

Loughborough University Institutional Repository

Doctoral students' use of examples in evaluating and proving conjectures

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

Citation: ALCOCK, L. and INGLIS, M., 2008. Doctoral students' use of examples in evaluating and proving conjectures. *Educational Studies in Mathematics*, 69 (2), pp. 111-129.

Additional Information:

- The final publication is available at: www.springerlink.com

Metadata Record: <https://dspace.lboro.ac.uk/2134/8575>

Version: Accepted for publication

Publisher: © Springer

Please cite the published version.

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/>

Doctoral students' use of examples in evaluating and proving conjectures

Lara Alcock

Mathematics Education Centre, Loughborough University, UK

l.j.alcock@lboro.ac.uk

Matthew Inglis

Learning Sciences Research Institute, University of Nottingham, UK

matthew.inglis@nottingham.ac.uk

Abstract

This paper discusses variation in reasoning strategies among expert mathematicians, with a particular focus on the degree to which they use examples to reason about general conjectures. We first discuss literature on the use of examples in understanding and reasoning about abstract mathematics, relating this to a conceptualisation of syntactic and semantic reasoning strategies relative to a representation system of proof. We then use this conceptualisation as a basis for contrasting the behaviour of two successful mathematics research students whilst they evaluated and proved number theory conjectures. We observe that the students exhibited strikingly different degrees of example use, and argue that previously observed individual differences in reasoning strategies may exist at the expert level. We conclude by discussing implications for pedagogy and for future research.

1. Introduction

It is widely accepted that understanding the behaviour of those who are successful in advanced mathematics is relevant for the design of mathematical instruction, at least at the undergraduate level (e.g., Stylianou, 2002; Weber, in press). In this paper we consider one area of mathematical behaviour: the frequency and purpose of example usage during the evaluation of conjectures, and the construction of proofs. We approach this in two ways. First we discuss the use of mathematical examples in understanding and reasoning, and explain the relationships we see between example use and semantic/syntactic reasoning strategies (*cf.* Weber & Alcock, 2004). Second, we report two case studies of the activity of successful research students as they attempted to evaluate and prove conjectures about abundant, perfect and deficient numbers. We observe that the first student, Chris, used both specific and generic examples during his work, for multiple purposes including developing his understanding of abundant numbers, inductive testing of conjectures, checking his arguments, and generating counterexamples. The second student, David, used examples only for the unavoidable purpose of giving counterexamples, and for the arguably extra-mathematical purpose of illustrating his thinking when prompted to give more explanation. By exhibiting these data, we build on previous research which has suggested that some individuals may preferentially use semantic or syntactic reasoning strategies. In all we aim (i) to clarify the theoretical relationships between example usage and semantic and syntactic reasoning strategies, and (ii) to argue that the case studies we report indicate that individual differences in strategy preference may extend into experthood.

2. Theoretical background.

We begin our theoretical discussion by clarifying the sense in which we will use the term “examples” in this paper. We restrict our discussion to examples of mathematical concepts, so that our examples are illustrations or cases of mathematical objects in the sense discussed widely in process-object theories of mathematical development (e.g. Dubinsky, 1991). This use of the term example is relatively narrow; Watson & Mason (2002, p.378) include it in their definition of the term as “examples of appropriate objects which satisfy certain conditions,” but also list “examples of classes, examples illustrating concepts, worked examples demonstrating techniques, examples of problems and questions which can be resolved, ... examples of ways of answering questions, constructing proofs, and so on.” It is also different from what many students mean when they use the term to indicate a desire to see a worked example of a procedure. However, we believe that it is consistent with what mathematicians most often mean when they talk about an example of a concept.

Having established that we will use the term example in this sense, we now go on to discuss literature on ways in which examples of this type may be used by individuals for various purposes in understanding and reasoning. We then explain the relationship we see between these uses of examples and semantic/syntactic reasoning strategies, and review suggestive evidence that individuals may exhibit preferences for one or another overall strategy.

2.1 How might examples be used in mathematical understanding?

Much recent work has been dedicated to the design of situations in which learners generate examples in order to extend their knowledge of mathematical concepts (e.g., Watson & Mason 2002; 2005). Especially at the advanced level, such activity might be seen as exploring the extension of the concept definition, and thus obtaining a concept image (Vinner, 1991) that is consistent with this definition and faithfully represents the variation that is possible within the concept. Such use of examples to gain understanding has been reported both in mathematicians' descriptions of their own activity (Alcock, in press) and in the actual activity of students who are beginning to work with formal definitions (Alcock & Weber, in press; Weber, Alcock & Radu, 2005). Dahlberg and Housman (1997) argued that example generation for this purpose facilitated further work with the concept.

