WATCHING MATHEMATICIANS READ MATHEMATICS

Matthew Inglis and Lara Alcock
Mathematics Education Centre, Loughborough University, UK

This report contributes to the debate about whether expert mathematicians skim-read mathematical proofs before engaging in detailed line-by-line reading. It reviews the conflicting introspective and behavioural evidence, then reports a new study of expert mathematicians' eye movements as they read both entire research-level mathematics papers and individual proofs within those papers. Our analysis reveals no evidence of skimming, and we discuss the implications of this for research and pedagogy.

INTRODUCTION

Proof is central to mathematical practice, so understanding proof and proving is an important goal of most mathematical curricula (Hanna, 2007). Furthermore, at least in advanced mathematics courses, students spend considerable time learning mathematics by studying proofs (Selden & Selden, 2003). Consequently, several research groups have investigated the processes by which students engage with written proofs (Inglis & Alcock, 2012; Ko & Knuth, 2013; Mejía-Ramos & Weber, 2014).

A complementary approach is to examine expert mathematical practice, with researchers arguing that if we want students to develop expert-like behaviours, we require accurate understanding of those behaviours (RAND, 2003; Weber, 2008; Wilkerson-Jerde & Wilensky, 2011). In this report, we address an unresolved issue from studies on expert reading (Inglis & Alcock, 2012, 2013; Mejía-Ramos & Weber, 2014; Weber, 2008; Weber & Mejía-Ramos, 2011, 2013): that of whether mathematicians skim-read mathematical texts before carefully reading line by line.

The skimming hypothesis was generated when Weber (2008) interviewed eight mathematicians about their behaviour while validating research-level proofs. Many explained that they would often skim-read before reading line by line. For example, one described "first try[ing] to understand the structure of the proof, to get an overview of the argument that's being used" (p.441); another described first reading through the proof "to get the flow of it" and then going back to "get the details" (p.441).

Inglis and Alcock (2012) investigated this hypothesis by asking mathematicians and undergraduates to validate purported proofs and recording their eye movements as they did so. They found no evidence of initial skimming—participants typically did not fixate on the last lines of purported proofs until approximately half way through their reading attempts. Citing earlier methodological work (e.g., Nisbett & Wilson, 1977), Inglis and Alcock therefore suggested that introspective evidence about mathematical practice should be regarded with caution. Weber and Mejía-Ramos (2013), however,

2018. In NNN (Eds.). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 1, pp. XX-YY. Umeå, Sweden: PME.

criticised this argument, in part because the proofs Inglis and Alcock used were too short to give meaningful results about expert practice. Inglis and Alcock (2013) concurred that their purported proofs were considerably shorter than those encountered in mathematical research (largely because their expert/novice research design required proofs that were accessible to first-year undergraduates).

Certainly mathematicians believe that they skim-read: Mejía-Ramos and Weber (2014) reported that 92% of mathematicians responding to a large-scale survey agreed with the statement "When I read a proof in a respected journal, it is not uncommon that I skim the proof first to comprehend the main ideas of the proof, prior to reading the proof line-by-line". They also asked participants about their reading behaviour when refereeing; again, large majorities of participants claimed to skim-read and check for validity in this context. They therefore suggested that it would be strange if Alcock and Inglis's (2012) failure to find such behaviour reflected actual mathematical practice. But whether mathematicians actually skim-read remains an open question and, in this report, we investigate whether skimming is evident in mathematicians' eye movements when they read research-level mathematics.

METHODS

Participants, apparatus and procedure.

To determine whether mathematicians skim-read before reading line by line, we recorded mathematicians' eye movements while they read research papers drawn from their own fields. Participants were ten permanent members of staff (assistant professor level or above) from a UK University. All had doctorates and numerous published academic papers. Five were applied mathematicians, four were pure mathematicians, and one was a statistician. Eight different nationalities were represented.

