



Multimedia resources designed to support learning from written proofs: an eye-movement study

Somali Roy¹ · Matthew Inglis¹ · Lara Alcock¹ 

© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract This paper presents two studies of an intervention designed to help undergraduates comprehend mathematical proofs. The intervention used multimedia resources that presented proofs with audio commentary and visual animations designed to focus attention on logical relationships. In study 1, students studied an e-Proof or a standard written proof and their comprehension was assessed in both immediate and delayed tests; the groups performed similarly at immediate test, but the e-Proof group exhibited poorer retention. Study 2 accounted for this unexpected result by using eye-movement analyses to demonstrate that participants who studied an e-Proof exhibited less processing effort when not listening to the audio commentary. We suggest that the extra support offered by e-Proofs disrupts the processes by which students organise information, and thus restricts the extent to which their new understanding is integrated with existing knowledge. We discuss the implications of these results for evaluating teaching innovations and for supporting proof comprehension.

Keywords Proof · Intervention · Comprehension · Reading · Eye movements · Experiment

1 Introduction

Most students learn about proving at least in part via encounters with written proofs. In much undergraduate instruction, this is obvious: students are often expected to learn by watching and listening as a lecturer presents proofs at a board (Weber, 2004). Some students no doubt learn a lot from lecturers' explanations, which typically use examples, diagrams, and dynamic physical gestures, and are rich in references to prior knowledge (Greiffenhagen, 2008; Mills, 2014; Pritchard, 2010). But it is also true that students can fail to follow well-structured

Electronic supplementary material The online version of this article (doi:10.1007/s10649-017-9754-7) contains supplementary material, which is available to authorized users.

✉ Lara Alcock
l.j.alcock@lboro.ac.uk

¹ Mathematics Education Centre, Loughborough University, Loughborough LE11 3TU, UK

presentations because their thinking is derailed by unfamiliarity with new notation or by an early point of confusion (Gabel & Dreyfus, 2013), or because they attend to algebraic manipulations but fail to grasp the conceptual ideas that the lecturer is attempting to convey (Lew, Fukawa-Connelly, Mejía-Ramos, & Weber, 2016). For many, the primary product of lecture attendance is a set of notes that is competently written but not yet understood, and learning requires making sense of these during independent study.

In other pedagogical approaches, contact time is focused more explicitly on engaging students' reasoning skills. Mathematicians and mathematics educators have designed undergraduate instructional sequences to develop both content knowledge and the ability to construct mathematical arguments (Alcock & Simpson, 2001; Burn, 1992; Larsen, 2013; Stylianides & Stylianides, 2009a), and inquiry-based learning is attracting increasing attention (Laursen, Hassi, Kogan, & Weston, 2014). But critical reading remains a big part of a student's task. In inquiry-based learning and guided reinvention activities, students are deliberately encouraged to evaluate and critique mathematical arguments (Larsen & Zandieh, 2008). Further, students in almost all instructional situations are expected to use books or online materials to support their learning. Courses have reading lists and sometimes assigned texts and, other than in rigorously applied traditional Moore methods (Coppin, Mahavier, May, & Parker, 2009), students are expected to refer to such materials.

Reading to learn, then, is expected in undergraduate mathematics. And this is appropriate: we want educational systems to produce young people who can critically assess written arguments, and we want to develop the next generation of mathematicians and scientists—those people will spend much of their working lives learning from complex written materials. It is therefore appropriate to study interventions that might support effective mathematical reading. In this paper, we pursue this line of research, reporting two studies on the effects of an intervention designed to help undergraduate students read and understand mathematical proofs.

2 Research background

In undergraduate mathematics, deductive arguments abound, and students must make sense of these if they are to learn with understanding. There is room for debate about competent proof comprehension and validation—mathematicians do not always agree about the permissibility of “gaps” in even fairly simple arguments (Inglis, Mejía-Ramos, Weber, & Alcock, 2013). But they do tend to spot claims that are outright false and to recognise and reject arguments that prove the converse of a claim rather than the claim itself (Inglis & Alcock, 2012; Selden & Selden, 2003). Undergraduate students do not reliably do this: at least some are unduly swayed by the appearance of an argument (Segal, 2000), and some do not reliably spot logical errors in individual deductions or in the way that the global structure of a purported proof relates to a theorem statement (Alcock & Weber, 2005; Imamoğlu & Toğrol, 2015).

It is not obvious, however, whether these failures reflect fundamental inadequacy in logical reasoning or simple inattention (Hodds, Alcock, & Inglis, 2014; Weber, 2010; Weber & Mejía-Ramos, 2014). The latter is plausible because it appears that many students enrolled in transition-to-proof courses have not learned to read effectively, or even to see reading as a way to learn mathematics. There is evidence that some treat texts primarily as sources of problems and procedures to copy, tending to skip expository sections and miss many of the author's intended messages about the structure of the material and its central ideas (Randahl,

2012; Weinberg, Wiesner, Benesh, & Boester, 2010). Observational and interview studies indicate that undergraduates are less assiduous than more advanced mathematicians in attending to details, addressing gaps in prior knowledge, and identifying and resolving misunderstandings (Shepherd, Selden, & Selden, 2012; Shepherd & van de Sande, 2014).

When undergraduates are asked to evaluate individual proofs, outcomes are similar. Some are willing to make validation judgements quickly, even if they are not fully confident that they understand all aspects of an argument (Weber, 2010; Weber & Mejía-Ramos, 2014). Some will accept that a theorem is proved by an argument for its converse (Selden & Selden, 2003; Inglis & Alcock, 2012), and some will subsequently admit that they did not spot such an important fault because they did not read the theorem before attempting to understand the proof (Powers, Craviotto, & Grassl, 2010). This comparative inattention is reflected in studies of reading behaviour: Inglis and Alcock (2012) showed that undergraduates make eye movements consistent with searching for logical justifications, but do so substantially less than mathematicians. Overall, these results are consistent with students' self-affirmed beliefs about proof reading: unlike their lecturers, undergraduates often believe that a reader can expect to understand a proof within 15 minutes and without needing to construct extra justifications (Weber & Mejía-Ramos, 2014).

