Loughborough
University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

| | |
|---|---|
| Title: | UNDERGRADUATES' EXAMPLE USE IN PROOF |
| | CONSTRUCTION: PURPOSES AND EFFECTIVENESS |
| Running Head: | USE OF EXAMPLES IN PROOF CONSTRUCTION |
| Authors: | Lara Alcock, Loughborough University, UK |
| | Keith Weber, Rutgers University, USA |
| Correspondence: | Lara Alcock |
| | Mathematics Education Centre |
| | Loughborough University |
| | Loughborough |
| | LE11 3TU |
| | UK |
| | lalcock@lboro.ac.uk |
| | Tel:    +44 1509 228255 |
| | Fax:    +44 1509 228211 |

Abstract

In this paper, we present data from an exploratory study that aimed to investigate the ways in which, and the extent to which, undergraduates enrolled in a transition-to-proof course considered examples in their attempted proof constructions. We illustrate how some undergraduates can and do use examples for specific purposes while successfully constructing proofs, and that these purposes are consistent with those described by mathematicians. We then examine other cases in which students used examples ineffectively. We note that in these cases, the *purposes* for which the students attempted to use examples are again appropriate, but the implementation of their strategies is inadequate in one of two specific ways. On this basis we identify points that should be borne in mind by a university teacher who wishes to teach students to use examples effectively in proof-based mathematics courses.


Key words: advanced mathematical thinking, college, counterexamples, definitions, equivalence relations, examples, functions, mathematics education, proof, transition-to-proof courses, undergraduates

Introduction

A primary goal of advanced mathematics courses is to improve undergraduates' abilities to construct formal proofs. Numerous research studies have documented undergraduates' difficulties in this regard (e.g., Bills & Tall, 1998; Harel & Sowder, 1998; Hart, 1994; Moore, 1994; Sowder & Harel, 2003; Weber, 2001; Weber & Alcock, 2004), and there is a growing body of research on their specific difficulties with proof (for reviews of this literature, see Harel & Sowder, 2007; Selden & Selden, 2008; Weber, 2003). While there has been considerable research on undergraduates' difficulties with proof, some researchers have noted that there has been comparatively little work about the processes that undergraduates use when they attempt to construct proofs, and that more research of this type is needed (e.g., Harel and Sowder, 2007; Hart, 1994; Selden and Selden, 2003; Weber, 2001, 2006). In this paper, we examine just one aspect of undergraduates' proof construction: the ways in which they use examples of mathematical concepts to aid them in constructing proofs about those concepts.


The use of examples in proving

The term "example" has wide applicability in mathematics and mathematics education, with some authors using this word in a very broad sense. Watson and Mason, for instance, include "examples of classes, examples illustrating concepts, worked examples demonstrating techniques, examples of problems and questions which can be resolved, examples of appropriate objects which satisfy certain conditions, examples of ways of answering questions, constructing proofs, and so on" (Watson & Mason, 2002, 2004). Zazkis and Leikin (2008) use a task in which student teachers are asked to

generate examples of definitions. In this paper we restrict our use of the term to mean a mathematical object satisfying the definition of some concept. For instance, 6 is an example of an even number and $f(x) = x^2$ is an example of a continuous real-valued function. The latter could, of course, be represented graphically rather than via a formula, and we consider such a graph to be an example too. Various students in our study used graphical representations of examples, treating these either as specific or as generic representatives of a broader class. We believe that this use of "example" to mean either a specific or generic representative of some defined class is probably the most common intended meaning of the term when it is used by mathematicians and mathematics educators in the context of proof-oriented mathematics.

There is evidence that mathematicians consider the use of examples to be an important component of proof construction. In his discussion of the ways in which mathematicians solve problems and construct proofs[1], for instance, Polya (1957) suggests obtaining ideas for why a general assertion might be true by "looking at simple cases." Several authors, including Mason and Pimm (1984) and Balacheff (1988), argue that a demonstration of why a general assertion holds true for an arbitrarily chosen generic element can provide the basis for a formal proof of the assertion. Alcock (2004, 2009) found that mathematicians taking part in a research study indicated that examples were useful to them for three specific purposes: understanding the meaning of mathematical statements, generating ideas for how a statement might be proven, and checking that inferences drawn within a proof are not invalid. Alcock and Inglis (2009) exhibit a case

---

[1] This passage is referencing Polya's *How to Solve It*, a volume about how to solve mathematical problems. Proof strategies are not explicitly listed as a topic of discussion. However, to some (perhaps most) mathematicians, problem solving and proving are intricately linked. Proof construction is a problem solving task. Likewise, a problem is not considered fully solved until one can prove that its solution is correct. As a result, many of the problem solving heuristics Polya discusses may be interpreted as proving heuristics.

study in which a successful mathematics doctoral student uses examples extensively while evaluating and attempting to prove conjectures.

Mathematics educators have also commented on the use of examples in mathematical reasoning. Building on Polya's work, both Mason (1985) and Schoenfeld (1985) illustrate the variety of ways that Polya's "looking at simple cases" strategy can be implemented to construct proofs. Garuti, Boero, and Lemut (1998) contend that considering examples can not only help students to determine whether an assertion is true or false, but that it can and should also play a role in proving that the same assertion is true or false. Rowland (2001) argues for the potential of generic examples as tools to help make general proofs more accessible to undergraduate students. He reports improvement when teaching a number theory proof in a generic rather than a general version, though he and others (e.g., Mason & Pimm, 1984) caution that a specific example that is seen as generic by a teacher may not be recognized as such by a learner.