Each participant was asked to select a research paper that they planned to read but had not yet begun; these papers were forwarded to the researchers prior to the experimental session. To protect the anonymity of participants, we do not report which papers were chosen. However, they included published journal articles, pre-prints from the arXiv, and a short monograph. Topics included Bessel functions, algebraic geometry, group theory, and the modelling of physical and biological phenomena. The papers varied in length: the shortest was 4 pages and the longest 53.

Each participant took part individually in a quiet room. Eye movements were recorded with a Tobii T120 Eye-Tracker, set to sample at 60Hz. The T120 is a remote eye-tracker with two binocular infrared cameras under a 17" TFT monitor; it typically achieves eye-position tracking accuracy of 0.5°. Stimuli were displayed on a screen that participants viewed (without head restriction) from a distance of approximately 60cm. For each participant, the eye-tracker was calibrated with a 9-point display.

Participants were told that they would be shown their paper and that they should read it as if intending to write a review for MathSciNet, an online database of short reviews of published mathematical papers. All participants were familiar with the guidelines for MathSciNet (http://www.ams.org/mresubs/guide-reviewers.html), which state:

> In most cases the review should state the main results, together with enough notation to make the statements comprehensible to someone already familiar with the field. The main ideas of the proof should be sketched when this is feasible.

This instruction was designed to ensure that all participants would read for comprehension rather than some other purpose (such as checking validity). We believed that if skim reading were a common feature of mathematicians' reading behaviour, then these instructions would be likely to reveal it.

After the instructions were displayed and explained verbally by the experimenter, the first page of the participant's research paper was displayed and the experimenter left the room. Participants could move sequentially through the pages of their papers using cursor keys, and were provided with pen and paper to make notes if they wished. On completing the task, they stopped the recording and called the experimenter. There was no time restriction, and participants' reading times varied between 17 and 65 minutes.

Data analysis.

Our analysis uses the fact that, when viewing a static image, eye movements consist of fixations (short stationary periods, usually lasting 150-500ms) and saccades (rapid movements between fixations). During saccades, no information can be processed (e.g., Matin, 1974), so fixation locations suffice to determine the path of a participant's attention (for a substantial review of eye-movement research see Rayner, 2009). Our strategy was to create, for each participant, a scatter plot with time on the *x*-axis and paragraph in the paper on the *y*-axis. Because eye-movement data are noisy (blinks or random head movements can cause single fixations away from the location of attention (Inglis & Alcock, 2013), we then fitted curves to these plots using LOESS regression (also known as "locally weighted scatterplot smoothing"). This technique fits connected quadratics to local sections of a scatterplot (e.g., Cleveland, 1979), and permits fitting a curve to data without making *a priori* assumptions about the shape of the curve. If participants adopted initial skim strategies, we would expect their fixation plots to look like that shown in Figure 1.

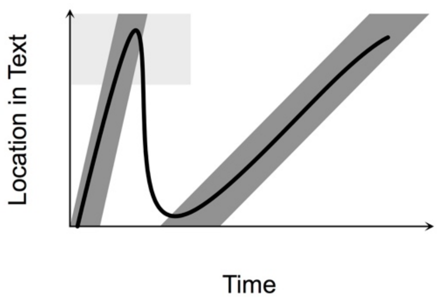

Figure 1: The type of fixation plot and LOESS curve we would expect if a participant had adopted an initial skimming strategy.

We operationalised this by evaluating whether each participant's LOESS curve entered the light grey box in the top left of Figure 1: if the focus of attention entered the last third of the reading material within the first third of their reading attempt, we coded this as a skim (cf. Weber & Mejía-Ramos, 2013).

RESULTS

We first examine global reading behaviour, reporting on each participant's reading of their entire paper. We take this approach because, in research-level mathematics, proofs cannot normally be read in isolation: papers typically introduce novel definitions, ideas and techniques before presenting a proof. We then examine local reading behaviour, illustrating participants' reading of their papers' first self-contained arguments. This allows us to compare more directly with earlier discussions of skim reading (Inglis & Alcock, 2012; Inglis & Alcock, 2013; Mejía-Ramos & Weber, 2013; Weber & Mejía-Ramos, 2013), which have typically involved single proofs.