Watson and Mason also considered "boundary examples"; cases in which an example makes clear why a condition for a given theorem is required, by showing that the theorem would fail if it were not included (Mason & Watson, 2001). This gives certain examples special status relative to a mathematical theory, a position foreshadowed by Michener, whose characterisation of a full understanding of a mathematical topic included knowledge of examples of the following kinds:

- start-up examples, which motivate basic definitions and results or set up useful intuitions;
- reference examples, which are referred to repeatedly because they are basic, widely applicable and link results and concepts together;

- model examples, which are generic examples that suggest and summarize expectations and default assumptions about results and concepts;
- counterexamples, which show that a statement is not true and sharpen distinctions between concepts (Michener, 1978, p. 366-367; descriptions slightly abridged).

This literature suggests that examples can play a role in developing conceptual understanding, and that certain examples may be significant in a mature understanding of a mathematical theory. In the second part of this paper, we (obviously) will not see the participants reach a full understanding, but we will see one participant, Chris, use examples to develop his understanding of abundant numbers and the relationships between them.

2.2 How might examples be used in mathematical reasoning?

Once one has access to an example space (Watson & Mason, 2005) for a concept, examples from this space can be used in reasoning. One obvious use is in refuting conjectures, either by citing known “standard” counterexamples (Michener, 1978), or by constructing new ones. Providing a counterexample is standard practice in mathematics, and so is an unavoidable use of examples at an advanced level. However, examples can also be used in reasoning in situations in which they are *not* strictly necessary. In particular, we now consider the use of examples as an aid in constructing arguments about general statements.

An individual might attempt to verify or justify a conjecture by checking it against a small number of examples. When the conjecture actually applies to a large (possibly infinite) set of examples, this kind of checking was characterised by Balacheff (1987) as *naïve empiricism*. Similarly, Bills and Rowland (1999) used the term *empirical generalisation* to describe generalisation from a small number of numerical observations, and Harel (2001) used the term *result pattern generalisation* to describe similar behaviour in the context of mathematical induction.

At a more sophisticated level, Balacheff (1987) identified the *crucial experiment*, in which a conjecture that seems to be true for some instances is checked against an example that is assumed to have no special properties (citing a case in which a student suggested using a huge figure to check a conjecture about the number of diagonals of a polygon). Balacheff noted that this approach at least recognises the need for generalisation and addresses it in some way. Weber (2008) described a case in which a mathematician used such reasoning to check an assertion within a possibly incorrect elementary number theory proof. In general there is evidence that successful mathematicians do not avoid using empirical strategies, but do appropriately qualify conclusions based on such reasoning (Inglis, Mejia-Ramos & Simpson, 2007).

At a level beyond empirical generalisation, an individual might treat an example as a *generic* representative of a class, so that operations upon it would apply similarly to other members of the class. Bills and Rowland (1999) used the term *structural generalisation* to describe generalisations made on this basis, and in the context of mathematical

induction, Harel (2001) used the term *process pattern generalisation* to describe similar reasoning. Mason and Pimm (1984) noted that treating an example as generic requires seeing the general in the particular, and Weber (2008) again offered illustrations of cases in which mathematicians use generic examples in their validation of possibly incorrect proofs.

Rowland (2001) considered in detail the idea of a “generic proof”, which he argued may be more illuminating for learners than a completely general proof. He listed the following as criteria for a good generic proof:

- the particular case used should be neither too trivial nor too complicated;
- we need to be able to “track” any value used throughout the proof;
- aspects that are invariant across other examples should be emphasized;
- the reasoning should be constructive, in the sense that where it is argued that objects exist, they should be identified;
- scaffolding will be needed to enable a transition to a formal, generalised argument (Rowland, 2001, pp.167-168; descriptions abridged).

Rowland’s primary concern was a teaching context, but we believe that the same principles apply for an individual seeking to use a generic example in their own reasoning. Further, we note that effectively selecting a generic example demands a good grasp of the example space one is trying to represent.

In summary, this work suggests that examples can play a role in the refutation and justification of general conjectures and in communicating mathematical arguments to

others, and it gives some indication of what it takes to use examples effectively for these purposes. In the second part of this paper, we will see both participants using counterexamples in refutations, and one, Chris, using both specific and generic examples in his reasoning.

2.3 Examples and semantic/syntactic reasoning strategies

We now relate our discussion of example use to the notions of semantic and syntactic strategies for mathematical understanding and reasoning, with a particular focus on proof production. We begin by observing that the majority of the uses of examples discussed above would not appear in a written proof's final form. While these uses may be valuable to an individual in understanding a concept or testing or justifying a conjecture, or in making a presentation more accessible to an audience, this type of reasoning does not usually form part of an "official" written proof, which instead is written using generalised symbolic statements. We explain what we mean by this below, following Weber and Alcock (in press) in describing syntactic and semantic reasoning relative to a representation system in Goldin's (1998) sense.