This provides some indication of how we would hope that students might use examples to understand statements and concepts and to construct and understand proofs. However, there has been only limited research on how undergraduates *actually* use examples in such contexts, and this research has produced mixed results. Example usage seems to help *some* undergraduates in these ways. For instance, Dahlberg and Housman (1997) reported that undergraduates who generate examples when learning about a formal mathematical concept tend to develop richer and more accurate images of that concept. Gibson (1998) illustrated how the inspection of examples (in the form of diagrams) helped some undergraduates overcome impasses in their real analysis proof-writing. On the other hand, it appears that many undergraduates will not generate

examples when working on such tasks. Moore (1994) found that undergraduates in a transition-to-proof course often could not prove statements about new concepts because they were both disinclined and unable to generate examples of those concepts. Housman and Porter (2003) found that when introduced to a new concept or asked to evaluate conjectures, many undergraduates did not generate examples unless they were explicitly prompted to do so. Similarly, the mathematicians interviewed by Alcock (2004, 2009) lamented that the undergraduates in their transition-to-proof courses rarely seemed to use examples when working on proofs.

One also needs to be aware that research on students' conceptions of proof reveals that undergraduates frequently hold beliefs about the relationship between examples and proof that are at variance with those held by the mathematical community. Some students behave as though simply checking to see that a general statement holds in several specific instances is sufficient to prove that the statement is true (e.g., Harel & Sowder, 1998, 2007; Recio & Godino, 2001). Other students believe that inspecting examples is useful in forming conjectures about the veracity of an assertion, but is not useful in generating a proof of that assertion (Harel, personal communication). This last is perhaps not surprising, given frequent exhortations by those teaching undergraduate mathematics that "an example is not a proof." Clearly this statement is true: an example by itself does not constitute a proof of a general assertion. Conversely, one does not need to consider examples of mathematical concepts to construct proofs about those concepts. So use of examples in proof construction is neither necessary nor sufficient for success. However, research demonstrates that mathematicians can and do use examples to aid

their proof construction, and suggests that it would be desirable for students to learn to do the same.

Methods

In this paper, we investigate the ways in which undergraduates enrolled in a transition-to-proof course used examples as they attempted to construct proofs. In the United States, transition-to-proof courses are common prerequisites for more advanced proof-oriented courses such as real analysis and abstract algebra. The goals of these courses include introducing students to formal notation, teaching them the rudiments of logic, and introducing them to standard proof techniques. Often, transition-to-proof courses mark students' first significant experiences with mathematical proof since their high school geometry course (Moore, 1994). For many students, a transition course is their first sustained study of mathematical proof altogether. We will address the following questions:

1. For what purposes do these undergraduates use examples in their proof construction?

2. **What factors inhibit these undergraduates from using examples productively in their proof construction?**

*Participants*

Eleven undergraduate mathematics majors volunteered to participate in this study. All of these participants were enrolled in a course entitled, "Introduction to Mathematical Reasoning," which was taught in Spring 2004 at a large state university in the

northeastern United States.  In this course, students practiced applying standard proof techniques and studied logic, sets, relations, functions, and some elementary group, number, and graph theory.  The participants came from three different sections of this course, each of which was taught by a different instructor.  All participants were interviewed near the end of this course, shortly after they had completed their second mid-term examination.

*Procedure*

Each interview proceeded as follows.  First, the participant was asked to describe their experiences in their transition-to-proof course, including their study habits, aspects of the course they found difficult, and their course performance. This portion of the interview was included primarily to help the participants become comfortable with the interview format.

Second, the participant was asked to complete the two proof construction tasks presented in Table 1.  Each of these tasks was chosen so that they could be sensibly approached with or without the use of examples.  For the relations task, for instance, a participant could examine particular functions that satisfy the relation or they could immediately try to prove from the definition that the relation was reflexive, symmetric, and transitive (see Alcock & Weber, 2005, 2009, for case studies of students who did each of these two things).  We note that the two tasks are of different types: one says "prove or disprove" and the other says "prove."  For a mathematician, these might prompt different types of activity, with the "prove or disprove" more likely to lead to initial exploration using examples.  We will see in the results sections that this did not

8

appear to be the case for our participants. The participants were presented with these tasks one at a time on separate sheets of paper, and were asked to describe what they were thinking about as they attempted each task. They worked without assistance from the interviewer until they either completed the task to their own satisfaction or reached a point where they felt they could no longer make progress. The interviewer then asked them why they had taken specific actions and, if appropriate, why they now found it difficult to proceed. Because we wished to avoid biasing the students' later responses in favour of discussion of examples, these questions focused on the student's choice of actions and conceptions of their own difficulties, rather than on the use of examples *per se*.

**\*\*\* Insert Tables 1 and 2 About Here \*\*\***

Third, in order that we might consider their proof attempts in a different context, the participant was asked to talk about their responses to two proof construction tasks on the mid-term examination they had taken the previous week. They were first asked to describe how they attempted to construct these proofs. They were then asked to discuss any difficulties they had experienced and whether they understood why their answers had been graded as they had. As the participants came from different sections of the transition-to-proof course, they completed different mid-term examinations. In choosing which questions to discuss with the participants, we strove to find questions that could have sensibly been approached with or without the use of examples. The particular questions that we discussed with the students are presented in Table 2.

Finally, the participant was asked to describe any general strategies that they used when writing and reading proofs.  Later, they were asked more specific questions about whether they focused on examples or "rules" during such work.

Participants were given as long as they wished to complete each of the interview tasks.  Most participants' interviews lasted approximately one hour.

*Analysis*

The interviews were transcribed and the transcripts were analyzed by the two authors.  First, each author independently read through two transcripts and identified each instance in which a participant introduced an example while completing one of the interview tasks.  Identified instances were compared and any discrepancies were discussed and rectified (discrepancies were few and usually the result of oversight rather than disagreement on what constituted an instance of example usage).