Global reading behaviour.

Figure 2 shows individual paragraph-by-time fixation plots for all ten participants. There appeared to be three broad categories of attention movement. Some participants (1, 2, 4, 5, and 7) read in an approximately linear order, beginning at the start of the paper and progressing to the end with few moves to non-adjacent paragraphs. Others (8, 9, and 10) moved their attention in a piecewise linear fashion: they started with a linear approach, then re-read certain sections in detail, again linearly. Finally, two participants (3 and 6) appeared to adopt different approaches. In the post-experiment debrief, Mathematician 3 reported that he had not understood the introduction to his paper and had therefore failed to make substantial progress beyond the first few pages. This is consistent with his eye movements, which include a series of linear attention moves within the first 30 paragraphs. Mathematician 6 had relatively few fixations (in any location) in the latter half of his reading attempt. He made a large number of notes, so we attributed this to his eyes being largely off screen during this time.

Despite this variety in reading behaviour, no mathematician used a skimming strategy: in no case did the LOESS curve enter the last third of the paper in the first third of the reading time. Some graphs (1, 4, 6, 7 and 8) *did* show a small number of single fixations in the key area, but these were so few that we attributed them to participants scrolling forward to the reference sections of their papers (they had to view each page in turn, explaining the "trails" of fixations leading up to the reference sections in plots 4, 6 and 8). Even for participants who read in a piecewise linear fashion, reading behaviour can be distinguished from the skimming strategy detailed by Weber (2007), because the second and third reading attempts did not involve the whole text and/or took place at a substantially faster rate than the initial reading attempt.

If initial skimming were a common feature of mathematicians' reading behaviour, it is extremely unlikely that we would have found no skims in our data. A skimming rate of zero out of ten is significantly lower than 50%, sign test $p = .002$, and significantly lower than the 92% figure found by Mejía-Ramos & Weber (2013), binomial test $p =$

1.03×10^{-11}. Of course, it is possible that our operationalisation of skimming was faulty, and we consider this possibility in the next section and the discussion.