In Goldin's (1998) terms, a representation system consists of primitive characters (e.g. letters, words, numerals), together with a syntactic structure comprised of rules for combining characters into permitted configurations, and rules for moving from one permitted configuration to another (Goldin & Kaput, 1996). Here we define the *representation system of proof* to consist of generalised symbolic statements which can be combined into permitted configurations via the rules of, for example, predicate

calculus, propositional logic and acceptable proof frameworks (in this we follow Weber & Alcock, in press). It is within this representation system that to-be-proved statements in mathematics, along with their proofs, are typically formulated. For instance, consider the following theorem and proof (used in this study):

Theorem. If p_1 and p_2 are primes, then p_1p_2 is not abundant¹.

Proof. If p_1, p_2 are distinct, we need to show that $1 + p_1 + p_2 + p_1p_2 \leq 2p_1p_2$. This can be rearranged to $(p_1 - 1)(p_2 - 1) \geq 2$ which is clearly true as at least one of p_1, p_2 is greater than 2. If $p_1 = p_2$, the inequality reduces to $1 + p_1 \leq p_1^2$ which is true for all integers greater than 1.

The theorem here is written entirely using general symbolic statements which are formed into a permitted configuration using propositional logic, i.e. the to-be-proved statement is written within the representation system of proof. Each step in this proof also involves manipulating configurations in ways that are acceptable within the same representation system as the to-be-proved statement. The premise is handled in two cases (a standard way of structuring a proof). In each, the definition of abundant is used to write a generalised symbolic version of what needs to be shown, algebraic manipulations are performed on the resulting inequality, and an observation is made about the rearranged inequality using general information about the primes p_1 and p_2 . p_1 and p_2 stand for general primes throughout.

¹ An abundant number is an integer n whose divisors add up to more than $2n$.

We note that this proof could be re-ordered and written more tersely – the phrase “this can be rearranged to” could be replaced by a sequence of algebraic manipulations or just the symbol “ \Leftrightarrow ”; the “we need to show that” phrase could be omitted with the inequality appearing only in a conclusion line, etc. In general, proofs include explanatory text designed to help a reader, so that not every phrase will be contained within the representation system of proof (as not every phrase obeys the same rules for formation into permitted configurations as the generalised symbolic statements within the proof). Nevertheless, the “backbone” of a typical proof (at least at the undergraduate level) consists of sequences of statements which are written within the representation system of proof in our sense, and follow from one another according to the rules of that representation system.

Given this characterisation, a proof production can take place following one of two strategies: via a syntactic strategy or via a semantic strategy (Weber & Alcock, 2004, in press). During a *syntactic* proof production the reasoner works entirely within the representation system of the to-be-proved statement, i.e. the representation system of proof. That is to say, they develop their understanding of the concepts involved in the proof by relating the relevant definitions (or components thereof) to known configurations and rules within the representation system of proof, and reason by moving between agreed configurations such as definitions and theorem statements by applying the rules of logic, standard proof frameworks and so on. The proof given above could have been constructed in this way. (We do not claim that an individual would arrive at a concise proof on a first try, but they may proceed by trying similar translations into

definitions, algebraic manipulations, invocation of standard proof frameworks etc. until finding a way to combine these to yield the desired result.)

In contrast to the syntactic strategy, during a *semantic* proof production the individual works, at least in part, within a different representation system to that of the to-be-proved statement. To explain what we mean, we consider the way in which characters or configurations in one representation system may sometimes be seen as representing characters or configurations in another. Goldin and Kaput (1996, p.400) illustrated this using a graph drawn in Cartesian coordinates to “represent” the equation $y + 3x - 6 = 0$. In that case, one configuration in one system represents one configuration in another, but in generalised symbolic statements it is often the case that a single letter or word is used to represent a whole class of objects; in our illustrative theorem, the configurations p_1 and p_2 each stand for particular but non-specific (in the sense of Mason & Pimm, 1984) prime numbers. Making use of this, one may understand and reason about the conjecture *semantically* by going *outside* the representation system of proof to consider examples of prime numbers, experiment with multiplying them together, check whether the outcome is abundant, search for patterns in factors that might explain the result, and so on. While this work might use some of the same underlying characters as the representation system of proof (numerals, letters etc.), it is not bound to follow the same rules as those that operate within that system. In particular, it may involve making holistic observations about properties that seem to be shared by examples considered, rather than deducing these properties by manipulating acceptable configurations according to logical rules.

In general, in order to understand semantically, an individual might explore an example space or relate particular examples to results within a theory; in order to reason semantically, an individual might move from the representation system of the to-be-proved statement to consider specific or generic examples and search for patterns or shared properties etc. A successful semantic proof production then, requires that the reasoner accurately translates aspects of the to-be-proved statement (and its associated concepts) into a new representation system, and that they reason productively within this new system. In the case of moving from a symbol that represents a whole class of objects to specific examples from that class, this means that these examples must actually have the properties they are supposed to have, and that one should be careful to avoid assuming that any observed properties of the specific examples hold for the whole class and not just for some special cases. In addition the reasoner must, when an adequate understanding has been generated within the new representation system, translate this understanding back into the representation system of proof. It may be non-trivially difficult to marshal general notation well enough to perform this final step, even when the structure of a general proof closely mirrors an argument based on a generic example (Rowland, 2001), and especially if it requires imposing a linear argument structure on a holistic observation about some property of the examples (for illustrations showing students having this latter difficulty, see Alcock & Weber, in press, Weber & Alcock, in press).

