Diagram construction and performance in advanced mathematics

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Additional Information:

- This is a conference paper.

Metadata Record: https://dspace.lboro.ac.uk/2134/17471

Version: Accepted for publication

Publisher: International Group for the Psychology of Mathematics Education / © The Authors

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
We report a study in which 100 students at the beginning of an undergraduate real analysis course were asked to construct diagrams to represent four general mathematical statements about functions. We present four theoretical criteria for analysing such diagrams and illustrate the range of student-produced diagrams; we then present an analysis showing that performance in the diagram-construction task was significantly related to subsequent performance in the course.

INTRODUCTION

There has been considerable research on students’ use of mathematical diagrams (Presmeg, 2006). Some has sought to clarify relationships between mental imagery, external representations, and successful reasoning (Duval, 1999). Some has classified students as visualisers or otherwise (Presmeg, 1986; Stylianou & Silver, 2004), and some has investigated whether students can interpret graphical information representing real-world situations (Leinhardt, Zaslavsky & Stein, 1990; Robert & Speer, 2001). Our work asks whether students can draw suitable diagrams rather than whether they are inclined to do so. Specifically, we asked students to construct diagrams to represent abstract statements from real analysis.

Real analysis lends itself to graphical representations because it involves real-valued functions and their properties. However, while students often see diagrams, they are less often asked to construct them. There is evidence that mathematicians and successful students can draw relevant diagrams and use them to construct mathematical arguments (Gibson, 1998; Stylianou & Silver, 2004), but studies at the undergraduate level are typically small-scale and focused on spontaneously-produced diagrams. Research is largely silent on the issue of whether a typical student can produce such diagrams, and thus on whether there is a systematic relationship between this skill and mathematical performance. This report addresses this gap by reporting a study that asked students to draw diagrams for four statements:

A: $f$ is bounded on the set $X$ if and only if $\exists M > 0$ s.t. $\forall x \in X \ , \ |f(x)| \leq M$.

B: Suppose that $f : [a,b] \rightarrow \mathbb{R}$ is continuous and that $y_0$ is between $f(a)$ and $f(b)$.

Then $\exists x_0 \in (a,b)$ s.t. $f(x_0) = y_0$.

C: If $f$ is continuous at $x = a$ and $f(a) > 0$ then $\exists \delta > 0$ s.t. $|x - a| < \delta \Rightarrow f(x) > 0$.

D: Suppose that $f$ is continuous on $[a,b]$ and differentiable on $(a,b)$ and that $f(a) = f(b)$. Then $\exists c \in (a,b)$ s.t. $f'(c) = 0$. 

Our first aim was simply to investigate the extent to which students embarking upon a real analysis course were able to produce diagrams to represent such statements. Our second aim was to find out whether ability to draw diagrams like these is systematically related to performance in the course.

THEORETICAL BACKGROUND

In our study, students were asked to construct diagrams to represent statements written in a typical combination of words and symbols. They were thus required to translate between representation systems (Goldin, 1998), a process Duval (1999) calls conversion. Students are required to perform many such conversions during their mathematical education, and ability to do this is seen as evidence of mathematical understanding: both policy documents (NCTM, 2000) and research-related arguments (Janvier, 1987) stress its importance in flexible mathematical problem solving.

Undergraduate mathematics can also involve diagrams, and the intention is often that a diagram be interpreted as generic – as representing a whole class of functions, say. An individual might draw a diagram to facilitate proof construction (Gibson, 1998) via semantic reasoning (Alcock & Inglis, 2008; Goldin, 1998; Weber & Alcock, 2009), and both mathematicians and undergraduate students can and do use diagrams to understand statements (Gibson, 1998) and to explore relationships (Weber & Alcock, 2009; Stylianou & Silver, 2004). Diagrams arguably have particular utility for such purposes, because they allow simultaneous external representation of multiple aspects of a problem (Pantazaria, Gagatsis & Elia, 2009). They can thus facilitate imagined variation of one or more of these aspects (Tall, 1995), recognition of relationships that may not be obvious from a problem statement (Pólya, 1957), and the correct set-up of equations necessary to solve a problem (Bremigan, 2005).

The extent to which a diagram is useful might vary, however. This observation is key to our study because we are interested in judging the value of diagrams produced in response to a direct request. In this paper we use four criteria to capture each diagram’s possible value in supporting further semantic reasoning.

Our first criterion is correctness. If a diagram does not correctly represent the relationships under consideration, this shows that the person who produced it does not understand the statement or was not (in this instance) able to convert between representation systems appropriately. Either way, an incorrect diagram will not reliably lead to productive and correct further reasoning.