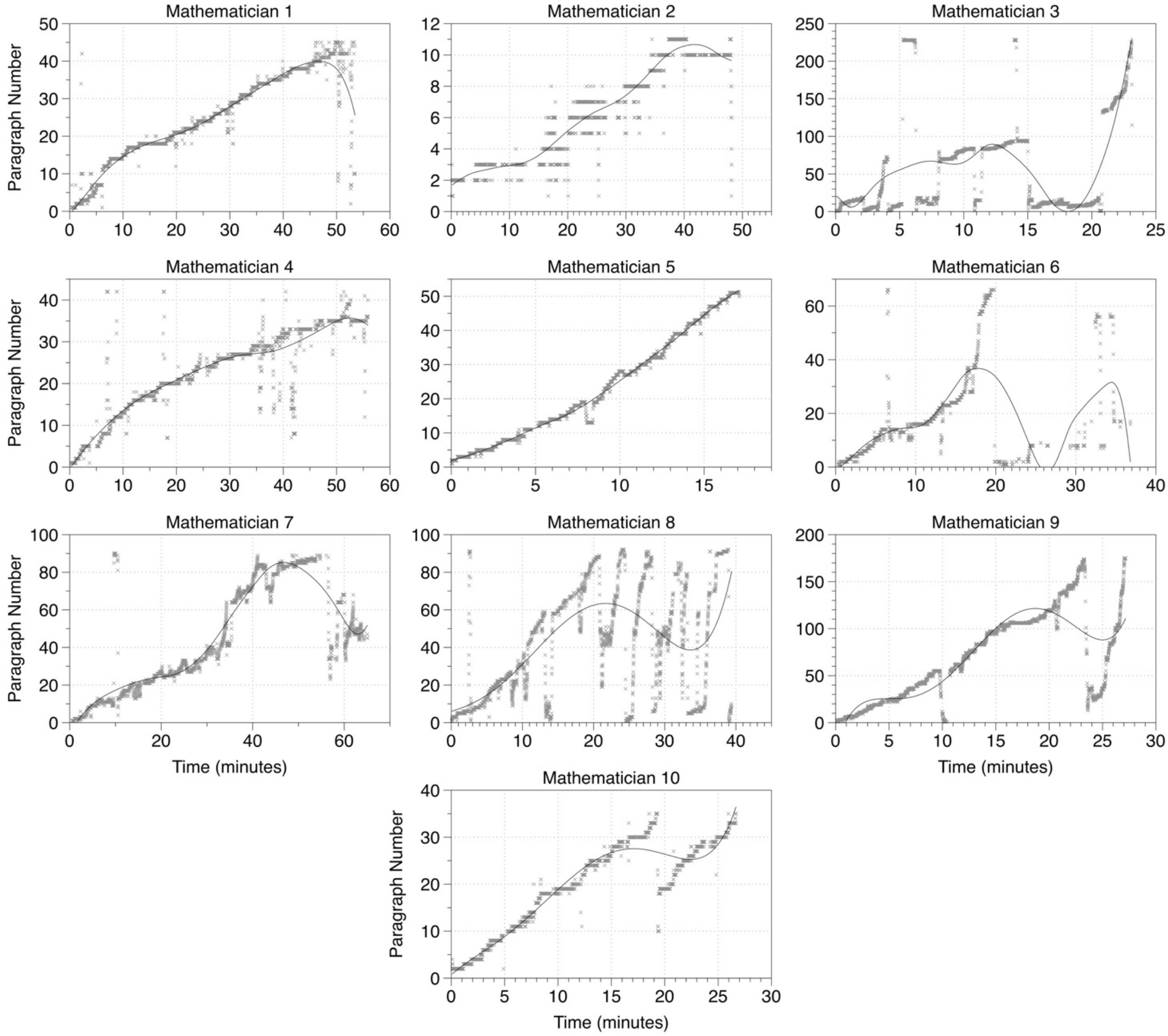

Figure 2: Paragraph Number by Time fixation plots for each participant, together with associated LOESS curves (second order, smoothing parameter 0.3).

Local reading behaviour.

We found no evidence of skimming in participants' attention while they read entire papers. But each of their papers included multiple shorter arguments, some of which formed self-contained paragraphs. Because our global analysis focused on between-paragraph eye movements, it is therefore possible that we missed the skimming behaviour hypothesised by Mejía-Ramos and Weber (2013) because this takes place within paragraphs. To investigate this possibility, we identified the first self-contained argument in each paper (typically a proof of a lemma or proposition, or the derivation of a model of a physical/biological process), and conducted a line-by-line analysis of the corresponding participant's attention for this argument.

Two illustrative fixation plots are shown in Figure 3. The wide graph shows every fixation on the relevant areas of each paper, although it is clear that many of these fixations did not contribute to genuine reading attempts (single fixations were probably due to random eye-movements or to flicking through the pages). Because of this we have magnified the sections of the plots that we judged to be the first attempt to read through the self-contained arguments, and plotted the associated LOESS curves. In our judgement, neither these participants nor any others could be said to have used a skimming approach—further graphs will be exhibited at the conference and the full set can be examined at https://doi.org/10.6084/m9.figshare.5733510.v1.