Figure 1 contrasts semantic and syntactic reasoning relative to a representation system of proof (the specific behaviours do not constitute an exhaustive list). Note that we do not

assume that semantic reasoning must necessarily involve example use. Informal rephrasings of mathematical ideas and physical gestures (e.g. Cook, Mitchell & Goldin-Meadow, 2008; Greiffenhagen & Sharrock, 2005) could also be characterised as semantic strategies since they also involve representation systems external to the representation system of proof. (We focus on examples in this paper, but note that other semantic work may also play a significant role in understanding and reasoning).

Insert Figure 1 here.

Of particular interest is the case of proof by counterexample. While one could, in theory, produce most proofs in an entirely syntactic manner, refutations often require considering the referents of general statements and selecting a counterexample from among them. So, as the to-be-proved statement is in a different representation system to that where the bulk of a counterexample search would take place, refuting a statement does, under the characterisation outlined above, require semantic work. However, this type of semantic work is somewhat different to the sort discussed above; crucially, there is no need for the reasoner to undertake the second translation between representation systems (back into the representation system of proof) that would be required if a typical proof was to be produced. That is to say that a proof by counterexample is presented within a different representation system to that of the to-be-proved statement. We characterise this type of proof production as involving semantic reasoning and a *semantic product*, in contrast to

the proof production discussed above which also involves semantic reasoning, but has a *syntactic product*. Producing a proof by counterexample is not unique in this respect: the production of a generic proof, and the use of an example as a pedagogical aid to understanding a proof, can also both be characterised as involving semantic reasoning and product (the reasoning and product are both located in a different representation system to the to-be-proved or to-be-explained statements respectively). The three types of proof production discussed in this section are summarised in Figure 2.

Insert Figure 2 here.

2.4 Possible preferences for semantic or syntactic strategies

Our reason for making the distinction where we do is that there may be considerable variation in the extent to which individuals use either a semantic or a syntactic strategy. Clearly successful mathematicians are likely to have many skills associated with both semantic and syntactic understanding and reasoning, which they deploy in response to different stimuli and/or desired outcomes. However, there is a small but accumulating body of evidence suggesting that some individuals might use one considerably more than the other. Both Pinto & Tall (1999, 2002) and Alcock & Simpson (2004, 2005) studied students' thinking in Real Analysis, and both teams noticed that some students regularly used visual representations to assist in their reasoning, whereas others rarely did this, and instead worked with mathematical concepts by memorizing definitions and attempting to

extract meaning from these definitions. Although neither team couched their research in terms of examples or semantic/syntactic reasoning, the visual images used by the students served as (often generic) examples in our use of the term, and the contrasting strategy is consistent with a syntactic approach. Weber and Alcock (2004) did use the semantic/syntactic distinction in discussing studies of undergraduate students in group theory and real analysis, and they observed similar tendencies in the context of a beginning course on proof (Alcock & Weber, in press; Weber, Alcock & Radu, 2005).

Less material is available on the habits of working mathematicians, but the possibility of individual preferences is consistent with introspective and philosophical reflections. Successful mathematicians have self-reported as having different reasoning styles from their colleagues (e.g. Hadamard, 1945; Poincaré, 1905), and these reports are consistent with large scale surveys of mathematicians' introspections (Burton, 2004). Most empirical studies of mathematical behaviour at this level have focused on similarities across sophisticated mathematical thinkers, often contrasting these with the behaviour of novices. Here we contend that such homogeneity at the level of overall problem solving may coexist with distinct strategies at a more detailed level; the remainder of this paper presents case studies showing different reasoning preferences in successful research postgraduates.

3. Method

Our case studies are taken from a larger study about how highly talented mathematics students evaluate conditional statements. Previous reports of this study have focused on

the importance of qualification in mathematicians' descriptions of their own certainty about mathematical inferences (Inglis, Mejia-Ramos & Simpson, 2007). However, analysis of this data in terms of semantic and syntactic reasoning has also proved illuminating. As stated in the introduction, we report just two case studies in this paper. Both participants were studying for a PhD in a prestigious UK university: Chris's thesis focussed on conformal field theory, and David was researching algebraic topology. Each participant had been highly successful in their undergraduate mathematics degree and both have now successfully completed their doctorates.

Neither participant had prior experience with the concepts of abundant and deficient numbers (the subject matter of the tasks), although of course they were well-acquainted with basic number theory concepts such as divisibility and factorisation. They were sufficiently advanced and successful that they did not (and would not be expected to) encounter any of the well-known problems with unpacking the logic of mathematical statements cited by Selden & Selden (1995; 2003) and, unlike the preservice teachers studied by Zazkis (2001), had no problem using formal mathematical terms to express divisibility relationships. Indeed, using this subject matter with these participants may provide a particularly "clean" context for studying reasoning strategies. As others have commented in relation to the *teaching* of mathematical reasoning and proof, the relevant examples in number theory are numbers and are therefore particularly familiar, accessible and unlikely to force the distraction from strategies that might result from having to simultaneously extend a conception (Selden & Selden, 2001). Here we note that much the same advantages might be said to apply in a research context.