If some student failures in comprehension and validation are due to inattentive reading, then it is plausible that reading performance might respond to guidance that leads to increased attention (cf. Hodds et al., 2014). And there is encouraging evidence that students can read more effectively, especially after appropriate tasks or training. Some students do spontaneously engage in useful behaviours: they identify proof frameworks, break proofs into sub-proofs, illustrate difficult assertions with examples, and compare what they read with their own proof attempts (Weber, 2015). Students in interview studies improve their validation performance when given opportunities to reconsider their answers (Selden & Selden, 2003), and observations from both peer assessment and intervention studies indicate that exposure to critiquing tasks can lead students to reflect on the shortcomings of their own written mathematics (Jones & Alcock, 2013; Stylianides & Stylianides, 2009b). Similarly, reflective reports suggest that when asked to critique written arguments, undergraduates quickly adopt stringent criteria about both content and communication (Kasman, 2006). Finally, evidence from experimental and eye-movement studies indicates that mathematical reading comprehension improves in response to light-touch self-explanation training. Hodds et al. (2014) reported a sequence of studies using a short self-study booklet that instructed students in how to construct self-explanations when reading mathematical proofs. In both lab and classroom studies, this training led to better subsequent proof comprehension and more expert-like reading behaviour.

None of this means that better reading would lead to perfect comprehension. But it does give cause for optimism: perhaps undergraduates do not routinely behave as we might hope when studying proofs, but can improve considerably if taught to read more effectively. One question, then, is how might a concerned lecturer teach students to study written proofs? In this paper, we address this question by evaluating one intervention designed to support effective proof reading: the provision of an *e-Proof*.

3 Intervention: design principles and early feedback

In this section we explain the rationale behind the design of e-Proofs, relate this to the literature on mathematics education and multimedia learning, and present small-scale evidence that students view e-Proofs favourably. In the following sections, we present two studies. Study 1

focused on outcomes, evaluating performance on comprehension tests after students studied an e-Proof or a standard textbook proof. Study 2 focused on processes, investigating the effects of e-Proofs on two aspects of students’ reading behaviours: distribution of attention and inferred processing demand.

3.1 Intervention design

E-Proofs were designed to address the natural weaknesses of traditional lectures as identified in the mathematics education literature (e.g., Lew et al., 2016). Their designer—the last author of this paper—observed that proof presentations in lectures require students to draw on background knowledge, validate justifications, and recognise larger-scale proof structures, all in rapid succession; further, that lecturers’ verbal explanations are ephemeral and typically no longer available when students re-read their notes (Alcock & Wilkinson, 2011). She thus set out to construct a resource that would capture the explanations that she would ordinarily offer verbally, but allow students to engage more fully with these by making them replayable and by highlighting the parts of the proof to which they referred. The result was eight e-Proofs designed for beginning real analysis. Each showed a single theorem and proof and offered multiple explanation screens. Each screen focused attention using greying out, boxes and arrows, and was accompanied by a short audio commentary that could be played by clicking a button, and replayed as many times as desired. A sample screen (one of ten) for an e-Proof as used in studies 1 and 2 is shown in Fig. 1. A

Cauchy's Generalised MVT: Suppose that f and g are continuous on $[a, b]$ and differentiable on (a, b) and that $\forall x \in (a, b), g'(x) \neq 0$.

Then $\exists c \in (a, b)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$.

Proof

Note that if $g(a) = g(b)$ then by Rolle's Theorem $\exists c \in (a, b)$ s.t. $g'(c) = 0$.

This contradicts the theorem premise so $g(a) \neq g(b)$.

Define $h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g(x)$.

h is continuous on $[a, b]$ and differentiable on (a, b) (sum and constant multiple rules).

Also $h(a) = f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}g(a) = \frac{f(a)g(b) - g(a)f(b)}{g(b) - g(a)}$.

and $h(b) = f(b) - \frac{f(b) - f(a)}{g(b) - g(a)}g(b) = \frac{f(a)g(b) - g(a)f(b)}{g(b) - g(a)}$, so $h(a) = h(b)$.

Hence, by Rolle's Theorem, $\exists c \in (a, b)$ such that $h'(c) = 0$.

But $h'(c) = 0 \Rightarrow f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(c) = 0 \Rightarrow \frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$.

So $\exists c \in (a, b)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$.



Fig. 1 A sample e-Proof screen. The accompanying audio commentary said “In the eighth line, we translate this information about h back into information about f and g by writing h -prime of c in terms of f -prime of c and g -prime of c . Differentiating h is simpler than it looks because the fraction is just a constant. We then rearrange to obtain the equation for the conclusion of the theorem”

full e-Proof as used in study 2 appears at <https://10.6084/m9.figshare.4239554.v1>, and more detail on the design appears in by Alcock and Wilkinson (2011).

3.2 E-Proof design principles

From a mathematical perspective, e-Proofs belong to a class of interventions that provide reading support targeted to specific content. Other such interventions have been designed to support careful reading of textbook sections by combining direct instruction on what to read with questions to prompt reflection (Alcock, Brown, & Dunning, 2015; Shepherd, 2005), or to promote careful reading of single proofs by asking comprehension questions tailored to those proofs (Conradie & Frith, 2000; Cowen, 1991). In reports on both types, authors have expressed the wish to help students engage in thought processes that will help them to understand a mathematical text. E-Proofs were also designed with this aim and, like all of these interventions, they required considerable time to construct but were designed for use in independent study so that they would not require lecture reorganisation.

From the perspective of the proof comprehension literature, the audiovisual explanations offered in e-Proofs address the dimensions put forward in models of proof comprehension by authors such as Lin and Yang (2007) and Mejía-Ramos, Fuller, Weber, Rhoads and Samkoff (2012). The seven-dimensional model of Mejía-Ramos et al., for instance, involves comprehension of (1) the meanings of terms and statements, (2) the logical status of statements, (3) the warrants or justifications used to deduce claims from other information, (4) the higher-level ideas or overarching approach, (5) the modular structure (if a proof can be broken into meaningful subsections), (6) transfer of the ideas to new contexts, and (7) application (illustrating the proof's workings with specific examples). Such models have typically been used for generating proof comprehension tests, and e-Proofs embody some of their dimensions in more obvious ways than others: it is less natural to discuss transfer while the screen still shows the proof in question, for instance. Also, the e-Proof designer attempted to capture somewhat "natural" explanations, rather than (say) to ensure that every e-Proof offered equal treatment of each dimension. Nevertheless, the audiovisual explanations typically included all of the first five dimensions of this model and sometimes the seventh.¹ In theory, then, they should assist students in focusing their attention in ways that contribute to proof comprehension.