Our second criterion is genericity. For a diagram to function as a generic example, it should be neither too trivial nor too complicated, and analogy with other instances should be readily achieved (Rowland, 2002). In our context, a diagram might be too trivial if it incorporates function properties that oversimplify the situation: a function might be drawn as constant or monotonic or always positive, for instance (Haciomeroglu, Aspinwall & Presmeg, 2010). A diagram might be too complicated if it includes potentially distracting irrelevant features such as multiple axis crossings.
or asymptotes. A diagram that is too simple might suggest invalid inferences, and one that is too complicated might impede focus on key properties.

Our third criterion for judging diagrams is quality of labelling. Incorrect labelling can result in misrepresentation, and a more subtle possibility is that some mathematical objects might not be explicitly labelled. This could be important because experts might be more consistent than novices in producing fully-labelled diagrams (Stylianou & Silver, 2004) and because quality of labelling might be a factor in enabling translation from a diagram-based insight to a formal argument, a process that can be difficult (Alcock & Weber, 2010; Weber & Alcock, 2009).

Our fourth criterion emerged during data analysis, so it is described in the Method section below. The Method section also describes data we collected in order to investigate any relationship between diagram construction and performance in the analysis course. Theoretically there could be such a relationship: ability to use and convert between a variety of mathematical representations might support understanding and semantic reasoning. On the other hand, performance in courses like real analysis is traditionally measured via formal work with definitions, theorems and proofs, and a student could learn to do such work without attending to diagrams.

**METHOD**

**Task design**

All participants were asked to draw a diagram to represent each of the four statements listed in the introduction; the order of presentation was randomised so that participants saw different versions of the task. We selected the statements from the real analysis course, using the following criteria. First, we wanted all terminology and symbols to be familiar to the students, so that ability to construct diagrams would not be confounded with ability to interpret the components of the statement. Second, we wanted statements which would be accessible but which the participants had not studied before, so that they would not try to ‘remember’ an appropriate diagram (for this reason we did not flag any of the statements as a definition or theorem). Third, we wanted statements for which participants would not be tripped up by inattention to the subtler points of calculus or analysis. For example, differentiability is key to statement D (Rolle’s Theorem), but attention to this was unlikely to be problematic because students tend to think about differentiable functions; completeness of the real numbers is key to the statement B (Intermediate Value Theorem), but this was unlikely to cause problems under typical naïve conceptions of continuity.

Because of these criteria, we did not expect the participants to have trouble in literal reading of the statements. Nevertheless, we wanted to exclude the possibility that any apparent difficulties in diagram construction resulted from an inability to read the statements. We thus also asked the participants to write out in words exactly what each statement said. Except for occasional minor awkwardness in English expression, there was no evidence that any participant had difficulty reading the
statements. To establish that this writing did not, in itself, improve diagram construction, we asked half of the participants to complete the drawing task first and half to complete the writing task first (the order in which the statements appeared in the writing task was randomised too). Results of this manipulation are reported later.

**Participants and administration**

A total of 100 students took part; 75 were in the second year of a single-honours mathematics degree and 25 were in the third year of a joint-honours mathematics degree. All had high pre-university mathematical attainment, all had studied two-semester courses in calculus and linear algebra, and all spent 50-100% of their study time on mathematics. The task was administered at the beginning of the first lecture in the real analysis course. Participants were given a booklet and asked to fill in a cover sheet stating that they understood that their responses would also be used for research and asking them to provide their ID number if they gave permission for the researchers to link this to information from the university database (all students gave this information). Drawing-first and writing-first versions of the task were interleaved so that students sitting next to each other did not receive the same type of task first. The participants were given 10 minutes to complete whichever task was first in their booklet, which asked them not to turn over until told to do so. When told to turn over, they then had the same amount of time to complete their second task.

**Data analysis**

Before analysing the student-produced diagrams, we collected diagrams for each statement from three mathematics lecturers (one author of this paper and two with recent calculus lecturing experience). This confirmed that expert diagrams were broadly similar. It also prompted us to introduce our further criterion for judging diagrams, for the following reason. Figure 1 shows two expert diagrams for statement C. Both are accurate, generic and fully labelled: as required, $|x - a| < \delta \Rightarrow f(x) > 0$. Also, however, outside the region where $|x - a| < \delta$, the function does take on values that are less than zero. This goes beyond the literal statement to indicate a sense of what is mathematically important about the claim.

![Figure 1: Expert diagrams for statement C.](image-url)
We thus awarded each student-constructed diagram a score of 0, 1 or 2 for correctness (to allow for partially correct answers) and 0 or 1 for each of genericity, labelling and this new criterion, which we termed completion. This gave us a score out of 5 for each statement and an overall score out of 20 for each participant.

