text ("table-text" shown subsequently) outperformed students who used the TUI after reading the text ("text-table" shown subsequently). Those results suggest that TUIs are best used following the PFL framework, rather than a "Tell & Practice" kind of instruction. Additionally, we found that this difference was associated with the quality of students' verbalizations: participants in the "Table-Text" were more likely to generate rules explaining the way the human brain processes visual information, and this behavior was a significant mediator for students' learning gains (computed from a pre- and post-test). The learning tests used contained five questions about the terminology used, five questions on students' conceptual understanding of human vision, and five transfer questions.

Our findings suggest that students in the "Table-Text" group were more likely to become curious about the human brain and process the content taught more deeply; students in the "Text-Table" group, on the other hand, were more likely to

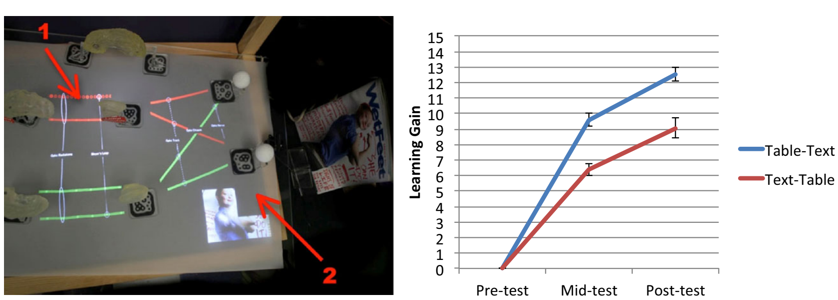


Fig. 2 On the left side, the first version of BrainExplorer. Students manipulate a small-scale brain where the TUI displays visual pathways between brain regions (in this example, a user cut the outer left optical radiation (1) and the top right corner of the visual field (2) is not perceived by the brain). On the right side, a plot shows the learning gains of students who followed a "Tell & Practice" ("Text-Table") versus a PFL ("Table-Text") kind of instruction. Whiskers show standard errors

gain a superficial understanding of those concepts and rely on rote memorization. This last observation was corroborated by another study that we conducted (Schneider and Blikstein 2016). In this experiment, we used a similar experimental design with a TUI teaching students about concepts in probability (Fig. 3, left side); in this particular study, we used video lectures instead of a textbook chapter. We found similar results (Fig. 3, right side) suggesting that TUIs are best used as a discovery learning environment. The learning tests used in this study contained questions from a standard textbook on combinatorics and probability. We found qualitative evidences suggesting that students in the "video → table" group were more likely to focus on memorizing formulas while watching the lecture and on trying to recall them while interacting with the TUI. On the other hand, students in the "table → video" group were more likely to explore the concepts taught using the TUI, for instance by spending more time interacting with the visualizations.

Those two studies suggest that learning is increased when a standard kind of instruction (e.g., reading a textbook chapter or watching a video lecture) is *preceded* (rather than followed) by a hands-on activity on a TUI. Coding students' utterance in study 1 suggested that students in the latter group were more likely to generate rules explaining the mechanisms associated with visual processing in the human brain. Qualitative excerpts from study 2 suggested that students in the former group were more likely to memorize and recall information when following a "Tell & Practice" kind of instruction. The current study replicates those results and provides two additional contributions to this line of research. First, we designed additional measures to further explain the difference found in the studies described previously; for instance, by capturing students' curiosity and the extent to which they processed the material taught. Second, we compared the efficiency of a TUI with a traditional PFL pen and paper activity (i.e., groups of students exploring those concepts with CCs). Since TUIs are expensive to design and to build, it is important to know whether pen and paper activities might be as good as TUIs, or if they provide other advantages. In the following section, we describe how we operationalized those two research questions.

General Description of the Experiment

Our goal with this paper is twofold. First, we want to replicate previous results (Schneider et al. 2013; Schneider and Blikstein 2015) showing that providing students with an opportunity to build prior knowledge with a TUI before following a more standard type of instruction is beneficial to learning (compared to a traditional T&P approach, where students are first taught some concepts and then practice their understanding of those concepts on a TUI). Secondly, we want to see how TUIs compare to CCs in terms of preparing students for future learning: in other words, is it worth spending time and energy building interactive hands-on activities? Do they provide any benefits compared to traditional "pen and paper" activities?

We designed the following experiment to investigate those two lines of research (see Fig. 7 for a complete description of the experimental design). Dyads in the control group (T&P condition) first read an abridged version of a textbook chapter on the human visual system and then completed another activity where they had to apply their new knowledge. Dyads in the treatment group (PFL condition) first discovered those concepts in a hands-on activity and then learned about them in a more traditional way, i.e., by reading an abridged textbook. Based on the PFL framework, our main hypothesis is that the treatment group should achieve higher learning gains compared to the control group. Additionally, we crossed another factor in our experimental design: the hands-on activity was either a set of CCs or a TUI. We did not have a strong hypothesis regarding this comparison, but we expected the TUI group to slightly outperform the CCs group on the final learning test; as mentioned, previous work suggests that TUIs facilitate students' collaboration, exploration, and engagement. Participants were counter-balanced across conditions. We designed measures to look at three potential predictors for learning: engagement (using a questionnaire with validated psychometric properties), curiosity (by having students list all the questions they would like to see answered after the first activity), and quality of their mental models (by asking students to draw a model summarizing their understanding of the topic taught after the first activity). Our goal was to see if

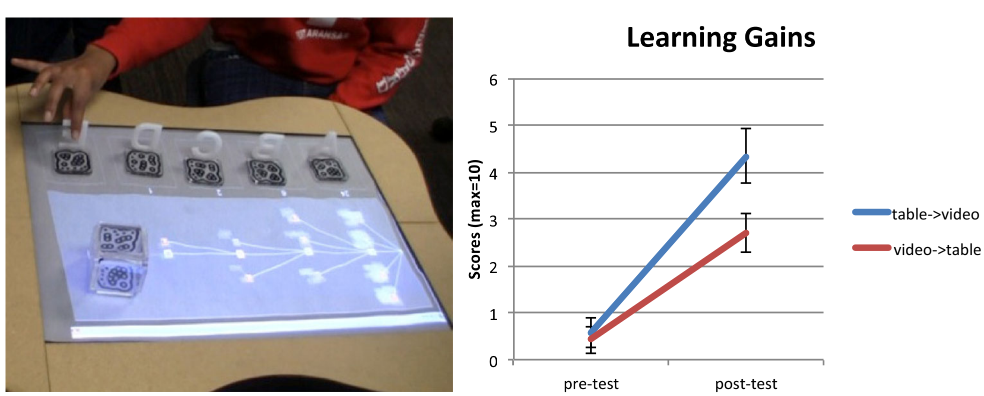


Fig. 3 On the left side, a TUI that we developed to promote students' exploration of concepts in probability (i.e., combinatorics). On the right side, students' scores on the pre- and post-tests

students would differ on those measures between our different conditions.

The Role of Collaboration in This Study

The role of students' collaboration is not central to our study. We decided to group students in pairs for theoretical and practical reasons: theoretical because several studies have shown that students' interactions can facilitate constructivist learning (e.g., Schwartz 1995); practical because students' interactions externalize their cognition and problem-solving strategies. Thus, we opportunistically decided to work with dyads to (1) facilitate their investigation of the material taught and (2) make learning processes visible to researchers. In our analyses, we qualitatively analyze students' collaboration merely as a way to illustrate key differences between our different experimental conditions and suggest avenues for future work.