3.3 E-Proof design and multimedia learning

From the broader perspective of learning resource design, e-Proofs conform to recommendations based on Mayer's (2001) theory of multimedia learning. Mayer views learning as an active process in which learners *select* which information to attend to, *organise* this into a coherent cognitive structure, and *integrate* it with relevant knowledge from their long term memory—see Fig. 2.

Mayer has related this theory to multimedia learning in particular via two key results: that working memory is limited (Baddeley, 1992), and that visual and verbal information are entered into and processed in an individual's cognitive system via separate channels (as per

¹ The model of Mejía-Ramos et al. had not been published when e-Proofs were designed, but its authors intended it to be as complete as possible, so it is encouraging that it does capture all the types of explanation that the designer spontaneously included.

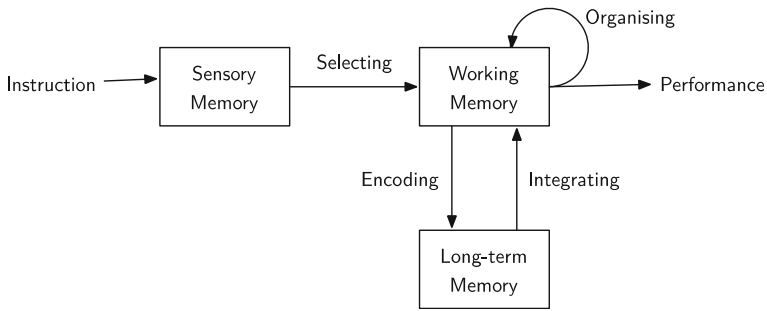


Fig. 2 Mayer's theory of learning

dual coding theory, see Clark & Paivio, 1991). On this basis, Mayer and Moreno (2003) offered recommendations for designing educational resources that render working memory load manageable by taking advantage of multimedia communication. E-Proofs are consistent with these recommendations: they move some essential processing from the visual to the auditory channel, allow time between successive bite-sized segments, provide cues to reduce processing of extraneous elements, avoid presenting identical streams of printed and spoken words, and present narration and corresponding animation simultaneously to minimise the need to hold representations in memory. In theory, then, e-Proofs should assist students in the complex task of proof comprehension by presenting information in a way that allows engagement without working memory overload.

3.4 Early feedback on e-Proofs

The eight e-Proofs initially constructed were used by the designer in a real analysis lecture course (on continuity, differentiability, and integrability) for approximately 115 second- and third-year undergraduates who were studying for UK degrees in mathematics or mathematics with another subject. There was no formal evaluation of e-Proofs at this stage; the designer simply used them in lectures, trying out different approaches such as playing all the animations in sequence or allowing students to discuss a printed copy of the proof before showing animations for the more challenging parts. The e-Proofs (along with lecture notes) were made available on the institutional virtual learning environment as the course progressed, and all eight remained available through the examination period.

From the beginning, it was clear that students liked this new type of resource. Positive comments were common on standard feedback forms, where a typical comment was, "I found hearing the lecturer explaining each line individually helpful in understanding particular parts and how they relate to the entire proof." Broader feedback-gathering exercises revealed similar attitudes. Based on responses to an earlier survey and on focus group interviews (Roy, 2014), we constructed an online survey combining items on e-Proofs with other items on study behaviour. Thirty-eight students (approximately one third of the class) took part and participants were asked to state the extent to which they agreed with eight statements including:

- I understand more by studying a proof on paper for myself than I do when using an e-Proof;
- e-Proofs helped me understand where different parts of a proof come from and how they fit together.

Participants responded on a Likert scale, reverse-scored where necessary so that the response to each item ranged from 0 (strongly negative response to e-Proofs) to 4 (strongly positive response). The mean score was 25.53 out of 32, 95% CI [24.45, 26.60]. We report this result to provide indicative evidence that those undergraduates who took part held favourable views on e-Proofs.

We were aware, however, that students are likely to view extra resources favourably simply because they show that a lecturer wants to support learning. We were also aware that authors of reports on related interventions described their experiences but did not conduct experimental evaluation studies (Alcock et al. 2015; Shepherd, 2005), so they could not provide empirical evidence about causal effects on learning outcomes. For e-Proofs, therefore, we conducted such a study, measuring learning outcomes via immediate and delayed comprehension tests.

4 Study 1: learning outcomes

4.1 Study 1 method

Study 1 used an experimental design to compare the effect of studying an e-Proof with that of studying the same proof on paper with no additional explanation (the “textbook” version). Participants were enrolled in a real analysis course; all were in either the second or third year of a single or joint honours UK mathematics degree, meaning that 50% or more of their study time was spent on mathematics. They were randomised into an e-Proof group ($N=21$) or a textbook proof group ($N=28$), where group assignment was done using a random number generator. Both groups studied a previously unseen proof of Cauchy’s Generalised Mean Value Theorem for 15 min (see Fig. 1)². The e-Proof group sat in a computer lab and were each provided with earphones to permit individual study of the e-Proof. The textbook group sat in a classroom and each received a printed copy of the proof.³ Both groups had access to pens and paper; they were neither encouraged to use these nor prevented from doing so. For the immediate test phase, both groups were provided with a short version of the proof and an additional information sheet which contained definitions of continuity and differentiability and statements of relevant theorems. They were given 30 min to complete an immediate post-test comprising eight free-response questions designed according to Lin and Yang’s (2007) proof comprehension model, which has five facets: basic knowledge, logical status, summarisation, generality, and application. The test was scored out of 18, and included questions such as follows:

- With h defined as in the proof, what is $h'(x)$?
- Where does the proof show that h satisfies the conditions for Rolle’s Theorem?
- Find an interval where $f(x) = 1 + x^2$ and $g(x) = x^2 - 1$ satisfy the premises of Cauchy’s Generalised MVT.

Two weeks later, all participants were again provided with the short version proof and the information sheet, this time in a single lecture theatre. They took the same test as a delayed test,

² The version used in the study had line numbers so that participants could refer to these.

³ Because the computer room could not accommodate half of the class, a third group listened to a live lecture. This group is not relevant here, but the full study was presented at the 13th Conference of the SIG-MAA on Research in Undergraduate Mathematics Education (Roy, Alcock, & Inglis, 2010) and we are grateful to conference participants who gave useful feedback.

this time in 20 min due to classroom time constraints. The test—with scoring information—and both versions of the proof are provided in the [Supplementary Materials](#).