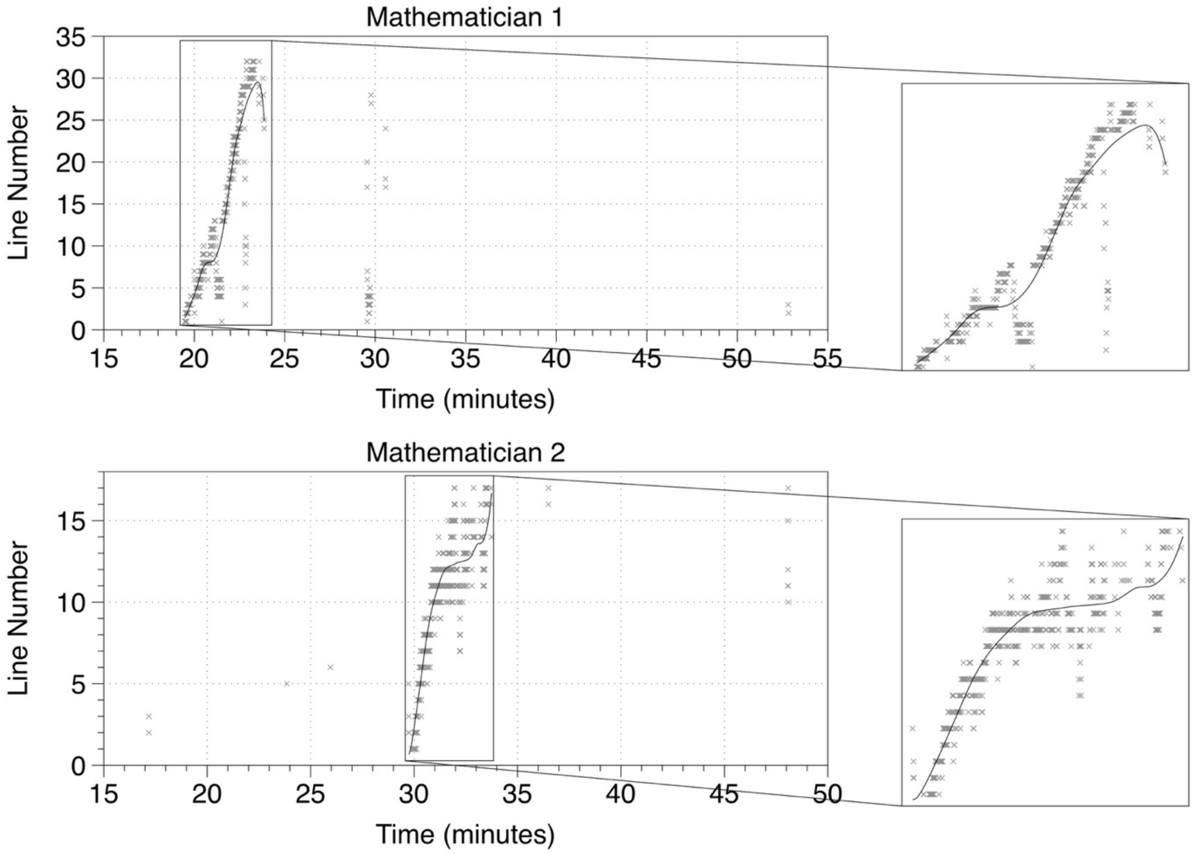

Figure 3: Line Number by Time fixation plots for the first argument in the paper for Mathematicians 1 and 2. The first clear-cut reading attempt is been magnified, together with its associated LOESS curve (second order, smoothing parameter 0.3).

DISCUSSION

Mejía-Ramos and Weber (2013) found that 92% of mathematicians claimed to understand the structures of proofs by skimming them before reading in detail. We have no reason to believe that our sample was unrepresentative of expert mathematicians—our participants worked in various areas of pure and applied mathematics and statistics, and were from eight different countries—yet we found no evidence of skimming in our data. The probability of this occurring if the introspective accounts are correct is vanishingly small, so we think it unlikely that skimming as operationalised in our study is fundamental to mathematicians' behaviour.