While scoring, we had reason to believe that seven students had misunderstood the task instructions (most had written instead of drawing; one had apparently begun trying to prove the statements); a further three were repeating the course. These ten students were excluded, leaving 90 participants for the descriptive analyses. For the remaining participants we collected prior performance scores from their earlier calculus course (as percentages); we considered calculus to be the most relevant as preparation for our task. We used the students’ eventual real analysis final examination scores in two ways, looking at both raw score and a standardised score which excluded points from question parts that involved drawing diagrams. Of the 90 students who completed the drawing task, ten did not take the analysis examination and for one a calculus mark was not available. Thus the analytical results are based on a total of 79 participants.

RESULTS

Descriptive results: student-produced diagrams

Scores were low: the mean out of 20 was 7.0 (standard deviation 5.19) and, of the 90 participants, 14 scored zero. Statement B (the Intermediate Value Theorem) appeared easiest, with the highest mean score of 2.5. Figure 2 illustrates the types of errors and misinterpretations that can arise by showing two low-scoring student-produced diagrams (more diagrams will be shown in the presentation if this report is accepted).

![Figure 2: Low-scoring diagrams for statement B.](image)

For statement D (Rolle’s Theorem) the mean score was 2.1; for statement A (boundedness definition), 1.6, and for statement C (lemma), 0.8. Figure 3 shows low-scoring participant-produced diagrams for statement C. Very few participants were able to correctly represent the meaning of this statement – delta was rarely labelled in any way – and none captured the completion aspect.
Analytical results: drawing scores and performance

As noted in the Method section, half of the students were asked first to draw and half were asked first to write out the statement in words. Diagram construction scores for the writing-first group ($n=39; \bar{m}=8.21, \text{s.d.}=4.97$) were slightly higher than those for the drawing-first group ($n=40; \bar{m}=7.08, \text{s.d.}=5.30$), but this difference was not statistically significant ($t=0.98, p=0.33$) so it is not used in further analyses.

Two linear models were considered, the first using the raw real analysis examination score as the dependent variable, and the second using the amended real analysis examination score as the dependent variable. In both cases, independent variables were the participants’ calculus score, drawing-task score, year of study and interaction terms between year of study and calculus and drawing-task scores. In both models, all of the interaction terms and also year of study were found to be non-significant and were thus excluded. Both calculus score and drawing-task score were found to be statistically significant in both models, as shown in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Raw analysis exam score</th>
<th>Amended analysis exam score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE B</td>
</tr>
<tr>
<td>Calculus</td>
<td>0.48</td>
<td>0.11</td>
</tr>
<tr>
<td>Drawing</td>
<td>1.57</td>
<td>0.36</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.49</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary of regression analysis for variables predicting raw and amended analysis examination scores; *$p < .005$.

In both cases, the estimated coefficients indicate that each additional 1% scored in calculus is associated with approximately an additional 0.5% in real analysis. More interestingly, each additional point out of 20 scored in the drawing task is associated with an additional 1.5% in real analysis. The standardised coefficients indicate that the predictive power of the drawing task score is on a par with that of prior attainment in calculus, even when performance in real analysis is measured exclusively via standard formal work.
DISCUSSION

The low scores on our drawing task indicate that constructing diagrams was not easy for participants. This could be considered unsurprising given that these students had no specific training in constructing diagrams for statements of this type, but it provides evidence regarding whether we can expect students at this level to make good use of diagrams in semantic reasoning. If students cannot produce such diagrams when specifically asked to, it seems unlikely that they would use them effectively as a natural part of reasoning. Of course, our study does not provide information on whether students can correctly interpret diagrams provided by others. Interpretation might be considerably easier than construction, and further research would be required to investigate whether this skill is related to academic success.

The relationship between drawing-task score and examination performance indicates that skill in producing diagrams might be an important factor in successful learning of advanced mathematics. It should be interpreted with caution, because all of the participants were enrolled in one course; this study does not enable us to tell whether this skill would be useful in any real analysis course, or whether features of the teaching simply made it useful in this course. It certainly does not provide evidence that this diagram construction skill is useful across the curriculum; it could be that it is of benefit in real analysis but not, say, in abstract algebra. Nevertheless, our findings provide reason to investigate diagram construction at this level more broadly, perhaps as one of a number of distinct mathematical skills that might benefit students in different mathematical domains.

Finally, examining expert-produced diagrams and scoring student-produced diagrams prompted us to articulate a clearer theoretical conceptualisation of what constitutes a good diagram. But student-produced diagrams also provide a valuable window into individual comprehension and mathematical reasoning, and we plan to report further on this issue in future work.

References