During the research, each participant was individually interviewed for approximately an hour in a private office or seminar room. The participant was first presented with some written information about abundant, perfect and deficient numbers (shown in Figure 3, adapted from the materials used by Markowitz & Tweney, 1981). When participants indicated that they were ready to proceed, they were given a written conjecture from Figure 4 and asked to determine, with proof, whether it was true or false (proofs appear in the Appendix). After the conjecture had been solved to the participant's satisfaction, the interviewer presented a new conjecture.

Insert Figures 3, 4 and 5 here.

The procedure took the form of semi-structured clinical interviews (Ginsburg, 1981). The interviewer prompted participants when their methods were unclear and asked for clarification when necessary. In order to maximise the use of interview time, participants were given a card of examples (Figure 5) if they began to make numerical calculations of the divisibility properties of any specific integers. Each interview was recorded and transcribed for analysis. The analysis followed the quasi-judicial approach developed by Bromley (1986): transcripts were independently coded to identify each instance of example use.

Eight interviews were conducted in the larger study and we report data from two of these in the next section. Our goal in presenting these case studies is not to attempt to generalise (or overgeneralise) to all mathematics research students by claiming that the participants' behaviour is typical (Bromley, 1986; Yin, 2003). Rather, we aim to give "existence proofs" of the theoretical classification discussed in the first part of this paper by illustrating that some research students adopt largely semantic strategies and some adopt largely syntactic strategies. In doing so we will demonstrate the utility of this theoretical framework for analysing mathematical behaviour, and give an indication of the range of possible behaviour from highly successful mathematics students.

4. Case studies

Both case study participants used skills associated with syntactic and semantic reasoning relative to the representation system of proof. Both correctly manipulated algebraic expressions, and both used counterexamples. Both produced syntactic products in some cases and semantic products in some cases. However, there was a striking difference between the two regarding the amount of example-based semantic reasoning conducted in production of either of these types of product. The first interviewee, Chris, used examples for multiple purposes including developing his understanding of abundant numbers (and the relationships between them), inductive testing of conjectures using crucial experiments/generic examples, checking his generalised arguments were correct by relating them to specific examples, and (of course) generating counterexamples. During the hour-long interview he considered a total of 28 different examples, and in only one of the seven conjectures did he produce a syntactic product with no

consideration of specific examples at all. In contrast David, the second interviewee, considered only 9 examples, each of which involved a semantic product: his examples were either counterexamples or were used, not to support his reasoning, but to explain his general argument to the interviewer. In no cases did David retranslate his semantic work back into the representation system of proof. A breakdown of the number of examples used by Chris and David is given in Table 1.

Insert Table 1 here.

We present illustrations of this difference below, using conjectures 1, 2 and 6. These allow us to illustrate the participants' responses to one conjecture in which semantic products were appropriate, one in which either semantic or syntactic products were appropriate and one in which a syntactic product was appropriate.

4.1 Conjecture 1: A number is abundant if and only if it is a multiple of 6

This conjecture is false (in both directions) so giving one or more counterexamples (semantic products) is appropriate.

David

In response to this conjecture David remarked, "Well, I guess 6 is perfect, right?" He then considered $6n$ where $n \neq 1$.

D: ...as divisors you've got of course 1,2,3 and 6, and n of all these, n , $2n$, $3n$, $6n$. So the sum of all them is going to be $12n$, is that right? So, but you may have more divisors I

guess. [...] Yeah so it may be perfect I guess. I mean 6 is a multiple of 6 and that's perfect.

He confirmed that 6 thus acted as a counterexample and, queried about the only-if direction, he asked, "Have we got any examples down here?" and, when provided with the card of examples, said "So 20, that's not a multiple of 6."

Chris

Chris remarked that he would look for a counterexample, and that "perhaps 6 multiplied by a large prime would be." He suggested 6×13 , broke this down into its prime factorisation and after some calculation of the various factors found that this did not work. The interviewer offered the card of examples, and Chris observed that the first few abundant numbers "are all built up of 2's, 3's and 5's." He then considered the prime factorisations of 6, 28 and 496 (the first perfect numbers). Observing that "every abundant number must be a multiple of 6" seemed to be true for the small numbers, he continued,

C: I should probably check it for a very large multiple of 6. Well, if I was going to go very large I'd use a calculator. But, I don't know, something like 1200, or something like that. Is 6 by 2 by 2 by 2 by 2 by 5 by 5 I think. Okay, so all the factors...it's going to have a lot of factors.