4.2 Study 1 results

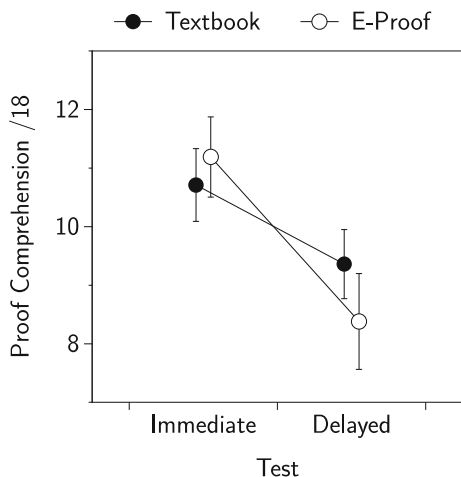
The mean scores of the two groups at immediate and delayed test are shown in Fig. 3.

An analysis of variance (ANOVA) with one within-subjects factor (time of test: immediate, delayed) and one between-subjects factor (group: e-Proof, textbook) revealed a significant main effect of time, $F(1,47) = 28.213, p < 0.001$. Both groups performed significantly better in the immediate test than in the delayed test, which is accounted for by the 2-week delay and the shorter time available to answer the same questions. The main effect of group was not significant, $F(1,47) = 0.006, p = 0.938$, but there was a significant time \times group interaction, $F(1,47) = 5.659, p = 0.021$. The drop-off in the score between the immediate and delayed tests was greater for the e-Proof group than for the textbook group (2.8 versus 1.3 points out of 18). This cannot be accounted for by the delay or the intervening teaching because both groups had the same delay and the same intervening teaching. Thus, it indicates that the e-Proof group exhibited significantly poorer retention.

4.3 Study 1 discussion

This outcome was a reminder that well-intentioned interventions might not promote good retention of studied material, even when students report that these interventions are helpful. They were also a salutary reminder that evaluation is tricky. Had we stopped at an immediate test, we would have concluded that e-Proofs did no harm and that, as students liked to have them, providing them was a good idea. It took a delayed test to reveal that e-Proofs did not affect immediate test performance, but did apparently prompt students to learn in a way that led to poorer retention. These points are not new; they are (for instance) consistent with recent rigorous studies showing that student evaluations are not consistent with teaching effectiveness as measured by performance in subsequent courses (Braga, Paccagnella, & Pellizzari, 2014; Carrell & West, 2010). But we believe they bear repeating here because some reports of interventions in mathematics education are framed in a way that focuses on instructor or

Fig. 3 Mean proof comprehension test scores for e-Proof and textbook groups at immediate and delayed test. Error bars show ± 1 SE of the mean



student experiences (Trenholm, Alcock, & Robinson, 2012), and because student experience surveys are becoming increasingly influential in the UK context where this study took place (Cheng & Marsh, 2010).

This result does not, however, imply that e-Proofs are never useful. Perhaps this particular e-Proof was weak on its own terms. Mathematicians do not always agree about what makes a proof pedagogically sound (Lai, Weber, & Mejia-Ramos, 2012), and similar resources designed by a different lecturer might be more effective. Or perhaps e-Proofs are not ideal for first encounters with proofs, but would help students struggling with previously studied proofs or using them as revision tools. Both of these are real possibilities that would merit further research. We considered them seriously, but remained troubled because we did not understand the cognitive mechanisms by which this e-Proof affected retention. Without this understanding, we believed that attempts to design better e-Proofs or to help students use them more effectively would be based on intuitive guesswork rather than evidence. We therefore decided to step back and study in detail the processes involved in studying this type of resource. Our purpose in study 2 was to use an experimental design and eye-movement data to investigate the effects of e-Proofs on reading processes.

5 Study 2: reading processes

5.1 Study 2 methodology

Eye-movement studies have been used extensively to investigate reading behaviours and thereby to infer information about the cognitive processes underlying reading in English and a variety of other languages (see Rayner, 2009, for an extensive review). These studies use the fact that when reading fixed materials, eye movements are not smooth but instead consist of short stops called *fixations* and rapid movements called *saccades*. In ordinary silent reading in English, fixations typically last around 225–250 ms (Rayner, 2009), but individual fixation durations vary considerably. So eye-movement studies commonly use *mean fixation durations* as a measure of processing demand, where mean fixation durations are higher when people engage in more demanding tasks. Evidence for this comes from educational and psychological studies that vary the complexity of stimuli or other task demands (Amadiou, Van Gog, Paas, Tricot, & Mariné 2009; Rayner, 1998; Gould, 1973; Van Gog, Paas, & Van Merriënboer, 2005). For instance, mean fixation durations increase as text becomes more conceptually difficult (Jacobson & Dodwell, 1979), and when fixating on words that are less frequent in the language (Inhoff & Rayner, 1986). Note that mean fixation durations are used to infer moment-to-moment processing demand (Rayner, 1998) rather than conscious intention or experience. Higher mean fixation durations should not be interpreted to mean that individual readers are more motivated to read the material, have decided to exert more effort, or even are aware that they are doing so. They simply provide a behavioural measure indicating that a task is more or less demanding.

Fixation locations are used in the obvious way to measure locus of attention, where appropriate analysis software (e.g., Tobii Technology, 2010) allows the user to define *areas of interest* within the material to be read. Areas of interest are typically used in two ways: researchers study attention distribution by summing the durations of all fixations in each area to calculate their associated *dwelling times*; they also study reading behaviours by comparing patterns of saccades between areas. In recent research in mathematics education, studies using

dwell times have indicated that mathematicians, more than undergraduates, focus attention on words in proofs (Inglis & Alcock, 2012). Saccade patterns have been used to document strategies for comparing fraction magnitudes, by examining when people shift attention from numerator to denominator for a single fraction and when they shift attention from the numerator of one fraction to the numerator of another (Obersteiner & Tumpek, 2016). Saccade patterns have also indicated that mathematicians, more than undergraduates, shift their attention between lines of a proof when reading it for validation (Inglis & Alcock, 2012), and that self-explanation training promotes more expert-like mathematical proof reading behaviour (Hodds et al., 2014).

5.2 Study 2 method

In the study presented here, we used a within-subjects design to compare students' reading behaviours when studying an e-Proof and a textbook proof. We used four proofs: two for which there were pre-existing e-Proofs from the real analysis course, and two in elementary number theory for which similar e-Proofs were created. Both e-Proof and textbook versions were formatted for use on an eye-tracker screen—the only difference between the two versions of each was the extra animations and audio provided for the e-Proofs. For each of the four proofs, we formulated a short multiple-choice comprehension test, designed this time according to the framework set out by Mejía-Ramos et al. (2012). We switched to short multiple-choice tests because our main interest in this study was in reading processes rather than learning outcomes, but we still wanted to ensure that the participants made appropriate effort to read for comprehension. We switched to the alternative model when it became available because it is more comprehensive than the geometry-focused model of Lin and Yang (2007). The proofs and tests appear in the [Supplementary Materials](#).