Hypotheses

Concerning those measures (engagement, curiosity, quality of their mental models), our hypotheses are as follows: first, the main difference between the PFL and T&P groups should be about the quality of students' mental model and their curiosity. Since the PFL framework is about helping students construct some prior knowledge, we believe that the PFL activity should better help them build a "proto-theory" of how the human visual system works; this difference should be reflected in their drawings after the first activity. In this process, students should be more likely to ask themselves questions and develop their curiosity about the topic taught. This should then help them make sense of a standard instruction (reading a text about the visual system) during the second activity. On the other hand, we expect students in the control group (T&P) to focus on memorizing the content of the text and spending their time recalling this information when completing the second activity (i.e., using the TUI or working on the CCs). Secondly, we expect the main difference between the TUI and CCs groups to be about their levels of engagement. Since TUIs have been shown to promote collaborative hands-on learning (Schneider et al. 2011), we believe that the physicality of the system should invite dyads of students to be less intimidated by the complexity of the domain taught and explore the problem space to a greater extent. Those two hypotheses motivated the use of measures described subsequently (i.e., middle test measuring the quality of the students' mental model and their curiosity, and a questionnaire measuring their engagement). Finally, we did not have any hypothesis regarding an interaction effect between our two factors (i.e., we do not have any reason to believe that CCs or the TUI should have a differentiated effect on students when following the PFL or T&P instructional sequence.) Our hypotheses are as follows:

> Hypothesis 1.0: students learn more when completing a hands-on activity *before* rather than *after* following a standard kind of instruction (i.e., reading a textbook chapter)
> Hypothesis 1.1: this difference is associated with higher levels of engagement, curiosity, and better mental models.
> Hypothesis 2.0: students learn more when completing the hands-on activity with a TUI rather than with a set of CCs.
> Hypothesis 2.1: this difference is associated with higher level of engagement.

We describe on how we operationalized those research questions in the following section. Additionally, we will explore students' behaviors when interacting with the TUI by looking at the log files generated by the system and qualitatively analyze students' collaboration by looking at the videos recorded during the experiment.

Methods

Participants

Forty students from a community college participated in this study (13 males, 27 females; mean age = 21.28, $SD = 4.08$). Students signed up for the study as part of a psychology class. The only prerequisite for participating was to have no prior knowledge on the topic taught (neuroscience and the visual system). Dyads of students were randomly assigned to each experimental condition.

We checked the students' grade point average (GPA), since populations in community colleges are known to be heterogeneous. Because there was an interaction effect: $F(1,16) = 4.46$, $p = 0.049$ (Fig. 4) between the two levels of our factors, we used GPA as a covariate for the analyses presented in the "Results" section.

Material

This study included three different activities (Fig. 7). In the PFL condition, students first explored the domain taught with either a TUI or a set of CCs. The TUI is described subsequently (Fig. 6); the CCs (Fig. 5) included six diagrams of the human brain, each one featuring a different lesion. Half of the CCs showed the correct answer to students (i.e., cases 1, 3, and 6). The right column showed potential answers for the remaining cases.

After finishing the first activity, students answered the following two questions: (1) "By the end of this first activity, what are the questions that you would like to see answered about the way the human brain processes visual information?" and (2) "Please draw a simple model that summarizes your

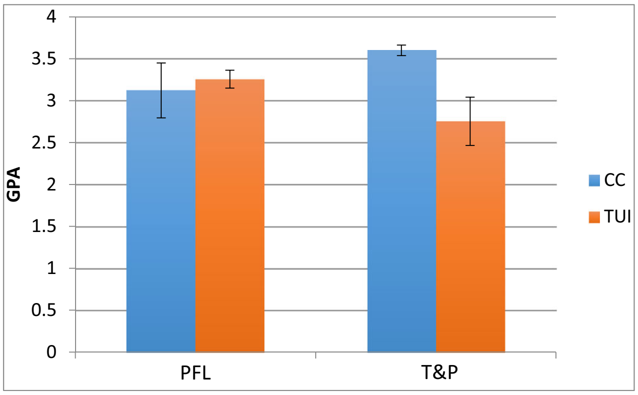


Fig. 4 Student's GPA in each experimental condition. CCs stands for Contrasting Cases, TUI for Tangible User Interface, PFL for "Preparing for Future Learning," and T&P for "Tell-and-Practice." The interaction effect is significant ($p = 0.049$)

understanding of the way the human brain processes visual information." During the second activity, students read an abridged version of a textbook chapter explaining how visual information is processed in the human brain (available at http://goo.gl/47RIwv). They then took a post-test (similar to the pre-test) that included questions on the terminology used (students had to correctly label different brain regions and neural pathways on a given diagram), on the effect of various lesions on the visual field (given a particular lesion on a brain diagram, students had to predict its effect on a person's visual field among several answers—similar to the options shown on the right side of Fig. 5), and on more general scenarios involving human vision (transfer questions; e.g., "patient X is likely to have a lesion in region Y of the brain; should he be allowed to drive?"). Finally, we asked them to fill the engagement questionnaire designed by O'Brien et al. (2008). This questionnaire was developed for researchers in HCI (Human–Computer Interaction) and measures six dimensions of an activity that relates to users' engagement (33 items on a Likert scale): focused attention (9 items; e.g., "I forgot about my immediate surroundings while doing X"), perceived usability (8 items, e.g., "I felt frustrated/discouraged/annoyed while doing X"), esthetics (5 items, e.g., "X was aesthetically appealing"), endurability (5 items, e.g., "I consider my experience with X a success"), novelty (3 items, e.g., "I continued to do X out of curiosity"), and involvement (3 items, e.g., "I was really drawn into X"). In the T&P condition, the order of the learning phases was reversed (students first read the text and then completed the TUI or CCs activity).

The TUI used in this study is an improved version of a system previously developed in our lab (Schneider et al. 2013), called BrainExplorer. BrainExplorer (Fig. 6) allows students to physically deconstruct a small-scale replica of a brain while an interactive tabletop displays visual pathways between brain regions, shown as red and green lines on Fig. 6. Users can then cut those pathways with their fingers to create lesions and observe their effect on the visual field of a subject (similar to what is shown on Fig. 2). Two eyes, with a webcam placed between them, capture the field of view of this brain and show occlusions on the corresponding visual field—

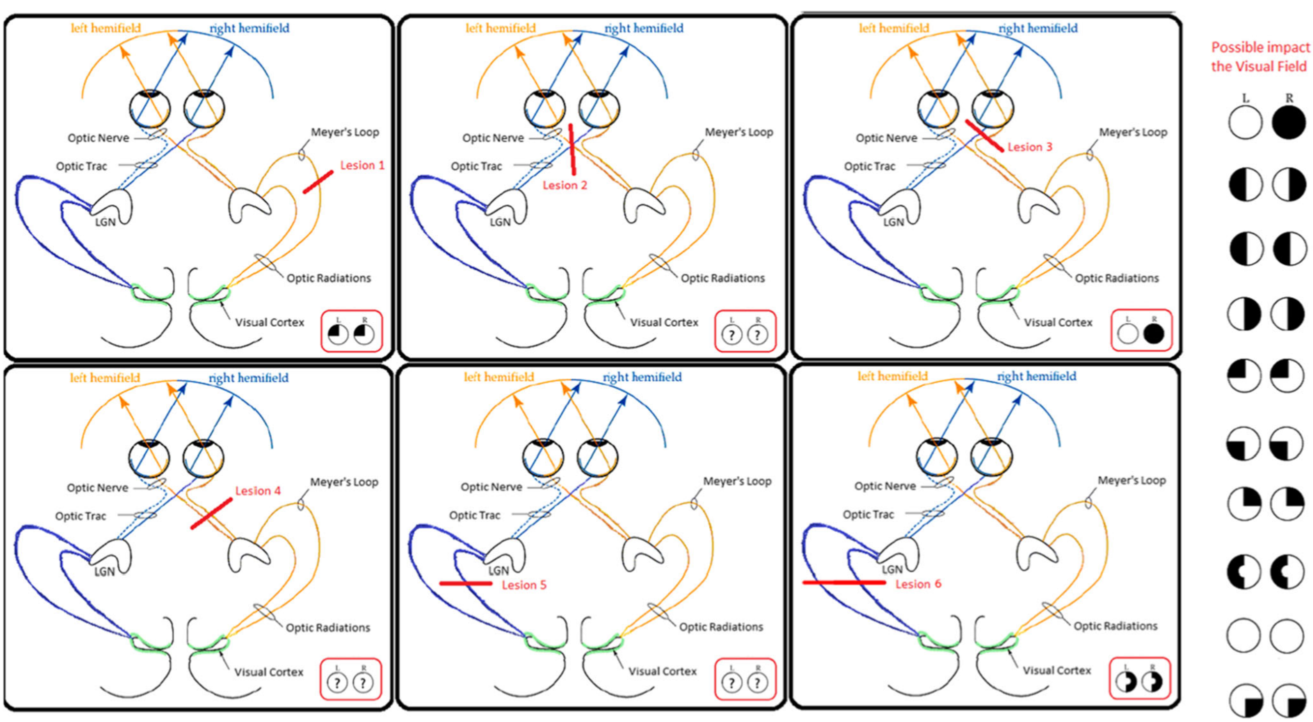


Fig. 5 The set of CCs used in the study. Answers are given for diagrams 1, 3, and 6. Potential answers for the remaining cases are shown on the right column. The diagram used in the study is accessible at http://www.scribd.com/doc/98921800 (last access 01/2018). Originally retrieved from Washington University in St-Louis (http://thalamus.wustl.edu/)