The interviewer decided to save time at this point by saying that 1200 would be abundant, and Chris continued thinking, commenting that he was "just checking that there wasn't an obvious pattern." He then changed tack, formulated a syntactic argument to show that numbers of the form $n = 6q$ (where q is any natural) were abundant by listing their factors and noting that $q + 2q + 3q + 1 > 6q$ (he counted 1 as an "extra factor", missing the case in which $q = 1$). When asked about the other direction he said,

C: These first ones it's certainly true for, I chose a largish one and it appeared to be true, so yeah I'm going to try and prove that that's true. Um, I mean I could try it for 6 million or something like that.

At this point the interviewer said "Is it true for these first few?" and Chris realised that 20 is not a multiple of 6, and laughed about having not noticed this. The interviewer also asked about his earlier argument in relation to the case where $q = 1$ and Chris confirmed that this case gave $n = 6$, which is perfect.

Comparison

Where David gave single counterexamples very quickly, Chris expended considerable energy exploring the example space(s) relevant to the conjecture and using "large numbers" to conduct what appeared to be crucial experiments (in the sense of Balacheff, 1987). One could say that Chris was simply less efficient in his counterexample search, but we suggest that this behaviour indicates a different orientation, in favour of developing understanding and arguments using a semantic strategy by looking at examples.

4.2 Conjecture 2: If n is perfect, then kn is abundant for any $k \in \mathbf{N}$

This conjecture is false for $k = 1$ but true for $k \neq 1$ so a counterexample (a semantic product) is appropriate for the conjecture overall but a general proof (a syntactic product) is appropriate for a minor modification of the conjecture.

David

David saw an analogy to his previous thinking about the factors of $6n$, which he generalised to say,

D: k times each divisor of n is a divisor of kn , right? [...] Why would it be greater than?

Um, I don't know, why couldn't it be perfect?

He considered the numbers on the card as possible counterexamples, but commented that he couldn't see any and that what was needed was to find some divisors of kn that were not of the form km for m a divisor of n . He then said,

D: k 's going to be a divisor of k , so k divides kn , okay? And k may or may not be a divisor of n . [...] Er, so that proves it for any k that doesn't divide n .

He then considered the other case, suggesting a general inductive argument.

D: ...well suppose k divides n and k^2 doesn't divide n . Um, then k^2 divides kn , so that's a number that would make this strictly greater than, and then I guess you're looking at some sort of induction thing.

When asked for clarification by the interviewer he said,

D: ...you're saying let k be a natural number [...] if k doesn't divide n then the theorem's true I think. If k divides n but k^2 doesn't divide n then the theorem's true, if k^2 divides n but k^3 doesn't divide n then the theorem's true. And I think you can just carry on.

He confirmed that he thought the conjecture plausible at this stage, and when the interviewer asked about the possibility that $k = k^2 = k^3$ he laughed and said,

D: ...this has got to be false for $k = 1$, so actually I could have just said, if I'd have realized, that that's false. But possibly for k above 1 this is true, and if k is above 1 then you haven't got this problem.

Chris

Chris commented that the conjecture "seems reasonable" and began this time with a syntactic approach, introducing general notation for the divisors of n and, like David, explaining that,

C: ...summing the elements of this set [*the divisors of kn*] is going to give you k times er, this sum here [*the sum of the divisors of n*]. I'm including 1 and n in this set, and we know that the divisors add up to $2n$. So this is $2kn$. So that shows that if I can find just one more divisor of kn , then we're done, and it shows that kn is abundant. [...] So if k were...let's assume k is bigger than 1, so then k itself is a divisor so we're done. [...] Oh, so, of course, k could be equal to one of these [*the divisors of n*]. So if k did not divide n ... If the highest common factor of k and n were 1 and k was not equal to 1, then we'd be done I think.

Asked by the interviewer whether this meant that k was a “spare divider,” he tentatively agreed, and then used an example to further develop his understanding.

C: Yeah, yes. Umm, if, let's see, if k were 2. And 2 appeared in here [...] yeah there is still going to be a spare divisor here I think. So why is that? Umm, [*long pause*] so 6 was an example of a perfect number. If we had 2 or 3 in this set, 2 and 3, or rather 1, 2, 3 and 6 would be the divisors of 6. Now if I multiply this up by any number, I'm still going to get a strictly larger set of numbers although, I'll try and write down a proof.

I: So you're going to multiply that up by k ?

C: Yeah, so if k were 2 then the divisors will be 1, 2...oh dear, what am I doing? Multiplying by 2. Yes, so you'd get the 1, 2, 4, 6, 12, is a strictly larger set.

Asked about where the 1 came from, he summarised his argument (using configurations from the representation system of proof) and adjusted it appropriately:

C: My argument was multiplying each of these up by k sums to $2kn$, but 1 is not included in the set unless k is 1. So this is true for k greater than 1. Um, if k were equal to 1 then it's going to be perfect because you're left with the same number.