Thirty-four students took part in exchange for an £8 inconvenience allowance; each completed the study individually in an eye-movement laboratory. Each participant saw two proofs in textbook format then two in e-Proof format, where the textbook proofs appeared first so that we could assess students' reading behaviour before this was influenced by the implicit guidance provided by e-Proofs. To ensure a fair comparison of e-Proof- and textbook-prompted reading behaviour, participants were randomly assigned to one of two groups: for each proof, half of the participants saw an e-Proof and half saw the textbook format. Before each proof, participants saw an instruction page stating what theorem and what type of proof (e-Proof or textbook) they would be asked to read and indicating what key to use to move to the next page. Participants were encouraged to read the proofs for as long as they wished in order to maximise their comprehension. They were also told that when they had finished reading, they would be asked to rate their confidence in their own understanding and to answer a multiple-choice comprehension test.

Eye movements were recorded using a Tobii T120 remote eye-tracker (Tobii Technology, 2010) set to sample at 60 Hz and calibrated at the start of the study for each participant. We treated each line of each proof as an area of interest, and we report data using both mean fixation durations and total dwell times. Because of inconsistent recording (which typically occurs if a participant makes unusually exaggerated head movements or is wearing heavy eye makeup), data were excluded for one participant for Theorem 1, two for Theorem 3, and one for Theorem 4. All analyses were performed after excluding outlier fixations lasting longer than 3 standard deviations above the mean (i.e., fixations longer than 1031 ms).

5.3 Study 2 results

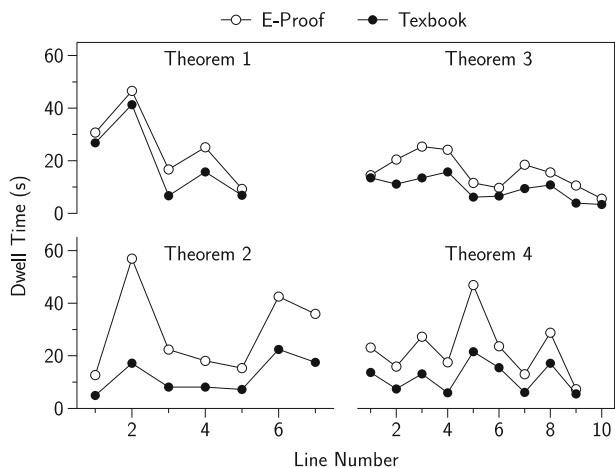
Although comprehension test scores were not the primary focus of study 2, between-groups differences were investigated using Kolmogorov-Smirnov Z tests (because scores were out of 3 or 5, parametric tests were not appropriate). No significant between-group differences were observed for any proof, all $ps > 0.1$. So, as in study 1, there was no evidence of comprehension differences at immediate test. We report the reading process data in three stages, focusing first on attention distribution, then on global processing demand, then on processing demand in detail.

5.3.1 Study 2 results: attention distribution

Did the participants distribute their attention differently when reading an e-Proof and when reading a standard proof? We anticipated that they would: the e-Proofs' creator believed that students did not always notice key ideas in a proof, and intended to point these out using the annotations and commentary. The global answer, however, was no. Figure 4 shows the mean total dwell times for each line of each proof, comparing these results for participants who read the e-Proof and those who read the standard proof. As is clear from these graphs, different lines of the proofs attracted different proportions of participant attention, but participants in both groups spent longer and shorter times on the same lines.

Statistical analyses confirmed these observations. For Theorem 1, for instance, the main effect of condition was borderline significant $F(1,31) = 4.126, p = 0.051, \eta_p^2 = 0.117$, reflecting the fact that reading the e-Proof took longer on average than reading the standard proof. The distribution of attention varied by line: for Theorem 1, for instance, a 5×2 ANOVA with one within-subjects factor (line number: 1, 2, 3, 4, 5) and one between-subjects factor (condition: e-Proof, textbook proof) showed a significant main effect of line number, $F(4,124) = 55.053, p < 0.001, \eta_p^2 = 0.631$. For example, participants took longer to comprehend Line 2 compared to the other four lines. The line number \times condition interaction was not significant, $F < 1$, so attention distributions were not significantly different. Results for the other three proofs were similar in terms of main effects but different in that the line number \times condition interaction

Fig. 4 Mean total dwell times for each line of each proof, separated by proof type



effects were also significant. Inspection of the graphs indicates that the e-Proofs did not substantially alter overall patterns of attention, but did tend to amplify differences in dwell times; it seems that they encouraged even longer dwell times on those lines that naturally attracted more participant attention in the textbook version (this was particularly the case for line 2 of Theorem 2 and line 5 of Theorem 4).

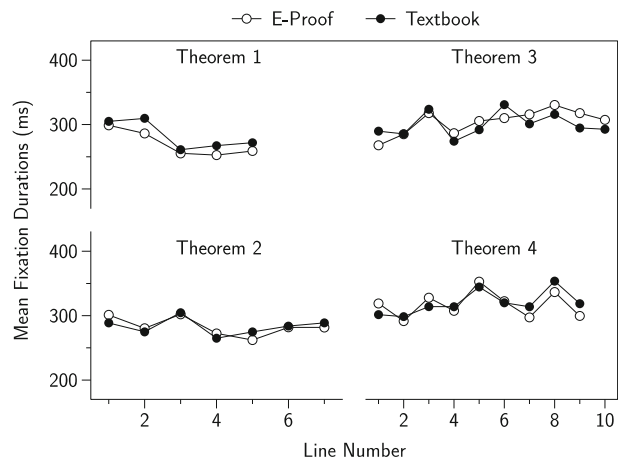
5.3.2 Study 2 results: processing demand

Next, we investigated whether e-Proofs affected processing demand. Again there were reasons to believe that they would: observational studies indicate that students typically read with less care than experts (Weber & Mejía-Ramos, 2014), and e-Proofs are designed to promote more expert-like reading, which might lead to longer mean fixation durations. On the other hand, e-Proofs are designed to make proof comprehension easier, so they might make study of proofs less demanding. In fact, they turned out to make no significant difference, as represented graphically in Fig. 5. For Theorem 1, for instance, a 5×2 ANOVA with one between-subjects factor (condition: e-Proof, textbook) and one within-subjects factor (line number: 1, 2, 3, 4, 5) showed a significant main effect of line number, $F(4, 124) = 20.999, p < 0.001, \eta_p^2 = 0.404$; this indicates differences in the cognitive demand associated with each line. But there was no significant main effect of condition, $F(1, 31) = 1.115, p = 0.299, \eta_p^2 = 0.035$, and no significant line number \times condition interaction, $F < 1$. Results were similar for the other proofs, indicating that overall, there was no systematic association of proof format with processing demand.