The effort to identify the purposes for which undergraduates used examples began with open coding.  The authors together examined instances in which individuals used examples, grouping similar uses into categories and producing initial names and definitions for these. After these categories were defined, each author independently inspected each instance in which a participant used an example and coded this as having been used for one or more of the purposes that had previously been defined. During this process, the authors met to compare their codings and discuss disagreements. Occasionally, definitions of categories were refined, new categories were formed, or multiple existing categories were collapsed into a single category.  This process continued until the coding was complete.  The result of this process was a set of four

major categories that was grounded to fit the available data (*cf.*, Strauss & Corbin, 1998).

These were:

- Understanding a statement;

- Evaluating the truth of an assertion;

- Generating a counterexample;

- Generating a proof.

After this open coding process, a graduate student affiliated with this project coded each instance of example usage as being a member of one or more of the authors' categories. The graduate student shared her findings with the authors of the paper and any disagreements between the authors' and the graduate student's codings were resolved.

The authors then looked at each instance of example usage and assessed whether the example usage had been appropriate and successful. In instances in which the example usage had been either inappropriate or unsuccessful, the authors wrote a description of what had occurred and a reason for this occurrence. Finally, for each proof construction task, each proof that a student produced was coded as being a valid proof (V), a valid proof with the exception of minor errors (E), or an invalid proof (I).

Results

The results presented here are in three sections. In the first section, we give an overview of the extent to which the participants used examples in their reasoning and the purposes for which they used these examples. This shows that, as in other studies, there were some who used examples regularly and some who never did so. In the second

11

section, we provide illustrations of specific instances in which participants used examples. This shows that at least some undergraduates can and do successfully use examples for purposes similar to those cited by mathematicians. In the third section, we provide illustrations in which participants used examples but were not successful in their proof attempts. This allows us to pinpoint two specific junctures at which example use might go awry.

*Students' use of examples*

Table 3 presents a summary of participants' performance and example usage on the relation task, the function task, and the mid-term examination questions. This shows that 6 of the 11 students used examples for at least some purposes during their attempts to construct proofs. Collectively, the participants were able to construct valid proofs or mostly valid proofs in only 5 of the 22 interview proof-construction tasks attempted; a finding that once again illustrates the serious difficulties that undergraduates have with proof. The table also shows the students' eventual grades for the Introduction to Mathematical Reasoning course (D is the lowest passing grade but a student is only allowed one mathematics course at grade D if they are to graduate as a mathematics major). One can see that in general, the few proofs that were at least partially valid were produced by those who went on to obtain higher grades.

**\*\*\* Insert Table 3 About Here \*\*\***

*Understanding a statement (US)*

The most frequent purpose for which participants used examples was to understand a mathematical statement. In these cases the participant would respond to a general statement pertaining to a class of objects by examining how that statement applied to a particular object within that class. In the excerpt below[2], Brad had just read the problem statement for the relation task and was attempting to make sense of the relation described (for which two functions $f$ and $g$ with domain $D$ are related if and only if there exists an $x$ in $D$ such that $f(x) = g(x)$).

B: Alright, I'm just going to like write out some examples to try and…like, set a D. And then…yes, write out a function or two. I don't know if that's going to help me.

[Brad writes  $D = \{1, 3, 5\}$   $f(x) = x^2$   $g(x) = x$ ]

B: Is it… you're not allowed to answer me, right? Would this be an example? Like where $f$ of $x$ is equal to 1, and $g$ of $x$ is equal to 1…and since $x$ is 1, like 1 is in the domain, $f$ is related to $g$?

In the above excerpt, Brad appeared initially unsure as to the nature of the defined relation. He introduced two functions, apparently using these in an attempt to develop a concept image of the new relation, a strategy that students successfully used in the studies by Dahlberg and Housman (1997) and Housman and Porter (2003). Later in the interview, Brad was asked to reflect upon this example usage.

---

[2] Transcripts of students' utterances were lightly edited to increase their readability. Repeated words, stutters, and expressions such as "um" were removed. Short sections of text carrying no meaning were also removed. No new words were added to the transcripts and in all cases, we believe that the meaning of the text was not changed. "…" denotes a short pause or that short segments of the text were removed. "[…]" denotes that a longer segment of the text was removed.

I:    I'm interested in the role that the example played for you.

B:    I think just to get an idea of what the relation was. […] when we're given a relation, a lot of times it's … "greater than or equal to" or something like that. This is kind of like a lot harder to see. […] I guess it just gives you something concrete that you could go to and related back to in one form. Like, take one example that might be this so that—because this is really general. And you can't really put your hands on this. You know I can't like, get a grasp of it.

*Evaluating the truth of an assertion (ET)*

When participants were asked to prove or disprove a general assertion about a class of objects, they would sometimes choose specific objects to see whether the assertion held for those objects. For instance, on her mid-term examination, Lisa was asked to prove or disprove that if $n$ was an integer greater than one, $n$ was divisible by a prime. On her exam, before offering a proof, Lisa wrote:

$n = 2$     $2/2 = 1$    √

$n = 3$     $3/3 = 1$    √

$n = 4$     $4/2 = 2$    √

When asked to describe her thinking on this question, Lisa replied:

L:    I was just trying to use induction, like if you know $n$ and $n+1$ then you know you can keep going. So I did the first few and they worked. So I just tried to set it up for $n+1$.

In describing their proving processes, a number of participants indicated that checking to see whether an assertion was true for several particular instances was one of the first things they did. Two excerpts illustrating this, from the interviews with Dan and Karen, are provided below:

I: When someone asks you to prove a statement, what are some of the first things that you do?