displayed on the bottom right corner of the screen. The new version of this system allows collaborative exploration of the visual system, since users create lesions on the multi-touch interface (as opposed to using a single infra-red pen). Figure 6 (top) shows the different brain regions and visual pathways that users could cut, and Fig. 6 (bottom) shows two students reflecting on the effect of a lesion on the visual field of this brain.

Design

We used a 2 × 2 between-subjects experimental design (Fig. 7). The first factor had two different hands-on conditions: TUI and CCs. The second factor sequenced the two learning activities in different ways: either with the hands-on activity first (PFL) or second (T&P).

Procedure

The experimenter ran students in groups of two in a private room. Upon their arrival, the experimenter welcomed

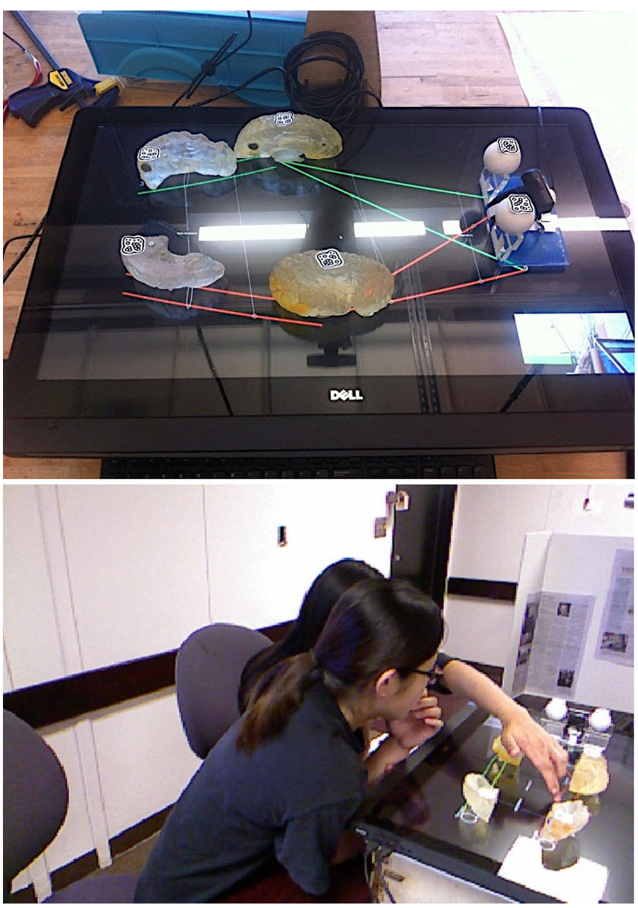


Fig. 6 The TUI used in this experiment (BrainExplorer). The system on the top shows the tangibles that students can manipulate with the field of view of this brain (bottom right corner of the table). The image on the bottom shows two students interacting with the system during our study

them and told them that they would complete two small learning activities in groups. They were also told that the topic taught was about neuroscience and the human visual system. After completing a pre-test, students completed two learning activities: In the PFL condition, they first did the hands-on activity (i.e., TUI or CCs) and then read a text about the visual system. In the T&P condition, they first read the text and then practiced their understanding of the topic on the hands-on activity. A second factor was crossed with those two conditions: students either worked with the CCs or TUI for the hands-on activity. Thus, referring to Fig. 3, students in the PFL + CCs condition discovered those concepts with CCs and then read a text. Students in the PFL + TUI condition followed the same procedure except that they used the TUI instead of the CCs. Students in the T&P + CCs first read a text and then applied the concepts they just learned on a set of CCs. Students in the T&P + TUI followed the same procedure except that they reinforced their understanding of the visual system by using the TUI. Each group spent exactly 15 min on the two activities. A small number of groups were done a few minutes early and were told to check their answers before moving on to the next step.

Between the two activities, the experimenter gave students two questions to answer individually (10 min): first, they had to list the questions that they wanted to see answered about the human visual system after the first activity (i.e., a measure of curiosity); second, they had to draw a model summarizing their understanding of the concepts taught (see the "Material" section for more information). The students individually completed a post-test (15 min) and were thanked for their participation.

Coding

The pre-tests and post-tests were coded in a binary fashion (1 point for a correct answer, 0 point for an incorrect answer). For the middle test, we counted the number of questions students had and applied a simple rating scheme to their models: 1 = no useful information shown (no or very few keywords); 2 = some useful information drawn, mostly keywords about the terminology used (no or little conceptual understanding of the effect of lesions on the visual field); 3 = significant signs of understanding (i.e., the drawing does not only contain keywords, but also sentences describing mechanisms by which the human brain processes visual information). Figure 8 provides an example for each category. A second researcher double-coded ~23% of the models drawn by the participants (10/43), and we found an inter-rater reliability of 0.85 using Krippendorff's alpha (Hayes and Krippendorff 2007).

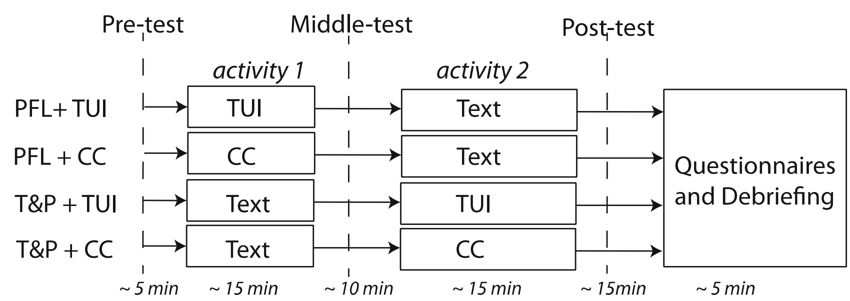


Fig. 7 The four experimental conditions of our study (factor 1 = PFL versus T&P; factor 2 = TUI versus CCs)

Results

Since our samples are not independent (i.e., members of a dyad influenced each other) and since the intraclass correlation for the learning test is significant ($p < 0.001$; Kenny et al. 2006), it is advised to conduct analyses at the dyad level. Since this reduced our statistical power, we will also report results where $p < 0.1$ with a moderate effect size. We start by testing our hypotheses, then we will explore the logs generated by the TUI to compare how students interacted with the system when following a T&P and PFL sequence, and finally illustrate differences between the TUI and CCs in the PFL sequence using qualitative excerpts.