We briefly discuss two possible accounts for this finding, drawing out the implications of each. One account is that mathematicians simply do not skim. This would raise methodological concerns: where introspective claims are inconsistent with behavioural evidence, we must decide how to interpret the results of methodologically distinct studies. In such a situation, one might argue that introspective evidence should simply be ignored (e.g., Lyons, 1986; Nisbett & Wilson, 1977). Alternatively, however, it could be that we incorrectly operationalised what it means to skim when reading mathematics. When 92% of participants agreed that they would often "skim [a] proof to comprehend the main ideas…prior to reading [it] line-by-line"", perhaps they were referring to a much longer process than either we or Weber and Mejía-Ramos (2013) believed. Perhaps, for instance, the entire reading attempts we recorded in this experiment (which lasted up to an hour) should be classified as skim-reads. Perhaps it is only after a relatively long "skim" that mathematicians go back and re-read mathematical arguments line by line, or perhaps in normal circumstances mathematicians only skim and line-by-line reading is relatively rare. We suggest that disentangling these possibilities requires ethnographic studies of mathematical practice (cf. Greiffenhagen & Sharoock, 2011). Such studies would form a worthwhile contribution to the literature on mathematicians' reading behaviour.

In the meantime, we can comment on a broader issue. Our data revealed considerable variety in mathematicians' reading behaviours, as is apparent in Figure 2. It thus contributes to a growing body of evidence on diversity in expert mathematical behaviour (e.g., Inglis, Mejía-Ramos, Weber & Alcock, 2013; Weber, Inglis & Mejía-Ramos, 2014). We do not yet know what causes these differences. Is behaviour driven by individual differences among mathematicians? Or perhaps by the mathematical content or structures of papers or proofs? What prompts a decision to re-read a section, or to skip ahead? However, we can observe that such findings complicate arguments that we should teach students expert-like behaviours (e.g., RAND, 2003; Wilkerson-Jerde & Wilensky, 2011). If expert behaviour is heterogeneous, as suggested by this study and others, then basing instruction upon it is a non-trivial task.

References

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, *74*, 829–836.

Greiffenhagen, C. & Sharrock, W. (2011). Does mathematics look certain in the front, but fallible in the back? *Social Studies of Science*, *41*, 839–866.

Hanna, G. (2007). The ongoing value of proof. In P. Boero (Ed.), *Theorems in school: From history, epistemology and cognition to classroom practice*. Rotterdam: Sense, pp. 3–18.

Inglis, M. & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*, 358–390.

Inglis, M. & Alcock, L. (2013). Skimming: A response to Weber and Mejía-Ramos. *Journal for Research in Mathematics Education*, *44*, 471–474.

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.

Ko, Y.-Y. & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practices of importance for mathematics majors. *Journal of Mathematical Behavior*, 32, 20–35.

Lyons, W. E. (1986). *The Disappearance of Introspection*. Cambridge, MA: MIT Press.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin*, *81*, 899–917.

Mejía-Ramos, J. P. & Weber, K. (2014). Why and how mathematicians read proofs: Further evidence from a survey study. *Educational Studies in Mathematics*, *85*, 161-173.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–295.

RAND (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. Santa Monica, CA: RAND Corporation Mathematics Study Panel.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

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.

Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for Research in Mathematics Education*, *39*, 431–459.

Weber, K., Inglis, M., & Mejia-Ramos, J. P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, *49*(1), 36-58.

Weber, K. & Mejía-Ramos, J. P. (2011). Why and how mathematicians read proofs: An exploratory study. *Educational Studies in Mathematics*, *76*, 329–344.

Weber, K. & Mejía-Ramos, J. P. (2013). On mathematicians' proof skimming: A reply to Inglis and Alcock. *Journal for Research in Mathematics Education*, *44*, 464-471.

Wilkerson-Jerde, M. & Wilensky, U. (2011). How do mathematicians learn math? Resources and acts for constructing and understanding mathematics. *Educational Studies in Mathematics*, *78*, 21–43.