Comparison

Here both participants constructed a syntactic product for the minor modification of the conjecture. David worked on this using a syntactic strategy by considering completely general numbers and factors and an inductive argument. Chris also began with syntactic manipulations but broke off from this approach and used a specific example (with $k = 2$ and $n = 6$) to increase his conviction in his general argument, obtaining a new insight as he did so and incorporating this into his final general statement, phrased (and later written) using the representation system of proof. That is, Chris used a semantic strategy to support the production of a syntactic product.

4.3 Conjecture 6: If n and m are abundant, then nm is abundant

This conjecture is true so a general proof (i.e. a syntactic product) is appropriate.

David

David's approach involved the manipulation of algebraic representations of abundant numbers and their divisors. He used the definition of abundant, writing down

$$\sum_{k \in \text{div}(n)} k \sum_{k' \in \text{div}(m)} k' > 4mn \text{ and continuing:}$$

D: Yeah, I've multiplied those two inequalities. I'm wondering how, now I'm doing a very sort of blind thing like that, I'm wondering whether all these terms are going to be divisors of nm . Well, if k' is a divisor of m , and k is a divisor of n , then kk' is a divisor of nm right?

I: Yes.

D: So, so yeah, so this is greater than or equal to, sorry, so, it's probably equal is it actually? Is it equal? I don't know, anyway it's certainly, things of kk' yeah, of course, this is fundamental theorem of arithmetic isn't it? You're not going to get, you're going to get

any other, 'cos you've got 1 in there. I think that's actually equal to, but it's certainly greater than or equal to the sum of the divisors maybe k'' , k'' in divisors of mn .

I: Is that right?

D: It's certainly less than or equal to, because every single kk' of this form is some k'' , but I think k'' is you see, if k'' divides nm , so what are we saying? If k'' divides nm then, err, k'' divides n , so it can be written as $n=ck''$ and it divides m , so $n=c'k''$, so erm, so we've got a divisor [...]. Well I guess, the other thing is, do you get things more than once? Actually, that might not be right, because you might get terms more than once here, if you multiply these two sums together, erm, [*long pause*] if n and m ...I mean we could go back to our examples, but I don't particularly want to because we'd be multiplying high numbers [*laughs*].

We note David's comment about "blindly multiplying" inequalities, and his aversion to using examples. Only when asked for clarification by the interviewer did he introduce one to illustrate his general argument:

D: There might be a 3 there and a 2 there, and there might be a 2 there and a 3 there. So, we've got duplications and that would only be counted once in there, so you can't, so that's not actually true.

This example might be thought of as a sort of "semantic by-product" of an interviewer question, rather than an integral part of David's reasoning. As with all of his semantic work, it was not used further in order to support the production of a syntactic product; David did not use it to check his reasoning or to obtain further insight that he would translate back into the representation system of proof. Instead, he returned to re-examine the earlier conjectures, focusing on Conjecture 2 and observing,

D: ...what I'm going to prove is a stronger thing: if n is abundant then nm is abundant for any m in \mathbf{N} , right? And I think you can just prove [...] in the same way as Conjecture 2,

because you start off with the sum of the divisors is greater than or equal to $2n$ [...] and then you have the same process.

Chris

In contrast, Chris began with an informal plausibility argument indicating that he thought the conjecture was probably true.

C: Well, abundant numbers seem to be things which have many factors. And, I don't know, when I'm multiplying together something with many factors obviously you get even more factors.

However, he did not immediately decide to try to write a general proof. Instead he considered the given examples and a specific “big” example.

C: I think I wrote down the prime factorisation of the perfect numbers earlier, and they didn't seem to be very illuminating to me. So, I mean it was something like $2^4 \times 3 \times 5^2$, $2 \times 3 \dots$ I mean I haven't got much intuition about what the prime factorisations of these perfect number should look like. So, okay,

I: So you think it's true?

C: Umm, I'm not sure, I mean I haven't actually tried an example, because I'm too afraid to work out the prime factorisation of something that big...I suppose, well, ok, $12 \times 18 \dots$ Right [*laughs*] so umm, so the factorisation is going to be $3 \times 4 \times 3 \times 2 \times 3$ so 196 factors to $2^3 \times 3^3$ in fact.

I: Well, actually, that's going to be abundant.

C: Right, okay, good. Umm, okay, I mean it certainly looks plausible.

At this point the interview had to be drawn to a close so we did not see whether Chris would go on to try to formulate a general argument.

Comparison

Once again we notice a difference between David's primarily syntactic strategy of manipulating definitions and working with entirely general representations of numbers and their factors, and Chris' primarily semantic strategy of experimenting examples of specific numbers and their factorisations. We note in particular that even though Chris initially thought the conjecture was probably true, he did not immediately attempt to construct a general proof, but rather returned to his earlier strategy of trying to develop his intuition by exploring the relevant example spaces.