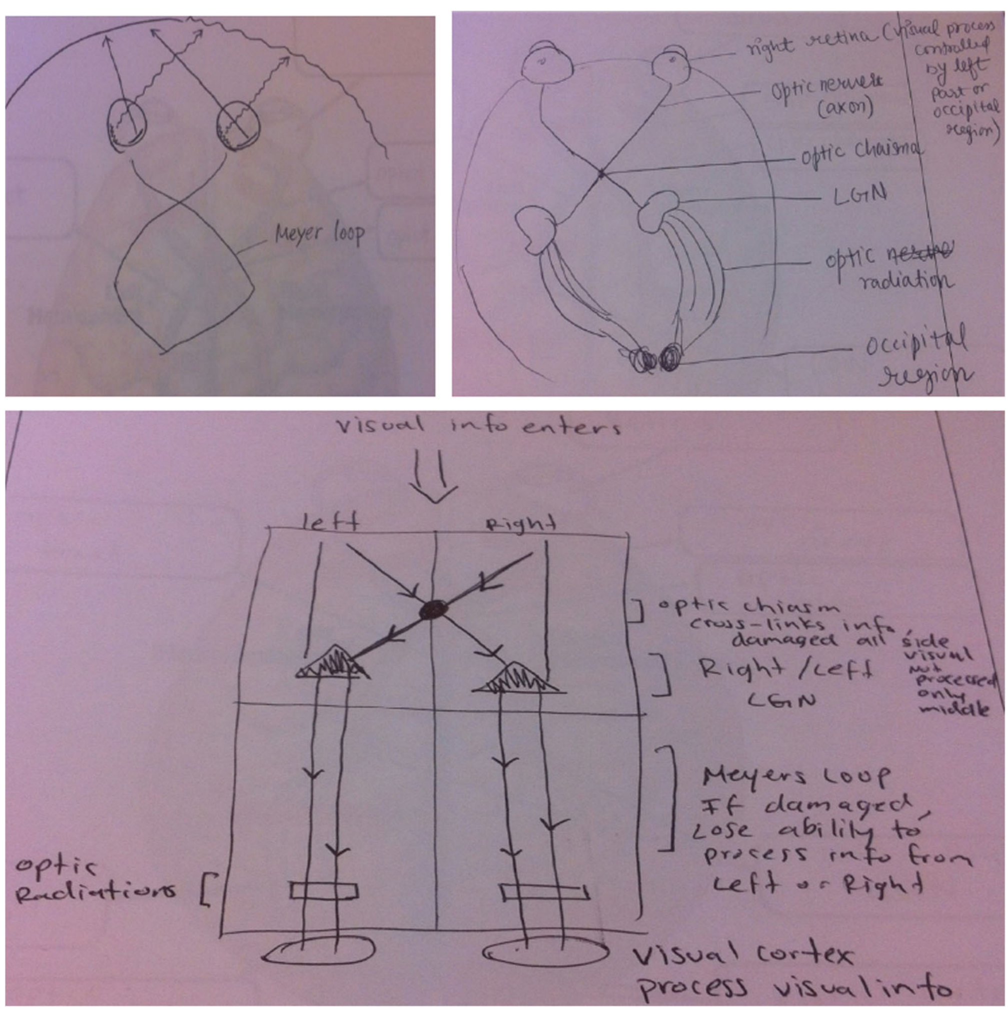


Fig. 8 Three examples of models drawn by our participants. The model on the top left received 1 point (= no or little useful information); the one on the top right received 2 points (= some useful information, mostly about the terminology); and the one on the bottom received 3 points (= clear signs of conceptual understanding)

Learning Gains (Hypotheses 1.0 and 2.0)

The results supported our two main hypotheses. Participants did not score any points on the pre-test, which confirms that they had no pre-existing knowledge on the topic taught. Thus, we only considered their results on the post-test. Scores are shown on Fig. 9. An ANCOVA revealed that students in the PFL condition outperformed students in the T&P condition: $F(1,16) = 15.45$, $p < 0.001$, Cohen's $d = 1.41$ (for the PFL group, $M = 12.95$, $SD = 3.15$; for the T&P group, $M = 7.85$, $SD = 4.06$). Additionally, students using the TUI outperformed students in the CCs group: $F(1,16) = 5.32$, $p = 0.036$, Cohen's $d = 0.48$ (CCs: $M = 9.35$, $SD = 3.90$; TUI: $M = 11.45$, $SD = 4.74$). All distributions were checked for normality and homogeneity of variance.

Curiosity and Mental Models (Hypothesis 1.1)

Additionally, we looked at the effect of our two factors (PFL vs. T&P, TUI vs. CCs) on the results of the middle-test: (1) the number of questions students would like to see answered about this topic (a proxy for students' curiosity) and (2) the quality of the model they drew—illustrated in Fig. 8 (as a reminder, scores varied among 1 pt = no or little useful information, 2 pts = some correct terminology, and 3 pts = clear signs of conceptual understanding). For the first factor, we found that students in the PFL group created significantly higher quality models compared to the students in the T&P group: $F(1,16) = 7.38$, $p = 0.016$, Cohen's $d = 1.32$ (for the PFL group, $M = 2.00$, $SD = 0.56$; for the "T&P" group, $M = 1.35$, $SD = 0.41$). There was no significant difference in terms of the number of questions they asked themselves: $F(1,16) = 2.75$, $p = 0.118$, Cohens'$d = 0.81$. For the second factor (TUI vs CCs), both comparisons were not significant ($F < 1$). We then correlated those two measures with our main dependent variable. Both measures were significantly associated with higher learning gains: $r(20) = 0.40$, $p = 0.043$ for the number of questions students asked themselves and $r(20) = 0.55$, $p = 0.007$ for the quality of students' model.

Engagement (Hypotheses 1.1 and 2.1)

Finally, we looked at the engagement questionnaire administered at the end of the study (see Table 1). We found that participants in the PFL condition were more engaged (aggregate measure of all the items) than the students in the T&P group: $F(1,16) = 6.1$, $p = 0.025$, Cohen's $d = 1.05$. Results were significant on the "endurability" sub-dimension ($p = 0.045$) and marginally significant on the "Focus" ($p = 0.096$), "involvement" ($p = 0.068$), novelty ($p = 0.066$), and aesthetics ($p = 0.071$) sub-dimensions. Similarly, students in the "TUI" group were marginally more engaged compared to the "CCs" group: $F(1,16) = 3.4$, $p = 0.064$, Cohen's $d = 0.80$. They found the TUI to be more usable ($p = 0.033$) and rated the endurability and esthetics' dimensions marginally higher ($p = 0.108$ and $p = 0.114$, respectively). Across all participants, being engaged was significantly correlated with higher learning gains ($p = 0.015$); more specifically, involvement and endurability were the only sub-dimensions significantly correlated with students' learning ($p = 0.011$ and $p = 0.03$, respectively).

Linear Regression

We ran a linear regression to find how much variance of the learning gains our main predictors could explain (i.e., number of questions on the middle-test, complexity of students' mental model, endurability, involvement). We found that the quality of students' mental model was the strongest predictor ($\beta = 0.43$), followed by the endurability variable ($\beta = 0.30$), students' involvement ($\beta = 0.22$), and curiosity ($\beta = 0.12$). Altogether, these four variables explained more than half of the variance of students' learning gains: $R^2 = 0.58$, $F(4,18) = 4.78$, $p = 0.012$ (Table 2).

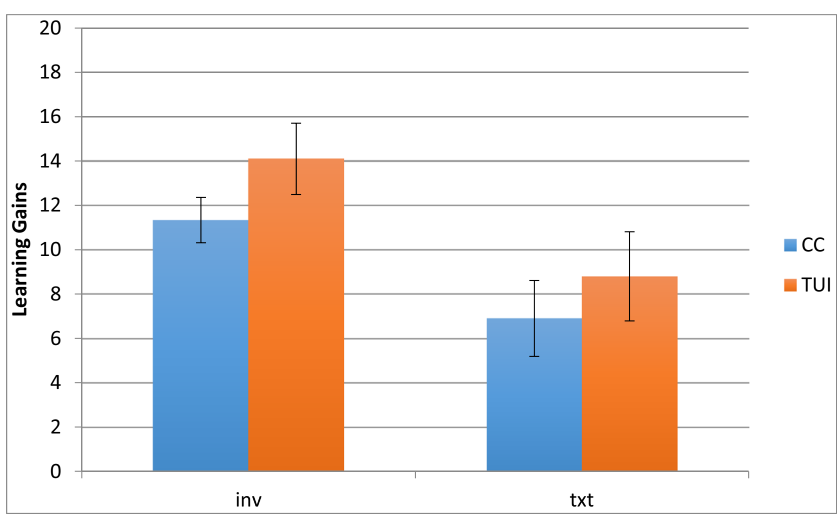


Fig. 9 Results on the learning test (standard errors shown). CC stands for Contrasting Case, TUI for Tangible User Interface, PFL for "Preparing for Future Learning," and T&P for "Tell-and-Practice"