D: *Pause.* The first thing that I do is see whether or not I think it's true or false […]

I: Okay. So how do you do make that decision? How do you decide whether it's true or false?

D: Just work out examples in my head.


I: Do you have any general strategies that you use when you're asked to write a proof?

K: I just put – at first I try to find an example that works. And then I try to think of one that won't work. And if I can't find one, I just go off on the assumption that there's a way to prove it.

We will see more instances of students attempting to use examples for this purpose in the third results section.


*Constructing a counterexample to disprove a statement (CX)*

Some students used examples appropriately as counterexamples. For instance, on his mid-term exam, Dan was asked to determine whether the power set of a set is totally ordered under the inclusion relation. Dan wrote:

Not symmetric.
If $A \subseteq B$, that does not mean $B \subseteq A$.
Example. Let $A = \{1, 2\}$ and $B = \{1, 2, 3\}$. $A \subseteq B$. But $B \not\subseteq A$ since 3 is not an element of A.

In describing his thought processes for that question, Dan stated:

D: Okay for symmetry, A was related to B…that does not mean B is related to A. And I gave an example, say you have set A, which has the elements 1 and 2, and set B which has the elements 1,2 and 3. Um…A is under inclusion…all the elements of A

15

are in B, but all the elements of B are not in A since 3 is not…in A. So they wouldn't

be symmetric.

In the reflective segment of their interviews, it was common for participants to list

constructing counterexamples as a primary means of proving that a statement was false.

This was the case even among those who did not use examples, or describe their use, for

any other purpose.


*Generating a proof (GP)*

When constructing a proof of a general assertion, participants would sometimes

inspect a specific example with the purpose of understanding *why* that assertion should be

true. They would then attempt to use the gained understanding as a basis for constructing

a formal proof. This is illustrated in the excerpt below, as Ellen describes the thought

processes that she used to answer the second question on her mid-term exam. In this

question, a relation was defined on $\mathbf{R}$ x $\mathbf{R}$ such that $(a, b)$ is related to $(c, d)$ if both $a \leq c$

and $b \leq d$. The question asked the student both to prove that any two-element subset of $\mathbf{R}$

x $\mathbf{R}$ had an upper bound with respect to this relation and to find the least upper bound of

{(-1, 2), (3, -4)}. Ellen's solution to these tasks is provided in Figure 1 and the diagram

she used to generate this solution is presented in Figure 2. Her description of her solution

to this problem is given below:

E:  For the second part, where it says prove any two-element subset of R-squared has an

upper bound, um I guess I realized that it would be um...like I sort of skipped it and

went to the [next] part where he gave actual numbers. And I just realized that I could

draw a graph of it. So I did over here [referring to the graph in Figure 2] and I

realized the upper bound would be the corner of the sort of rectangle that it would

make…so I just thought that any upper bound would sort of just be to the right and up of your two points. So for this one, it was obvious that it was the corner of the square. The least upper bound would be sort of on the lines and then um, so I went back to the second part then and [inaudible] since it's real numbers…I guess pairs of real numbers so either way you can go an infinite amount. So for any, um, two pairs that he takes, you can always go um larger I guess. So that's why there's an upper bound.

**\*\*\* Insert Figures 1 and 2 About Here \*\*\***

Ellen described the way in which she found a least upper bound for a particular two-element set, and then saw how the visual and graphical technique she used to generate the least upper bound for two specific ordered pairs could be generalized to generate upper bounds for any two ordered pairs. This is somewhat different from the previous instances in that the example was provided to Ellen in the question. However, her professor stated that there was only one student who successfully answered this question, so her use of it was apparently not obvious to her classmates (Ellen was also the only student who provided valid proofs for both interview tasks). Later in the interview, Ellen described the ways in which visual examples help her generate proofs.

> E: [Visual examples] just help me understand why it [the statement to be proven] is true. And what makes it true. And so, by that I sort of see what I have to prove within the proof. But the actual way to prove it comes from the definitions and things that we learned in class.

*Ineffective attempts to use examples*

In the previous sub-section, we illustrated the different ways that participants used examples to understand mathematical assertions and to construct proofs. However, as Table 3 illustrates, there were many cases (eight out of nine) in which participants used examples in their proof attempts but were unable to construct a proof. In this section, we illustrate two specific ways in which the use of an example might fail.

### *Using examples that did not satisfy the definitions*

There were several instances in which participants instantiated general assertions with specific examples but did not choose these examples appropriately, so that they were led to false conclusions or invalid arguments. We illustrate this with the work of two participants on the relation task.

Lisa began the task by reminding herself that an equivalence relation would need to be reflexive, symmetric and transitive. She made a partially successful attempt to state the meaning of each of these terms before considering examples:

L: …Say $f(x)$ can be, like, $x+2$, you know, I would usually pick the same one because it

has to be equal. The $g(x)$ would be… the same thing… you're always using the same $x$.

You see what I'm saying?

Lisa then tried to check that the interviewer followed what she was saying; the interviewer asked her to continue at the time being and said that she would ask questions later if she didn't understand. The following discussion ensued.

L: So basically I'm just trying to find something related and then prove it. Like if I can

have an example for myself, then that works. Then I'll prove it… In this case, if I pick

$f(x)$… to equal … hmm [mumbles]… [very long pause]… see this is where, this is

18

where I go wrong. Like I pick this to be $x + 2$, so that's always going to work. But $f(x)$

can be anything and so could $g(x)$. So once I pick these examples and then I show

myself that that example works, and then it's okay because the $f(x)$ that I chose is right.

You know what I'm saying? So even though like $f(x)$ can be $x + 3$ and then it's not true,

so…

I: So then what's not true, exactly?