5.3.3 Study 2 results: processing demand with audio on and off

There was, however, a difference at a more detailed level, and we believe this is key to understanding the unfortunate retention effect found for e-Proofs in study 1. Because e-Proofs offer audio explanations but do not force the listener to use these, we were able to separate fixation durations for e-Proofs into two categories: those during which the audio was playing and those during which it was not (eight participants were excluded from this analysis because they did not use the e-Proofs' audio features for any e-Proof, relying solely on the text and animations). An ANOVA with one within-subjects factor (condition: e-Proof audio-on, e-Proof

Fig. 5 Means of the participants' mean fixation durations for each line of each proof, separated by proof type



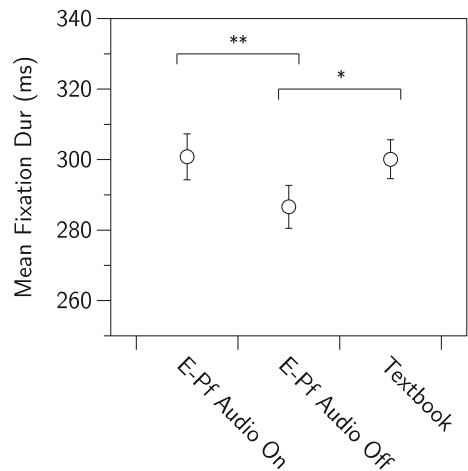
audio-off, textbook) showed a significant difference in mean fixation durations across the three conditions $F(2,50) = 3.651$, $p = 0.033$, $\eta_p^2 = 0.127$. Means were 300.1, 286.6, and 300.8 ms, respectively and, as shown in Fig. 6, Bonferroni-corrected paired t tests ($\alpha = 0.17$) showed that mean fixation durations were significantly higher for both textbook proofs and e-Proofs with audio on than they were for e-Proofs with audio off, $t(33) = 2.666$, $p = 0.012$, $d = 0.457$ and $t(25) = 3.143$, $p = 0.004$, $d = 0.616$, respectively. This indicates that, when viewing the written texts of e-Proofs without listening to the audio, the processing demand reflected in the undergraduates' mean fixation durations was significantly less than that associated with reading a text with no audio available. One possible interpretation is that e-Proofs undermined the participants' self-reliance—their reading reflected processing as demanding as that in normal reading when listening to the audio explanations, but less demanding when viewing the text independently.

5.3.4 Methodological comments on both studies

We conclude this section with two methodological points.

First, as noted in the Study 1 discussion, an e-Proof offers a specific set of audiovisual explanations, and it could be that an alternative set would result in different outcomes. However, study 2 used a counterbalanced experimental design involving four different proofs from two subject areas, so we consider it unlikely that alternative audiovisual explanations would evince dramatically different reading processes—if e-Proof effectiveness varied greatly according to the specific explanations offered, we would expect to see greater variety in the reading behaviours than was evident here. Of course, there remain other possible relationships between our findings and learning in context. Perhaps e-Proofs are not effective for retention after initial study of a proof, but are very helpful for students who have not independently reached a satisfactory understanding or who choose to use them while revising for an examination. Perhaps e-Proofs would be effective if they were modified so that they required more active engagement, requiring the student to respond in some way to explanations or other prompts. These suggestions are open to confirmation or refutation in future empirical studies, where such studies might also take into account contextual factors such as prior knowledge,

Fig. 6 Means of participants' mean fixation durations for e-Proofs with audio on, e-Proofs with audio off, and textbook proofs ($*p < 0.05$, $**p < 0.01$). Error bars show ± 1 SE of the mean



study goals, availability of one-to-one help, time spent in independent study, individual differences in take-up of e-Proofs as one among a set of learning resources, and so on.

Our second methodological point is that studies 1 and 2 both took place in “artificial” situations—study 1 involved individual learning in classroom settings different from the normal lecture theatre (a computer lab in the case of the e-Proof group), and study 2 involved individual reading on a screen in a lab. We do not claim that reading on a screen is the same as reading on paper, though it is now common—students regularly access mathematical content online or via electronic lecture notes and textbooks. And we do not claim that students would behave identically in typical lecture or group learning situations. We do claim—as is routine in experimental studies—that removing complex contextual factors allows us to infer causal connections between controlled variables and resultant comprehension outcomes and reading processes. In study 1, the between-groups difference provides evidence that study of an e-Proof resulted in poorer knowledge retention than study of a textbook proof. In study 2, mathematical reading processes were revealed via a behavioural measure that captures reading processes more directly than test performance; this provided evidence that the availability of audio explanations led students to exert less processing effort when these were turned off.

5.4 Study 2 discussion

The results of study 2 provide a process-based account for the comparatively negative effect of e-Proofs on retention that we found in study 1. Recall that e-Proofs were expected to support proof comprehension by using a design consistent with Mayer and Moreno’s (2003) recommendations for multimedia learning resources. They did not support comprehension in a straightforward way, but we suggest that our results as a whole can nevertheless be understood in terms of Mayer’s (2001) framework. When engaging with multimedia resources, learners need to select new information from audio and visual stimuli, organise that information to construct a coherent representation, and integrate it with their existing knowledge to create new knowledge (Fig. 2). When undergraduates study an e-Proof, they engage with multiple tasks: reading the written text, listening to the audio, following the highlighted arrows and boxes, and clicking the mouse to move between slides. The form of the explanations offered means that there might be interference between the audio explanations and the learner’s interaction with the text, so perhaps there should be caveats to the dual-coding-based recommendation that audio and visual information should differ. And the fact that explanations are available means that learners are not fully in control of selecting what they will attend to and in what order; removing choice about what to attend to could mean that attention is not well directed toward ideas that learners can link to existing knowledge. Moreover, the learning process is subject to numerous short interruptions. Because processing occurs in limited-capacity working memory, such interruptions could interfere with, rather than help, organisation. Thus the new information might remain comparatively poorly structured and not well integrated into long-term memory. In such circumstances, learners might struggle to reuse that knowledge, especially after a time gap.