Table 1 Engagement scores between our experimental groups

	Focus	Involvement	Novelty	Endurability	Esthetics	Usability	Engagement
TUI vs CCs	$F<1$	$F<1$	$F<1$	$F=2.9$ $p=0.108$ $d=0.45$	$F=2.8$ $p=0.114$ $d=0.73$	$F=5.4$ $p=0.033$* $d=0.85$	$F=3.4$ $p=0.064$ $d=0.80$
PFL vs T&P	$F=3.1$ $p=0.096$ $d=0.45$	$F=3.8$ $p=0.068$ $d=0.71$	$F=3.9$ $p=0.066$ $d=0.72$	$F=4.7$ $p=0.045$* $d=0.77$	$F=3.7$ $p=0.071$ $d=0.70$	$F<1$	$F=6.1$ $p=0.025$* $d=1.05$
Correlations with learning	$r=0.33$ $p=0.152$	$r=0.56$ $p=0.011$*	$r=0.33$ $p=0$ .154	$r=0.49$ $p=0.030$*	$r=0.36$ $p=0.122$	$r=0.21$ $p=0.380$	$r=0.53$ $p=0.015$*

Notes: Rows 2 (TUI vs CCs) and 3 (PFL vs T&P): MANCOVA between the two levels of our two factors on the dimensions of the engagement questionnaire. Row (Correlations with learning) 4 reports correlations with students' learning gains. Degrees of freedom are indicated as discussed previously: $r(20)$ for the correlations and $F(1,16)$ for the F-tests; d is Cohen's d

Log Analyses Using Learning Analytics Techniques

As a post-hoc analysis, we explored the log files collected by the TUI to describe how participants from the PFL condition differed from the participants in the T&P condition. The subsequent analyses are exploratory, and their goal is to suggest potential explanation for the differences observed previously; since the sample size for those analyses is small (half of our original pool of participants since we do not have logs for the participants who worked on the CCs), the patterns reported here need to be replicated in future studies to be conclusive. They currently only offer support for formulating future hypotheses.

Those logs describe how many times and in which order students made a lesion on the visual pathways displayed between brain regions. We then aggregate similar types of lesions into four different categories (Fig. 10): when the brain loses half of its vision vertically, horizontally, or by a quarter. There is a fourth category, labeled "exception," that describes atypical cases (e.g., when lesions cannot be mirrored because they involve both sides of the brains).

On average, we found that students in the PFL condition made nearly twice as many lesions as students in the T&P condition: $F(1,10)=5.76$, $p=0.04$, Cohen's $d=1.63$ (T&P $M=34.00$, $SD=15.96$; PFL $M=66.00$, $SD=22.68$). As shown in Fig. 11, we can observe that this difference is mostly driven by lesions resulting on a loss on the quarter of the visual field. Indeed, when testing for significance on each separate category, we found this difference to be statistically significant only for the "quarter" category: $F(1,10)=5.97$, $p=0.04$, Cohen's $d=1.66$ (T&P $M=12.00$, $SD=8.51$; PFL $M=29.33$, $SD=12.06$).

Table 2 Linear regression with students' learning gains as the dependent variable ($R^2=0.58$)

Variable name	β	*t* test
Endurability ("students' perception of success")	0.304	$t(15)=1.35$, $p=0.199$
Involvement	0.224	$t(15)=0.94$, $p=0.361$
Curiosity (number of questions asked)	0.119	$t(15)=0.60$, $p=0.555$
Quality of students' mental model	0.434	$t(15)=2.34$, $p=0.035$*

We then analyzed the students' *sequence* of actions when performing those lesions. To do so, we generated two Markov chains (Fig. 12). Markov chains are generated from a sequence of "states" pre-defined by the researcher. For example, in prior work, Schneider and Blikstein (2015) identified when students were in an "active," "semi-active," or "passive" state from motion sensor data when completing a hands-on learning task. They then built transition matrix representing student's probability of staying in one state or transitioning to a new one. This allowed them to identify cycles of cognition and action, which are known to support collaborative problem-solving (Tschan 2002). Markov chains provide a visual representation of the probabilities of staying in a state versus transitioning to a new one.

Proportionally, we found that students in the PFL condition were much more likely to transition between the "horizontal" state and the "vertical" state (39% versus 0.01% for the students in the T&P condition); they were also twice as likely to transition between the "quarter" state and the "horizontal" state (53% versus 27% in the T&P condition). In the T&P condition, this probability mass was redirected toward the "exception" state, which suggests a more random sequence of actions; in the PFL condition, it is likely that students were trying to understand how horizontal lesions were connected to vertical lesions and loss of a quarter of the visual field, which is critical for understanding on how the human brain divides visual information into vertical and horizontal halves.

Qualitative Excerpts

The quantitative results show the frequency of students' actions and the probability of moving between states for each experimental group, but they do not reveal the mechanisms behind those transitions. To illustrate how the dyads interacted, we present in this section two qualitative excerpts to highlight key differences between the two groups, CCs/TUI

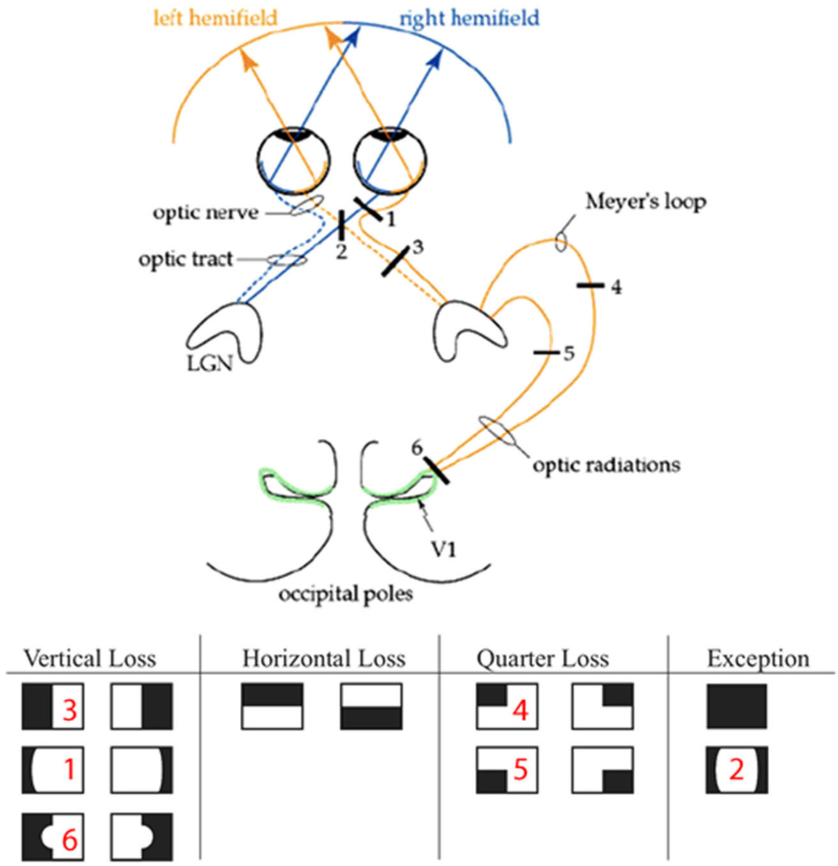


Fig. 10 The four categories of lesions that could be made on the TUI. In the diagram, we show where the lesion would have to take place for visual impairments on one side of the brain). Horizontal losses could be made by combining lesion 4 or 5 on both sides of the brain. The "exception" category represents corner cases (the first one can be made by cutting all connections before or after the LGN)

and PFL/T&P. We inspected the videos of the experiment and chose those two groups because they were representative of key differences we observed between the two conditions. These comparisons also describe how collaborative episodes unfolded based on the affordances of the activity.