L: Well, $f(x)$ wouldn't equal $g(x)$.

I: OK.

L: So…

I: So… where are you in trying to prove or disprove this at the moment?

L: [pause] Well right now I'm thinking it's not going to work…

Lisa appeared to be trying to use examples for the purposes of understanding the

statement and providing a counterexample. However, like some other interview

participants, Lisa seemed to assume that relations, or at least equivalence relations, were

synonymous with equality. This lack of understanding of the concept prevented her from

selecting examples that were genuinely relevant to the situation (it is true that the

functions that Lisa generated do not satisfy the relationship in the problem statement, but

this is not because they are not equal). Throughout her proof attempt, Lisa never

explicitly referred back to the definition of the relation that appeared in the problem

statement.

The second illustration presents Andy's performance on the same task. In the

excerpt below, Andy introduces example functions to see whether the relation he is

considering is symmetric and transitive:

A: Alright. This is saying, for reflexive, $f(x)$ is related to $f(x)$. And $x$ is in $D$. Then it

will be yes because that's just like saying something is related to itself. Of course it

would be related to itself. $f(x)$ is related to $g(x)$, … let's say…$x$ squared. Let's say $f(x)$ is $x$ squared. Is equal to $x$ squared. So $x$ squared is equal to $x$ squared. $g(x)$ would be $x$ to the fourth. $f(x)$ would be $x$ squared and $g(x)$ would be $x$ to the fourth. Is $x$ squared related to $x$ to the fourth? Yes. Is $x$ to the fourth related to $x$ squared? Yes. But do they equal, no they do not equal. So it would be false. […]

I:   Okay. One more question about that. So you said that $x$ squared is related to $x$ to the fourth. Why is that?

A:   In the sense that they're both parabolas… $x$ to the fourth would be skinnier than that [$x$ squared], but it's still relatively the same.

In this excerpt, Andy appears to believe that functions are related if they are "relatively the same" or share some distinguishing feature. However, to Andy, this relation should qualify as an equivalence relation only if related functions are all equal, and he appears confident that the relationship he sees between his examples therefore gives him a counterexample. Like Lisa, Andy did not make explicit use of the definition of the relation that was provided in the problem statement; neither student seemed aware that it would be appropriate to check carefully whether their examples satisfied this definition in order to check that insight gained would apply to the problem situation.

Clearly these students had weak understanding of the concept of an equivalence relation and did not process the problem statement in a way that captured its logical structure. This led them to imagine that they had found counterexamples when their examples were, in fact, not suitable for this purpose. However, we note that in both cases, the purposes for which they attempted to use their examples *were* sensible: using examples to gain understanding is a reasonable strategy, as is attempting to disprove a statement by giving a counterexample.

*Inability to connect reasoning from examples to the language of formal proof*

In the interviews, we also saw several instances in which considering examples helped participants to reach and articulate a meaningful understanding of why assertions were true, but left them unable to translate this articulation into a formal proof. We illustrate this below, with Hannah's attempt to prove that an increasing function cannot have a global maximum. After reading the statement, Hannah immediately sketched a graph showing a generic increasing function and also representing the properties a global maximum would have to have. A reproduction of this is shown in Figure 3 (the original is small and the shading fainter).

**\*\*\* Insert Figure 3 About Here \*\*\***

Hannah then used these graphs to formulate an explanation for why increasing functions cannot have global maxima. In the excerpt below, Hannah explains to the interviewer the meaning of her diagrams and how they establish the statement to be proven:

I:   You're talking and drawing things on this picture but it's a very small version. Do you want to draw a bigger version of that and talk me through what you're thinking?

H:   Okay… well, drawing the graph of what some… some increasing $f$ of $x$ may look like… So increasing so it's just going on to infinity. Forever and ever and ever increasing. For the second one, if it's going to have a global max, at some $c$. That means everything, of the rest of it, of the rest of the graph is under that $c$, because that's the maximum. [Here Hannah is referring to the shaded area of the lower graph in Figure 3].

I:   Okay. Okay. I understand what that picture is now! *Laughs*.

21

H: *Laughs*. Wherever the rest of the graph is, it has to be under there. So…what you just need to be able to *prove* [with comedic frustration] is that wherever this $c$ happens to be, this … could just even equal this $c$ and you're…home free. Because…as $f$ of $x$ equals $f$ of $c$…then, you've already proved that this…that you can't have a global maximum.

In this segment, as well as elsewhere in the interview, Hannah was able to explain why an increasing function does not have a global maximum in terms of her graphical representation of an increasing function, which she appeared (appropriately) to treat as generic. However, she also indicated that it was not immediately clear how she should go about writing this argument as a proof. Later in the interview, having struggled several times to begin a proof, Hannah described her frustration at not being able to prove what seemed to her to be an obvious result.

I: The one other thing that I noticed you said at the beginning…you read these a couple of times and you said something about, "Well it's intuitively…," and then you kind of stopped. Can you explain what was going on there?

H: Just, it's – it's so like… it's so obvious, it's so like a second-grader could tell you this, sort of. Maybe – sixth grade, I don't know. But that doesn't mean much because you still have to figure it…

I: Okay. So the difficulty you're having now is not understanding why this must be how it is,

H: Right! Right.

I: But, to pin it down in a way…

H: Right. Right.