This account is consistent with the retention effect of studying e-Proofs as found in study 1: if e-Proof readers succeed in following and understanding the provided explanations, they might attain good short-term understanding of those explanations but only weaker integration with their own existing knowledge. It is also consistent with the process effect of reading e-Proofs as found in study 2: if e-Proof readers concentrate primarily on following and understanding the provided explanations, one would expect them to exhibit longer fixation durations when audio is playing than when it is not.

Moreover, this account is consistent with the fact that students believed e-Proofs to be helpful. We do not suggest that they wilfully misreported their learning experiences, or even that they were misguided about how much they had learned. We suggest, instead, that they were reporting accurately on their experience of learning *in the short term*. Our results suggest that using e-Proofs might enable students to attain understanding with somewhat less effort than they would have to exert without the provided support. But learning with comparative ease in the short term is not necessarily the same as learning in a way that will promote deep understanding and long-term retention (Bjork, Dunlosky, & Kornell, 2013).

In general, it is not obvious how to balance these issues: there can be a fine line between productive struggle and distressing failure. But these results suggest to us that e-Proofs provide help that is too much or too directive, and that other approaches seem more promising for instructors wishing to support undergraduate proof comprehension. Specifically, as we observed in the Intervention section, e-Proofs fit into a class of interventions designed to support careful reading of mathematical texts by combining direct instruction on what to read with text-specific prompts. In this, they differ from approaches involving self-explanation training, which have the same overall goal but are not text-specific; instead they provide students with generic guidance on questioning their understanding of each part of a text and relating both the questions and self-generated answers to their existing knowledge (e.g., Hodds et al. 2014). When we began research in proof comprehension, we did not know which approach would be more effective. The evidence to date indicates that undergraduate mathematics students, at least those with experience similar to our participants, would likely be better served in the first instance by generic resources that teach them to self-explain.

6 Conclusion

The findings presented here pertain to a question raised by Weber and Mejía-Ramos (2014, p. 91): “If a student fails to understand a proof, is this because the quality of the presentation was inadequate or because the student did not exert sufficient effort to understand it?” In constructing e-Proofs, a lecturer does considerable presentational work to make individual proofs accessible. This sounds like it should help, but in our study it resulted in poorer retention, which should concern any lecturer or researcher designing an intervention to support students’ understanding of proofs. We conclude by offering comments on the relevance of this finding to broader issues in the learning of proof.

Teaching interventions happen all the time. Teachers and lecturers recognise that proof presents a challenge to their students, and they innovate on a daily and yearly basis in order to give better explanations or tasks related to specific proofs, and in order to help students experience the process of proving and learn about the meaning proof in the mathematical community. In the contemporary educational environment, they can also facilitate access to electronic resources and can tailor-make their own. This fosters creativity, which in itself generates positive feeling: there is no doubt that committed teachers and lecturers make sincere efforts to support their students. But it also means that innovation tends to run ahead of evaluation. One lesson confirmed by the studies reported in this paper is that it is risky to evaluate learning innovations by student feedback alone. No doubt this data has value—if a resource is not perceived as useful, there is not much hope of its being adopted by today’s discerning internet users. But we should be careful not to conflate opinions about the learning experience with actual learning outcomes (cf. Bjork et al. 2013).

Our work thus raises questions about the responsibility of teachers, lecturers, and particularly researchers in evaluating interventions, deciding to make resources available, and communicating with students about how best to use these. For mathematics students attempting to get to grips with proof, the contemporary environment creates a huge meta-problem: from the plethora of freely-available resources—lecturers’ webpages, open-source textbooks, online videos—how can they identify those that are useful, and how should they divide their time between them? This problem is well recognised in discussions about information literacy in both students (Kuiper, Volman, & Terwel, 2005) and teachers (Kovalik, Jensen, Schloman, & Tipton, 2010). But we suspect that it is not often discussed in mathematics degree programmes. In our experience, mathematicians sometimes provide links to online resources, and students certainly seek out video explanations for proofs they do not understand. But there is little information on how students might sensibly organise their independent study time. Even resources designed to address this gap (e.g., Alcock, 2013) do not usually discuss the likely merits of watching more lectures or online videos versus studying proofs in lecture notes or tackling proof construction problems. A student who is motivated—quite reasonably—by a sense of progress might well opt for one of the former, when in fact struggling independently to construct or understand difficult proofs might be a better use of time (cf. Bjork et al. 2013). Contemporary students must learn how to navigate the web in order to find high-quality learning resources; they might also need to learn when not to use them.

Acknowledgements This work was partially funded by a grant from the Mathematics, Statistics, and Operational Research Network of the Higher Education Academy (to LA and MI) and a Royal Society Worshipful Company of Actuaries Research Fellowship (to MI).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alcock, L. (2013). *How to study for a mathematics degree*. Oxford: Oxford University Press.
- Alcock, L., Brown, G., & Dunning, T. C. (2015). Independent study workbooks for proofs in group theory. *International Journal of Research in Undergraduate Mathematics Education*, 1, 3–26.
- Alcock, L., & Simpson, A. (2001). The warwick analysis project: Practice and theory. In D. Holton (Ed.), *The teaching and learning of mathematics at the undergraduate level* (pp. 99–112). Dordrecht: Kluwer.
- Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *Journal of Mathematical Behavior*, 24, 125–134.
- Alcock, L. & Wilkinson, N. (2011). e-Proofs: Design of a resource to support proof comprehension in mathematics. *Educational Designer*, 1(4). Retrieved from <http://www.educationaldesigner.org/ed/volume1/issue4/article14/index.htm>
- Amadiou, F., Van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction*, 19, 376–386.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64, 417–444.
- Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students’ evaluations of professors. *Economics of Education Review*, 41, 71–88.
- Burn, R. P. (1992). *Numbers and functions: Steps into analysis*. Cambridge: Cambridge University Press.
- Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy*, 118, 409–432.