TUI vs. CCs

Table 3 presents two exchanges that happened at the beginning of the experiment between group 18 (TUI–PFL condition) and group 24 (CCs–PFL condition). On the left side of Table 3,

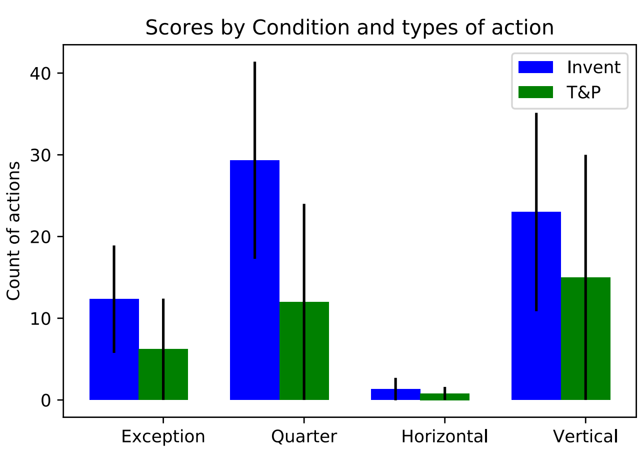


Fig. 11 Number of lesions for each category (whiskers show standard errors)

group 18 achieved a learning gain of 10.5; on the right side of Table 3, group 24 achieved a learning gain of 14.5. Both groups were randomly sampled from the dyads that completed the study. Those two groups obtained high learning gains between pre- and post-tests (compared to other groups).

An initial analysis of the video data of the experiment suggested that two those activities (CCs and TUI) successfully prompted productive discussion in these two dyads. We saw students explore the problem space ("my guess is that we should cut that one"), share information to build a common ground ("yes because the left side affects the right side and the right side affects the left side"), build hypotheses ("what happens if you just cut this"), question each other ("And this would affect the left side of both eyes?"), contradict each other and negotiate knowledge ("so I think that we have one rule. Opposite sides [of the brain] control opposite sides of the vision.") in rich ways. They were engaged by the material and collaborated with their peer even though they did not know each other beforehand. What is more, the way they collaborated varied based on their experimental condition.

For instance, Table 3 highlights the fact that students in the CCs group generated a massive amount of pointing to coordinate their attention. This observation is not anecdotal: we know from previous work that joint visual attention (JVA) is an important building block for establishing a common ground in social settings (Tomasello 1995). Pointing does not only

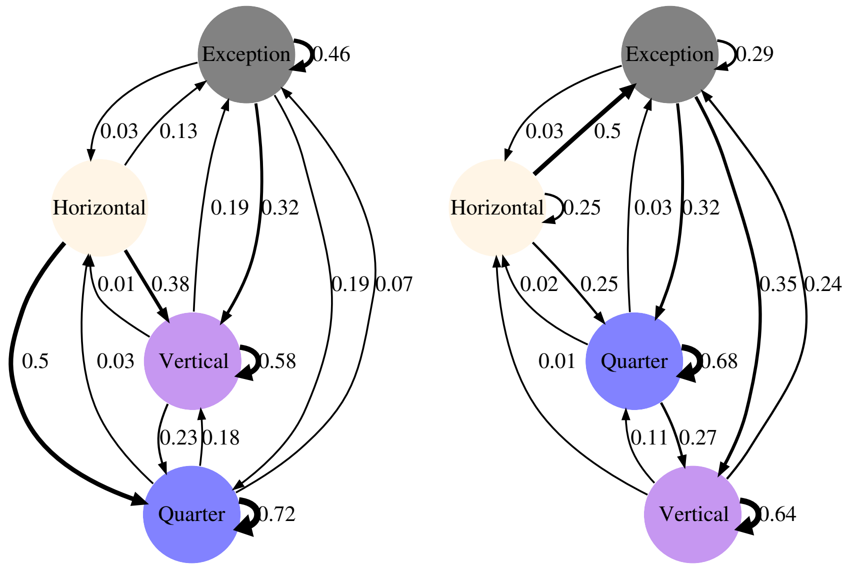


Fig. 12 Two Markov chains generated from each experimental condition (PFL on the left side and T&P on the right side). The "Quarter" state was the most visited space by students, as indicated in Fig. 10. Numbers between states indicate the probability of moving between states

externalize one's cognition, but also allows partners to synchronize their attention and contribute to the discussion. Thus it is not surprising that successful groups used a lot of pointing to complete the task. However, what is interesting is that the students in the TUI group *barely used any pointing*. Our observations suggest that the actions performed within the interface (i.e., cutting a visual pathway) fulfilled that function: using one's finger on the touch screen to create a lesion automatically attracted the attention of one's partner. This created a moment of joint attention, without having students purposefully point at an area of interest. Thus, it is possible that the TUI generated beneficial affordances for JVA, which might have had a beneficial effect on collaboration. This is a hypothesis that should be tested in future work, but points to an important benefit of interfaces that engage students in tangible manipulation.

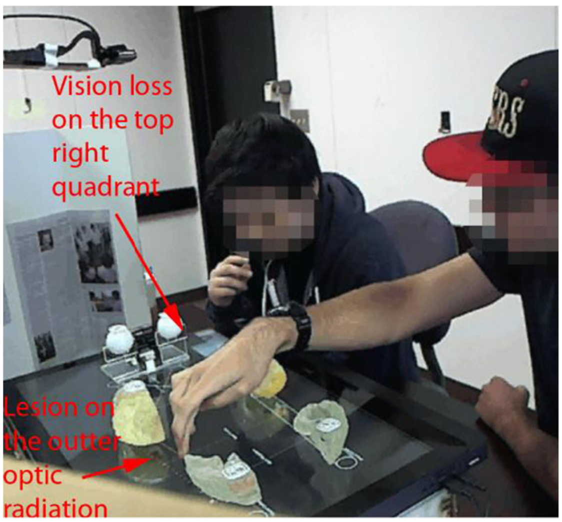
Another observation inferred from the data in Table 3 (and other videos) is that students in the CCs group tended to adopt a holistic approach and combine information from the different diagrams to infer the impact of a particular lesion on a visual field. Students in the TUI group, on the other hand, tended to adopt a more local approach by grouping and making sense of lesions in a sub-region of the brain (e.g., between the LGN and visual cortex). This suggests a more bottom-up approach in the TUI group, where pieces of information are slowly combined together to form a bigger picture. We hypothesize that students in this group might have achieved higher learning gains because this strategy worked particularly well in this context and for this content topic: the visual system of the brain is a case where the system can easily be decomposed into stable sub-systems (e.g., pathways processing the top/bottom, left/right quadrants of the visual field). It is the interaction of those subsystem that govern the human visual system, so it is reasonable to assume that a bottom-up approach is the optimal way of going about its study, and that the TUI primed students to adopt this more effective strategy. If this hypothesis is true, it also suggests that CCs would probably be more suited for problems that require a holistic, top-down approach where sub-system interacts with each other and are not stable, i.e., they vary their behaviors depending on the influence of other sub-systems. In future work, we plan to investigate which classes of problems or concepts that are more easily studied using CCs or TUIs (i.e., in other words, using a holistic or bottom-up approach).

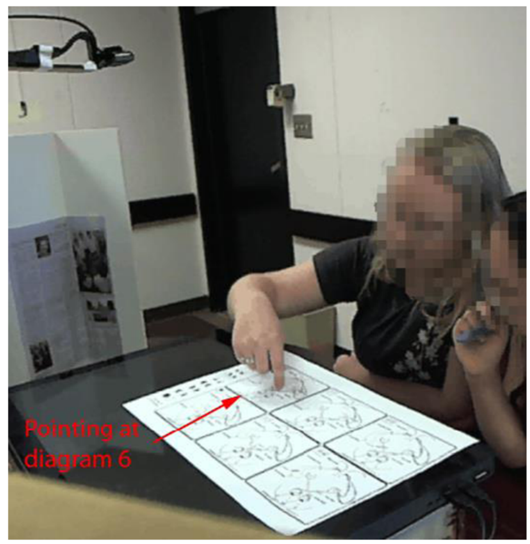
PFL vs T&P

Table 4 presents two exchanges that happened at the beginning of the experiment between Group 18 introduced previously (TUI–PFL condition) and Group 23 (TUI–T&P condition). Group 23 was randomly sampled from the dyads that completed the study and provided an illustrative contrast between the PFL and T&P conditions.