We note that once again, Hannah was attempting to use her (generic) examples for sensible purposes: to understand the statements and to generate a proof. Indeed, she

apparently succeeded in gaining insight into why increasing functions could not have global maxima. Her diagram and argument provided her with a strong conviction of the truth of the assertion (*cf.*, Alcock & Simpson, 2004), so much so that she believed a grade school student could see why it was true. Unlike Lisa and Andy, the examples that she constructed were correct relative to the definitions, and her inspection of them led her to produce a correct line of reasoning. She also recognized the need to couch her argument in the language of formal proof. However, she was unable to construct this proof. We suggest that this may be due, at least in part, to two connected factors: 1) a lack of detailed labels indicating connections between the diagram that she drew and the provided definitions, and 2) the fact that capturing insight the gained by examining a diagram requires imposing a linear logical order upon an intrinsically non-ordered representation.

Hannah's behavior was typical of other participants in this study. Brad, Karen, and Lisa all constructed examples of increasing functions and generated explanations based on these examples that were similar to Hannah's. (Brad's argument is studied in detail in Alcock & Weber, 2005, 2009). However, none came close to formalizing their arguments as proofs.

Discussion

This study confirms that examples can function in useful ways for students at the transition-to-proof level: in Brad we saw a case in which an example helped him make sense of a definition that he initially found confusing, and in Ellen we saw a case in which reference to a particular example led to the construction of a general proof and

23

provided confidence in a definition-based argument. Further, Ellen indicated that she could not construct her proof without considering that example. This study also confirms that there are pitfalls related to example use: in Hannah we saw a case in which a correct informal argument could not be formalized, and in Andy and Lisa we saw the use of inappropriate examples without an apparent inclination to check whether these corresponded to the given definitions. This latter especially is consistent with findings that when students' understanding of a concept is poor relative to the definition, the conclusions that they draw may be false (Vinner, 1991), but they may display considerable confidence in these (Alcock & Simpson, 2004) and not feel the need to verify their example-based conclusions by framing their arguments within the language of formal mathematics (Harel & Sowder, 1998).

These findings suggest that although the consideration of examples could be helpful for some undergraduates in the writing of some proofs, simply telling students to consider examples may not improve their proof writing performance. We suggest that the topic of using examples during proof construction could profitably be discussed in advanced mathematics classrooms. However, comments from mathematicians and mathematics educators suggest that this does not happen often. In describing the teaching of proof-oriented mathematics courses, Davis and Hersh (1981) claim that if examples of mathematical concepts are given at all, their presentation tends to be "parenthetical and in brief" (p. 151). Similarly, Dreyfus (1991) contends that advanced mathematical courses are taught in a highly formal manner in which informal aspects of mathematical concepts, such as examples of these concepts, are not covered. Our investigations into the teaching of transition-to-proof courses (Alcock, 2004, 2009) and other advanced mathematics

24

courses (Weber, 2004) also indicate that the use of examples does not receive the attention that it could, in part because professors wish to avoid confusion that could arise if students are encouraged to use examples on some occasions but are reprimanded for trying to "prove by example" on others.

We recognize this potential conflict, but wish to emphasize the fact that the students in our study, even when they used examples ineffectively, did <u>not</u> use them for inappropriate purposes. Although there may well be cases in which students offer a general example as a proof, our findings indicated that those in our study tried to use their examples for appropriate purposes but were ineffective for some other reason: for instance, although Lisa's and Andy's examples did not relate to the problem situation as they thought, they were trying to understand the situation and to produce counterexamples, both of which are perfectly sensible purposes.

We believe that it is important to recognize this because we do not want to give students the idea that their approach is completely wrong when in fact its intentions may be sensible. Rather, we want to help them improve on what is currently inadequate or ineffective about their approach. In our view, the major difficulties seen here are consequences of insufficiently explicit links between concept definitions and features of the examples used – the students either did not look for such links or were not able to use them to formalize an argument. Various authors have recently been discussing structured activities in which students generate examples with different combinations of properties in order to explore *example spaces*, with particular focus on *boundary examples* (Watson & Mason, 2004; Bills, Bills, Watson & Mason, 2004) or *pivotal examples* (Zazkis & Chernoff, 2008). We believe that instruction in the effective use of examples might

sensibly include both similar activities and explicit discussion of the purposes for which

it is appropriate to use examples in mathematical reasoning.

References

Alcock, L.J. (2004). Uses of example objects in proving. In M.J. Hoines & A.B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*: Vol.2 (pp.17-24). Bergen, Norway.

Alcock, L. (2009). Mathematicians' perspectives on the teaching and learning of proof. In F. Hitt, D.A. Holton & P. Thompson (Eds.), *Research in Collegiate Mathematics Education. VII* (pp.73-100). Providence: RI: American Mathematical Society.

Alcock, L. & Inglis, M. (2008). Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*, *69*, 111-129.

Alcock, L.J. & Simpson, A.P. (2004). Convergence of sequences and series: interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics*, *57*, 1-32.

Alcock, L.J. & Simpson, A.P. (2005). Convergence of sequences and series 2: interactions between non-visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics*, *58*, 77-110.

Alcock, L.J. & Weber, K. (2005). Referential and syntactic approaches to proof: Case studies from a transition course. In H.L. Chick & J.L. Vincent (Eds.), *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*: Vol. 2 (pp.33-40). Melbourne, Australia.

Alcock, L. & Weber, K. (2009). Referential and syntactic approaches to proof: Case studies from a transition-to-proof course. In F. Hitt, D.A. Holton & P. Thompson (Eds.), *Research in Collegiate Mathematics Education. VII* (pp.101-123). Providence: RI: American Mathematical Society.

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.) *Mathematics, Teachers and Children*. London: Hodder and Stoughton.

Bills, C., Bills, L., Watson, A. & Mason, J. (2004). *Thinkers: A collection of activities to provoke mathematical thinking*. Derby, UK: Association of Teachers of Mathematics.