- Cheng, J. H., & Marsh, H. W. (2010). National student survey: Are differences between universities and courses reliable and meaningful? *Oxford Review of Education*, 36, 693–712.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3, 149–210.
- Conradie, J., & Frith, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics*, 42, 225–235.
- Coppin, C. A., Mahavier, W. T., May, E. L., & Parker, E. (2009). *The Moore method: A pathway to learner-centered instruction*. Washington, DC: MAA.
- Cowen, C. (1991). Teaching and testing mathematics reading. *American Mathematical Monthly*, 98, 50–53.
- Gabel, M., & Dreyfus, T. (2013). The flow of a proof—the example of the Euclidean algorithm. In A. Lindmeier, A. & A.M. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group of the Psychology of Mathematics Education* (Vol. 2, pp. 321–328). Kiel, Germany: IGPME.
- Gould, J. D. (1973). Eye movements during visual search and memory search. *Journal of Experimental Psychology*, 98, 184–195.
- Greiffenhagen, C. (2008). Video analysis of mathematical practice? Different attempts to “open up” mathematics for sociological investigation. *Forum: Qualitative Social Research*, 9(3), Art. 32.
- Hodds, M., Alcock, L., & Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, 45, 62–101.
- Imamoğlu, Y., & Toğrol, A. Y. (2015). Proof construction and evaluation practices of prospective mathematics educators. *European Journal of Science and Mathematics Education*, 3, 130–144.
- Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, 43, 358–390.
- Inglis, M., Mejía-Ramos, J.-P., Weber, K., & Alcock, L. (2013). On mathematicians’ different standards when evaluating elementary proofs. *Topics in Cognitive Science*, 5, 270–282.
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431–439.
- Jacobson, J. Z., & Dodwell, P. C. (1979). Saccadic eye movements during reading. *Brain and Language*, 8, 303–314.
- Jones, I., & Alcock, L. (2013). Peer assessment without assessment criteria. *Studies in Higher Education*, 39, 1774–1787.
- Kasman, R. (2006). Critique that! Analytical writing assignments in advanced mathematics courses. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 16, 1–15.
- Kovalik, C., Jensen, M. L., Schloman, B., & Tipton, M. (2011). Information literacy, collaboration, and teacher education. *Communications in Information Literacy*, 4, 145–169.
- Kuiper, E., Volman, M., & Terwel, J. (2005). The web as an information resource in K–12 education: Strategies for supporting students in searching and processing information. *Review of Educational Research*, 75, 285–328.
- Lai, Y., Weber, K., & Mejía-Ramos, J.-P. (2012). Mathematicians’ perspectives on features of a good pedagogical proof. *Cognition and Instruction*, 30, 146–169.
- Larsen, S. P. (2013). A local instructional theory for the guided reinvention of the group and isomorphism concepts. *The Journal of Mathematical Behavior*, 32, 712–725.
- Larsen, S., & Zandieh, M. (2008). Proofs and refutations in the undergraduate mathematics classroom. *Educational Studies in Mathematics*, 67, 185–198.
- Laursen, S. L., Hassi, M.-L., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education*, 45, 406–418.
- Lew, K., Fukawa-Connelly, T. P., Mejía-Ramos, J. P., & Weber, K. (2016). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *Journal for Research in Mathematics Education*, 47, 162–198.
- Lin, F.-L., & Yang, K.-L. (2007). The reading comprehension of geometric proofs: The contribution of knowledge and reasoning. *International Journal of Science and Mathematics Education*, 5, 729–754.
- Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43–52.
- Mejía-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79, 3–18.
- Mills, M. (2014). A framework for example usage in proof presentations. *Journal of Mathematical Behavior*, 33, 106–118.
- Obersteiner, A., & Tumpek, C. (2016). Measuring fraction comparison strategies with eye-tracking. *ZDM Mathematics Education*, 48, 255–266.
- Powers, R. A., Cravioito, C., & Grassl, R. M. (2010). Impact of proof validation on proof writing in abstract algebra. *International Journal of Mathematical Education in Science and Technology*, 41, 501–514.

- Pritchard, D. (2010). Where learning starts? A framework for thinking about lectures in university mathematics. *International Journal of Mathematical Education in Science and Technology*, 41, 609–623.
- Randahl, M. (2012). First-year engineering students' use of their mathematics textbook—opportunities and constraints. *Mathematics Education Research Journal*, 24, 239–256.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Roy, S. (2014). *Evaluating novel pedagogy in higher education: A case study of e-Proofs* (Unpublished doctoral dissertation). Loughborough University, Loughborough, UK.
- Roy, S., Alcock, L., & Inglis, M. (2010). Undergraduates' proof comprehension: A comparative study of three forms of proof presentation. In *Proceedings of the 13th Conference on Research in Undergraduate Mathematics Education*. Raleigh, NC, USA. Retrieved from <http://sigmaa.maa.org/rume/crume2010/Archive/Roy%20et%20al.pdf>
- Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior*, 18, 191–210.
- Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34, 4–36.
- Shepherd, M. D. (2005). Encouraging students to read mathematics. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 15, 124–144.
- Shepherd, M. D., Selden, A., & Selden, J. (2012). University students' reading of their first-year mathematics textbooks. *Mathematical Thinking and Learning*, 14, 226–256.
- Shepherd, M. D., & van de Sande, C. C. (2014). Reading mathematics for understanding—from novice to expert. *Journal of Mathematical Behavior*, 35, 74–86.
- Stylianides, G. J., & Stylianides, A. J. (2009a). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, 40, 314–352.
- Stylianides, A. J., & Stylianides, G. J. (2009b). Proof constructions and evaluations. *Educational Studies in Mathematics*, 72, 237–253.
- Technology, T. (2010). *Tobii eye tracking: An introduction to eye tracking and Tobii eye trackers*. Stockholm: Tobii Technology AP.
- Trenholm, S., Alcock, L., & Robinson, C. (2012). Mathematics lecturing in the digital age. *International Journal of Mathematical Education in Science and Technology*, 43, 703–716.
- Van Gog, T., Paas, F., & Van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, 19, 205–221.
- Weber, K. (2004). Traditional instruction in advanced mathematics courses: A case study of one professor's lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior*, 23, 115–133.
- Weber, K. (2010). Mathematics majors' perceptions of conviction, validity and proof. *Mathematical Thinking and Learning*, 12, 306–336.
- Weber, K. (2015). Effective proof reading strategies for comprehending mathematical proofs. *International Journal of Research in Undergraduate Mathematics Education*, 1, 289–314.
- Weber, K., & Mejia-Ramos, J.-P. (2014). Mathematics majors' beliefs about proof reading. *International Journal of Mathematical Education in Science and Technology*, 45, 89–103.
- Weinberg, A., Wiesner, E., Benesh, B., & Boester, T. (2012). Undergraduate students' self-reported use of mathematics textbooks. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 22, 152–175.