Table 4 reproduces a behavior found in previous studies (anonymous for blind review). Students in the T&P condition were more likely to use declarative statements and memorize information ("if you cut those two lines on the right hand side, it affects the left," "if you cut through the inside, it's lower. And outside, it's upper."), refer to the instructional material they had studied beforehand (e.g., "yeah, they talked about that in the article"), and less likely to question one another (anonymous for blind review). Conversely, students in the PFL conditions, tended to ask themselves questions (e.g., "I wonder if …"), to try things out (e.g., "what happens if you just cut …"), to express curiosity (e.g., "what's interesting is that …") or to make prediction and verify them ("yeah that's what I think. It would affect the top right quadrant") (anonymous for blind review). We found those differences to be representative of several groups across those two experimental conditions, and they are consistent with our quantitative results as well.

Table 3 Qualitative examples highlighting differences between the TUI and CCs groups in the PFL condition

<table>
<tr><th>Group 18 (TUI–PFL)</th><th>Group 24 (CCs–PFL)</th></tr>
<tr><td>- Let's try to cut one and see what it does. Which one should we cut?
- It's up to you
- I don't know... my guess is that we should cut that one [she cuts the outer optic nerve several times - nothing happens]
- not much [she cuts the inner optic nerve]. Does it just get blurry or something?
- it didn't change anything, actually
- [he cuts the inner optic radiation] that's interesting
- what happens if you just cut this [he cuts the outer optic radiation] that didn't change anything.
- hum. [he cuts the outer optic radiations on the right side].


- Lower left. Yes. What happens if you cut both of the lines? [he cuts both optics radiations]. Ok so that cuts both of them.
- We should do the same thing for this one [he cuts the two outer optic radiations on the left side]
- yeah on the other side.
- what's interesting is that it kind retained the other side on both ways. Except that semi-circle.
- Yeah. It's pretty cool.
[continued below]</td>
<td>- [looking and pointing at the 3rd diagram] that's a total blindness for the right eye. So it looks like that would affect the left side of the left eye and the right side of the right eye [pointing at the 2nd diagram]. yeah because it goes in between the blue line and the orange line, so I think it's that [pointing at the 2nd answer on the right]
- ok. And so that one is the right side [pointing at the 6th diagram]. This is a weird one because it's half of the middle
- yeah [laughing]. And so this also cuts the Meyer's loop, but on the left side?
- Hum. That's the [pointing at the 5th diagram]. And so this goes like this, and this goes like that [pointing back and forth between the 5th and the 6th diagram]. And so I think that with that one is like this [pointing at the 6th diagram]. So I think that it would affect more of the right of the right eye [pointing at the 3rd diagram].


- aie aie aie [laughing].
- and so maybe this would affect the right side of the left?
- yeah that's what I think. It would affect the top right quadrant [pointing at the 7th answer]
- hum. And this would affect the left side of both eyes? [pointing at the 4th diagram]
- yeah that's what I think
- maybe this one is that, and that one is that? [pointing back and forth between diagrams and answers]
- [...] yes because the left side affects the right side and the right side affects the left side.
[The group goes on to solve the remaining lesions.]</td></tr>
</table>

Discussion

The main goal of this paper was to show that technologically *enhanced hands-on activities* (i.e., TUIs) have the potential to increase students' learning in traditional constructivist activities. We also wanted to replicate previous results showing the benefits of generating prior knowledge before receiving formal instruction. Our results suggest that PFL activities have a large effect on students' learning. When prompted to explore a domain by themselves before reading an abridged textbook chapter, students developed more refined mental models, became more curious, more engaged, and perceived themselves as being more successful (compared to the students in the T&P group); additionally, those differences were associated with higher learning gains. The fact that students created better models based solely on their analyses of the CCs or by interacting with the TUI is promising: it shows that providing students in the control group (T&P) with a traditional

Table 4 Qualitative examples highlighting differences between the PFL and T&L groups in the TUI condition

Group 18 (TUI–PFL)	Group 23 (TUI–T&P)
- so I think that we have one rule. Opposite sides [of the brain] control opposite sides of the vision. - Yes. And then when we cut both of the lines, it shrinks. - I wonder if that makes a difference where you cut. [cuts twice on the same line] - wait, do that again? It's bottom... ah, ok. [he cuts the same line several times]. Now the other side? Ok, it's on the same side.	- if you cut those two lines on the right hand side, it affects the left. - the left side affects the right side - yes, inside - but depending on where it's cut, it will affect the upper or lower vision
	
- so even we were to cut back here, same spot. So where it's cut doesn't necessarily matter. It's just which line is cut. It just has to be anywhere on the line. Except those two outside ones that don't seem to be doing anything. - and the other lines we haven't cut yet?	- yeah, they talked about that in the article - yes it's one fourth I guess. if you cut through the inside, it's lower. And outside, it's upper. And the same thing applies to the interior. - and those two do nothing. [silence. They keep on cutting lesions]

explanation of how the human brain processes visual information prevented them from creating their own model; worse, they did not even internalize the ideal model from the textbook in a proper way. Analyses of automated logs of students' actions suggest that students in the PFL condition differed from the students in the T&P condition in two ways: first, they performed more lesions (especially the ones related to a visual loss of a quarter). Secondly, they were more likely to transition between lesions that provoked a horizontal, vertical or a loss of a quarter of the visual field. Those results suggest a different mindset in terms of students' exploration and perseverance, and well as strategies used between the T&P and PFL conditions: students were more likely to generate strong contrasts between lesions, instead of just repeating the same actions. Additionally, our results suggest that using a TUI as a preparatory activity has a positive effect on students' learning compared to studying CCs. Participants learned more when using the interactive system and felt more engaged compared to the CCs. Marginally significant effects suggest that students using the TUI felt more successful with the task ("endurability"), which was correlated with higher learning gains. Based on the engagement questionnaire, they also found the activity to be more esthetically appealing and the TUI to be more usable, but those measures were not associated with higher learning gains.

However, those measures do not provide us with the full picture; it is likely that the TUI had a beneficial effect on students' learning beyond their level of engagement. We suggest a few hypotheses suggested by our qualitative analyses. First, BrainExplorer, and similar TUI-based exploratory systems, provide students with "on-demand" or "just-in-time" information about the visual system. From our observations, we saw students develop different strategies when using the system. Some were more comfortable using a "bottom-up" approach (i.e., they started by analyzing the pathways from the eyes to the LGN, and then from the LGN to the visual cortex); some others followed the same approach, but in reverse; finally, some students started by making as many lesions as they could, and then focused on more specific pathways. In comparison, students using the CCs were much more holistic, top-down, and uniform in terms of the strategies they developed. In future work, we plan to compare the variety of strategies used in both conditions (TUI vs. CCs) to test this hypothesis. Secondly, it is conceivable that using a TUI had an indirect effect on students' collaboration, which in turn had a positive effect on learning. For instance, it is conceivable that the TUI facilitated joint visual attention (because students did not have to explicitly point at objects of interest—cutting visual pathways on the touch screen accomplished the same role), which might have helped them explore the problem space together (Schneider et al. 2011), share information, and build hypotheses (Shaer et al. 2011) with the TUI. We plan to code students' quality of collaboration and correlate those measures with learning gains in future work. Finally, it is possible that the results were caused by a novelty effect; this is the most natural explanation for the significant effect found between TUI and CCs, since most students had likely not interacted with a TUI in the past. However, the statistical analyses performed on the "novelty" dimension of the engagement questionnaire did not support this explanation. Future work will look more closely at this possible confounding variable. Finally, it should be noted that our goal for the qualitative analyses was to illustrate strategies that we perceived to be specific to experimental condition for the purpose of generating new hypotheses for future study. Future work will look at those differences in a more systematic way and provide stronger evidences for generalizability of those effects.