Bills, E. & Tall, D. (1998). Operable definitions in advanced mathematics: The case of least upper bound. In A. Olivier and K. Newstead (Eds.) *Proceedings of the 22nd Conference for the International Group for the Psychology of Mathematics Education*: Vol. 2 (pp.104-111). Stellenbosch, South Africa.

Dahlberg. R.P, & Housman, D. L. (1997). Facilitating learning events through example generation. *Educational Studies in Mathematics*, *33*, 283-299.

Davis, P.J. & Hersh, R. (1981). *The mathematical experience*. New York: Viking Penguin, Inc.

Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory real analysis course. In A.H. Schoenfeld, J. Kaput, E. Dubinsky & T. Dick (Eds.), *Research in Collegiate Mathematics Education. III* (pp. 284-307). Providence: RI: American Mathematical Society.

Gray, E., Pitta, D., Pinto, M., & Tall, D. (1999). Knowledge construction and diverging thinking in elementary and advanced mathematics. *Educational Studies in Mathematics*, *38*, 111-133.

Harel, G. & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A.H. Schoenfeld, J. Kaput, E. Dubinsky & T. Dick (Eds.) *Research in*

*Collegiate Mathematics Education. III*, (pp. 234-283). Providence: RI: American Mathematical Society.

Harel, G., & Sowder, L. (2007). Toward a comprehensive perspective on proof. In F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 805-842). Charlotte, NC: National Council of Teachers of Mathematics.

Hart, E.W. (1994). A conceptual analysis of the proof-writing performance of expert and novice students in elementary group theory. In J. J. Kaput and E. Dubinsky (Eds.), *Research Issues in Undergraduate Mathematics Learning: Preliminary Analysis and Results*, MAA Notes 33, Washington, D.C.

Housman, D. & Porter, M.K. (2003). Proof schemes and learning strategies of above-average mathematics students. *Educational Studies in Mathematics*, *53*, 139-158.

Mason, J.H. (with Burton, L. & Stacey, K.) (1985). *Thinking Mathematically* (Rev. ed.). Harlow, Essex, UK: Addison-Wesley.

Mason, J.H. & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics*, *15*, 227-289.

Moore, R.C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics*, *27*, 249-266.

Pinto, M. & Tall, D.O. (1999). Student construction of formal theories: Giving and extracting meaning. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference for the International Group of the Psychology of Mathematics Education*: Vol.1 (pp.281-288). Haifa, Israel.

Pinto, M. & Tall, D.O. (2002). Building formal mathematics on visual imagery: A case study and a theory. *For the Learning of Mathematics*, *22(1)*, 2-10.

Polya, G. (1957). *How to Solve it* (2nd edition), Princeton University Press, Princeton, NJ.

Recio, A.M. & Godino, J.D. (2001). Institutional and personal meanings of proof. *Educational Studies in Mathematics*, *48*, 83-99.

Rowland, T. (2001). Generic proofs in number theory. In S.R. Campbell & R. Zazkis (Eds.), *Learning and teaching number theory: Research in cognition and instruction* (pp.157-184). Ablex Publishing, Westport, CT.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. Orlando, Fl: Academic Press Inc.

Selden, A. & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*, 4-36.

Selden, A. & Selden, J. (2008). Overcoming students' difficulties in learning to understand and construct proofs. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and teaching in undergraduate mathematics* (pp.95-110). Washington, DC: Mathematical Association of America.

Sowder, L. & Harel, G. (2003). Case studies of mathematics majors' proof understanding, construction, and appreciation. *Canadian Journal of Science, Mathematics, and Technology Education*, *3*, 251-267.

Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. London: SAGE.

Vinner, S. (1991). The role of definitions in teaching and learning mathematics. In D.O. Tall (Ed.), *Advanced Mathematical Thinking* (pp.65-81) Kluwer, Dordrecht,.

Watson, A. & Mason, J.H. (2002). Extending example spaces as a learning strategy in mathematics. In A. Cockburn & E. Nardi (eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*: Vol. 4 (pp.377-384). Norwich, UK,.

Watson, A. & Mason, J.H. (2004). *Mathematics as a constructive activity: Learners generating examples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Weber, K. (2001). Student difficulty in constructing proofs: the need for strategic knowledge. *Educational Studies in Mathematics*, *48*, 101-119.

Weber, K. (2003). Students' difficulties with proof. In A. Selden and J. Selden (Eds.) *Research Sampler*, 8. Available at:http://www.maa.org/t_and_l/sampler/rs_8.html Last downloaded: March 20, 2007.

Weber, K. (2004). Traditional instruction in advanced mathematics courses. *Journal of Mathematical Behavior*, *23*, 115-133.

Weber, K. (2006). Investigating and teaching the processes used to construct proofs. In F. Hitt, G. Harel & S. Hauk (Eds.), *Research in Collegiate Mathematics Education*. VI (pp.197-232). Providence: RI: American Mathematical Society.

Weber, K. & Alcock, L.J. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*, 209-234.

Zazkis, R. & Chernoff, E.J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics*, *69*, 195-208.

Zazkis, R & Leikin, R. (2008). Exemplifying definitions: A case of a square. *Educational Studies in Mathematics*, *69*, 131-148.

Table 1.

*Proof construction tasks attempted by participants*

Relation task.

Let $D$ be a set. Define a relation $\sim$ on functions with domain $D$ as follows.

$f\sim g$ if and only if there exists $x$ in $D$ such that $f(x) = g(x)$. Prove or disprove that $\sim$ is an equivalence relation.

Function task.

A function $f:\mathbf{R}\rightarrow\mathbf{R}$ is said to be <u>increasing</u> if and only if for all $x, y \in \mathbf{R}$, ($x > y$ implies $f(x) > f(y)$).