Limitations

It is worth mentioning that we do not take those results as evidence that TUIs are better learning activities than CCs beyond the scope of this experiment. Both activities can take many forms, and their efficiency is strongly influenced by a variety of design choices. Our findings merely suggest that the TUI introduced in this paper (BrainExplorer) seemed to promote higher learning gains compared to the CCs presented previously.

It should be noted that the effect size is between small and moderate. More analyses are needed both in collecting qualitative segments suggesting explanation for the higher learning gains found in the TUI (versus CCs) condition, and in analyzing the logs of the system to determine students' strategies when exploring BrainExplorer. Additionally, students in our experimental conditions had different GPA. We tried to control for it by using this variable as a covariate in our analyses, but it should be acknowledged that the learning activities presented in this paper might have benefited high/low GPA students in different ways.

Related to this point, it should be noted that our approach was more ecological than purely experimental. We wanted to compare two learning activities that could be used in a constructivist fashion (CCs vs TUI), where multiple variables differed between those two conditions (e.g., the tangible-based vs. paper-based activities, multiple static representations vs. one interactive systems). Thus, we cannot make a causal claim regarding which aspect of TUIs was most beneficial to learning. Our hypothesis (to be tested in future studies) is that tangibility supported collaboration, which indirectly impacted learning (as suggested by Shaer et al. 2011 and Schneider et al. 2012); while interactivity and dynamic feedback supported students' exploration of the domain, which helped them more efficiently go through cycles of predicting, testing and verifying hypotheses.

A final limitation of our study is the small number of participants who took part in this experiment. Most analyses were conducted by comparing the two levels of one factors (i.e., CCs vs TUI or PFL vs T&P). At the individual level, each condition contains 20 participants which is reasonable from a statistical point of view. At the dyad level, each condition contains ten groups which decreases our statistical power and limits the generalizability of our findings. This limitation should be kept in mind when interpreting the results presented in this paper.

Conclusion

The motivation of this paper was to compare two collaborative learning activities (i.e., CCs and TUI) used either in a typical "Tell & Practice" (T&P) or Constructivist fashion (PFL). As mentioned in the introduction, constructivist activities are difficult to design and evaluate. The goal of this paper was to test the design of our system and compare it with state of the art constructivist frameworks (i.e., the PFL framework and CCs). To this end, we crossed those two approaches with our learning activities to discover the optimal combination for increasing students' learning gains. Our results suggest that, under certain circumstances, *minimally guided instruction can be beneficial to learning*. We found this intervention to significantly increase the quality of students' mental models, which had a positive effect on their learning. This measure, associated with students' curiosity, involvement and perception of being successful at the discovery task, predicted more than half of the variance of their scores on the post-test. This finding shows the positive effect of using constructivist-inspired preparatory learning activities for learning scientific concepts. Those results, combined with others (e.g., Schwartz and Bransford 1998; Schwartz and Martin 2004), confirm that there is still a considerable gap between educational research (that advocates a constructivist view of students' learning) and regular classroom instruction (that still prevalently use a "T&P" framework). The contribution of this paper is to propose an initial step toward closing this gap, but our results do suggest that combining the affordances the new technologies (e.g., TUIs) with existing educational frameworks (e.g., PFL) can provide students with compelling, carefully-crafted hands-on learning experiences that prepare them for future learning.

Acknowledgements We gratefully acknowledge grant support from the National Science Foundation (NSF) for this work through the CAREER Bifocal Modeling grant (NSF # 1055130).

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical IRB (institutional review board) of Harvard and Stanford University. Informed consent was obtained from all individual participants included in the study.

Conflict of Interest Bertrand Schneider and Paulo Blikstein declare that they have no conflict of interest.

References

Blikstein, P., & Wilensky, U. (2009). An atom is known by the company it keeps: a constructionist learning environment for materials science using agent-based modeling. *Int J Comput Math Learn, 14*(2), 81–119.

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: a simple proposal with multiple implications. *Review of Research in Education, 24*(1), 61–100.

Cuendet, S., Dehler-Zufferey, J., Ortoleva, G., & Dillenbourg, P. (2015). An integrated way of using a tangible user interface in a classroom. *International Journal of Computer-Supported Collaborative Learning, 10*(2), 183–208.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research, 68*(2), 179–201.

Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning, 6*(4), 491–514.

Edwards, L. D. (1995). Microworlds as representations. In *Computers and exploratory learning* (pp. 127–154). Berlin Heidelberg: Springer.

Falcão, T. P., & Price, S. (2011). Interfering and resolving: how tabletop interaction facilitates co-construction of argumentative knowledge. *International Journal of Computer-Supported Collaborative Learning, 6*(4), 539–559.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77–89.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379–424.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford Press.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661–667.

Kynigos, C. (2007). Half-baked logo microworlds as boundary objects in integrated design. *Informatics in Education-International Journal, 6*, 335–358.

O'Brien, H. L., Toms, E. G., Kelloway, E. K., & Kelley, E. (2008). Developing and evaluating a reliable measure of user engagement. *Proceedings of the American Society for Information Science and Technology, 45*(1), 1–10.

Olson, I. C., & Horn, M. S. (2011). Modeling on the table: agent-based modeling in elementary school with NetTango. In Proceedings of the 10th International Conference on Interaction Design and Children (pp. 189–192). ACM.

Papert, S. (1980). *Mindstorms: children, computers, and powerful ideas*. New York: Basic Books, Inc.

Piaget, J. (1928). *The language and thought of the child*. New York: Harcourt.

Piper, A. M., & Hollan, J. D. (2009). Tabletop displays for small group study: affordances of paper and digital materials. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1227–1236). ACM.

Schneider, B., & Blikstein, P. (2015). Unraveling students' interaction around a tangible Interface using multimodal learning analytics. *International Journal of Educational Data Mining, 7*(3), 89–116.

Schneider, B., & Blikstein, P. (2016). Flipping the flipped classroom: a study of the effectiveness of video lectures versus constructivist exploration using tangible user interfaces. *IEEE Transactions on Learning Technologies, 9*(1), 5–17.

Schneider, B., Jermann, P., Zufferey, G., & Dillenbourg, P. (2011). Benefits of a tangible interface for collaborative learning and interaction. *Learning Technologies, IEEE Transactions on, 4*(3), 222–232.

Schneider, B., Strait, M., Muller, L., Elfenbein, S., Shaer, O., & Shen, C. (2012). Phylo-Genie: engaging students in collaborative 'tree-thinking' through tabletop techniques. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 3071–3080). ACM.

Schneider, B., Wallace, J., Pea, R., & Blikstein, P. (2013). Preparing for future learning with a tangible user interface: the case of neuroscience. *IEEE Transactions on Learning Technologies, 6*(2), 117–129.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences, 4*(3), 321–354.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition & Instruction, 16*(4), 475–522.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *C&I, 22*(2), 129–184.

Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., & Wang, H. (2011). Enhancing genomic learning through tabletop interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2817–2826). ACM.

Skinner, B. F. (1986). Programmed instruction revisited. *Phi Delta Kappan, 68*(2), 103–110.

Thille, C., Mitchell, J., & Stevens, M. (2015, September 22). What we've learned from MOOCs. Retrieved from https://www.insidehighered.com/views/2015/09/22/moocs-are-no-panacea-they-can-help-improve-learning-essay

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: its origins and role in development* (pp. 103–130). Hillsdale: Lawrence Erlbaum Associates, Inc..

Tschan, F. (2002). Ideal cycles of communication (or cognitions) in triads, dyads, and individuals. *Small Group Research, 33*(6), 615–643.

Wilkerson-Jerde, M., Wagh, A., & Wilensky, U. (2015). Balancing curricular and pedagogical needs in computational construction kits: lessons from the DeltaTick project. *Science Education, 99*(3), 465–499.

Wise, A. F., Antle, A. N., Warren, J., May, A., Fan, M., & Macaranas, A. (2015). What kind of world do you want to live in? Positive interdependence and collaborative processes in the tangible tabletop land-use planning game Youtopia. In Proceedings of the 11th International Conference on Computer Supported Collaborative Learning—Volume 1 (pp. 236–243). International Society of the Learning Sciences.