A function $f:\mathbf{R}\rightarrow\mathbf{R}$ is said to <u>have a global maximum at a real number $c$</u> if and only if, for all $x \in \mathbf{R}$ ($x \neq c$ implies $f(x) < f(c)$).

Suppose $f$ is an increasing functions. Prove that there is no real number $c$ that is a global maximum for $f$.

Table 2

*Questions considered from participants' mid-term examinations*

Class A.

1. Let S be a set with at least two elements so that there are subsets other than S itself and the empty set. Consider the power set S, *P(S)*. That is, the set of all subsets of S, under the inclusion relation $\subseteq$.

> a. Determine whether or not this relation on the power set is: i) reflexive, ii) symmetric, iii) anti-symmetric, or iv) transitive. Justify your claims.
>
> b. Determine whether or not the power set is partially ordered under the inclusion relation. Briefly explain your answer. (Don't say only yes or no).
>
> c. Determine whether or not the power set is totally ordered under the inclusion relation. Briefly explain your answer. (Don't say only yes or no).

2. Let a sequence $\{x_n\}$ be defined recursively by the algorithm:

$$x_1 = 1, x_2 = 2, x_{n+2} = x_{n+1} + x_n, \text{ for } n \geq 1.$$

Prove $x_n \leq x_{n+1} \leq 2x_n$ for all $n \geq 1$.

Class B.

1. Define the relation $\propto$ on $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ by $(a, b) \propto (c, d)$ if both $a \leq c$ and $b \leq d$. Prove that $\propto$ is a partial order on $\mathbf{R}^2$. Prove that any two-element subset of $\mathbf{R}^2$ has an upper bound with respect to $\propto$. What is the least upper bound of $\{(-1, 2), (3, -4)\}$?

2. Let $f = (A, B, R)$ [*authors' note: This means f is a function from domain A to range B with relation R*] and let $C$ be a subset of $A$. Let $S = R \cap (C \times B)$ and let $g = (C, B, S)$.

> a. Prove that $g$ is a function.
>
> b. Suppose that $f$ is one-to-one. Prove that $g$ is one-to-one.

1. Prove or disprove the following:

      a. If $A$, $B$, $C$, $D$ are sets such that $A \times B = C \times D$ and $A$, $C$ are non-empty, then $A = C$.

      b. If $A$, $B$, $C$, $D$ are sets such that $A \times B = C \times D$ and $B$, $D$ are non-empty, then $A = C$.

      c. If $A$, $B$, $C$, $D$ are sets such that $A \times B = C \times D$ and $A$, $B$ are non-empty, then $A = C$.

2. Prove or disprove: Every integer $n$ such that $n > 1$ is divisible by some prime number. (Hint: Use complete induction or well ordering; do not try to use ordinary induction).

Table 3

*Summary of participants' example usage*

| Student Name | Course Section | Course Grade | RELATION TASK Example Use | Validity of proof | FUNCTION TASK Example Use | Validity of proof | MID-TERM Example Use |
|---|---|---|---|---|---|---|---|
| Andy | 1 | D | Yes (US) | I | No | I | Yes (CX) |
| Brad | 1 | B+ | Yes (US, GP) | E | Yes (US, GP) | I | Yes (US) |
| Carla | 1 | A | No | I | No | V | No |
| Dan | 1 | D | No | I | Yes (US, GP) | I | Yes (CX, US) |
| Ellen | 2 | A | No | V | No | V | Yes (GP) |
| Fay | 2 | C+ | No | I | No | E | No |
| Gail | 2 | C | No | I | No | I | No |
| Hannah | 3 | C | Yes (US, ET, GP) | I | Yes (US, ET, GP) | I | Yes (ET, CX) |
| John | 3 | C+ | No | I | No | I | No |
| Karen | 3 | D | No | I | Yes (US, GP) | I | Yes (ET, US) |
| Lisa | 3 | C | Yes (ET, US) | I | Yes (US) | I | Yes (ET, US) |

Codings indicate the validity of proofs:
(V) valid proof
(E) valid proof with the exception of minor errors
(I) invalid proof

Purposes for examples are given in parentheses:
(US) understanding a statement
(ET) evaluating the truth of a statement
(CX) generating a counterexample
(GP) generating a proof

*Figure 1. Ellen's solution on her mid-term examination*

Take $\{(a,b),(c,d)\} \subseteq \mathbb{R}^2$.

There exists an $e \in \mathbb{R}$ such that $e \geq a$ and $e \geq c$. (for any two real numbers, there is a real number that is greater)

Also, there exists an $f \in \mathbb{R}$ such that $f \geq b$ and $f \geq d$.

Then by definition of $\unlhd$, $(a,b) \unlhd (e,f)$ and $(c,d) \unlhd (e,f)$. Thus, $(e,f)$ is greater or equal to $(\unlhd)$ any two-element subset of $\mathbb{R}^2$, so $(e,f)$ is an upper bound with respect to $\unlhd$.

least upper bound of $\{(-1,2),(3,-4)\}$ is $(3,2)$

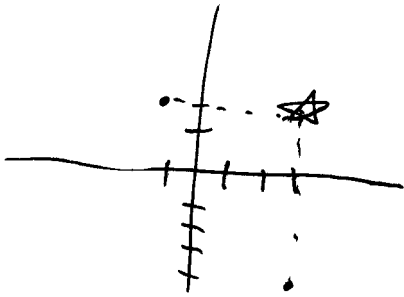*Figure 2. The diagram Ellen used to generate her proof*

*Figure 3.* A reproduction of Hannah's graph showing a generic increasing function and a shaded area indicating where other values of a function must lie relative to its global maximum