4.4 Summary

The illustrations above involved one conjecture in which semantic products were appropriate, one in which either semantic or syntactic products were appropriate and one in which a syntactic product was appropriate. In all three cases, David adopted a syntactic strategy, manipulating definitions and using standard proof frameworks such as inductive arguments in order to formulate (or at least outline) general proofs. He only considered specific examples in cases in which a semantic product was necessary (as a counterexample) or was appropriate for illustrating his general arguments for the interviewer. In the latter case, the example used could be thought of as a semantic by-product of an interviewer question; David did not use it to further inform his general reasoning. Finally, he did not appear to explore example spaces simply for the sake of developing understanding or intuition. In contrast, Chris adopted a semantic strategy, exploring the example space(s) associated with the conjectures for the sake of developing understanding or intuition (*cf.* Alcock & Weber, in press). He did this in cases in which he was looking for a counterexample and in cases in which he was looking for a general

proof, i.e. whether he thought a semantic or a syntactic product was appropriate. When he worked syntactically with generalised statements, he related these to specific examples, using these examples to check and further inform his reasoning. With one exception², these contrasting behaviours are representative of David's and Chris' responses to all of the conjectures they were presented with.

5. Discussion

In this paper we began by discussing the use of examples in mathematical understanding and reasoning, relating this to semantic and syntactic reasoning strategies relative to a representation system of proof. We then exhibited two case studies which suggest that there may be individual differences in the degree to which expert mathematicians use example-based semantic strategies. To be clear, we do not suggest either that some expert mathematicians use examples in all their mathematical activity, or that some never use examples. As would be expected at this level, David introduced and used appropriate examples at appropriate junctures, and Chris made sensible use of standard notation and structures. However, the degrees to which they invoked examples to support their reasoning were strikingly different.

Whilst indicative case studies such as these have limited generalisability (even regarding preferences of the individuals involved), we believe they provide suggestive evidence that it may be possible to distinguish primarily semantic and syntactic reasoning preferences amongst those who successfully pursue research level mathematics. There

² Both David and Chris worked syntactically on Conjecture 4, logically relating the statement in its general form to the other conjectures.

are several points worth making here. One suggestion to account for this distinction is that individuals' strategies may be topic-dependent, with some mathematical areas lending themselves more to one strategy or the other. This may be true, but in this study, along with previous studies on related issues, the interview tasks were the same for all participants. A second suggestion is that an individual's strategy may depend on their familiarity with the subject, with a natural progression from one type of strategy to another as familiarity increases. In the data reported here, neither participant was familiar with the notions of abundant or deficient numbers, and the research areas of the participants (algebraic topology and conformal field theory) do not indicate in any obvious way that this could be a deciding factor in this case. A third suggestion, which has important implications for mathematics education research, is that an individual might attempt a task with a somewhat arbitrarily chosen approach and, if this happens to be successful, be influenced in their approach to later questions, so that their overall strategy seems consistent. Whether participants are consistent in their reasoning strategy across long periods of time remains an open empirical question that would demand a longitudinal study.

If the hypothesis that distinct semantic and syntactic strategies exist at the expert level does stand, this has implications for our understanding of expert mathematical practice. In particular, it raises questions about the degree of homogeneity in successful mathematicians' behaviour. If there are, in fact, large individual differences in the behaviour of expert mathematicians, then basing our understanding of expert mathematical behaviour on studies with relatively small samples may be questionable.

Findings about individual differences also have implications for mathematics teaching, most clearly at the undergraduate level. If we are to teach students the skills needed for success in advanced mathematics, we need to know what those skills are and how they are employed by fluent mathematicians. Through studies like this we can identify some skills such as those associated with example use. For our participant Chris, for instance, the exploration of example spaces seemed to be an important part of activity in coming to understand and reason about a newly encountered concept. We do not know to what extent Chris' response is typical among expert mathematicians, but this does suggest that activities based on exploring example spaces (e.g. Watson and Mason, 2005) might develop skills that are routinely used by some successful mathematical thinkers. We also note that when thinking about his examples, Chris was aware of the relationships between these and more general representations of numbers and their factorisations; he translated back and forth between the two, using general formulations to indicate what kind of examples might be useful to look at, and using specific examples to check and to suggest refinements to his general statements. Skill in maintaining this awareness may be key to successful example-based semantic reasoning in situations in which a general statement is true so that a syntactic product is required. It is also a skill that may be worth researching further, especially in view of observations about how difficult it can be to translate an insight based on semantic reasoning back into the representation system of proof (Alcock & Weber, in press; Rowland, 2001; Weber & Alcock, in press).

Of course, if Chris is somewhat exceptional in his reasoning style, then research energy might be better devoted to study of how to develop skills more associated with syntactic reasoning, such as unpacking the logic of mathematical statements (Selden & Selden, 1995) and manipulating generalised symbolic statements. Larger studies would be necessary to provide more information on the relative prevalence of such skill use among experts, and might therefore lead to information about how to focus our attention in order to maximise the short and long term effectiveness of developments in pedagogical design.

Acknowledgements

We thank Pablo Mejia-Ramos, Dina Tirosh and Rina Zazkis for valuable comments on earlier drafts of this work.

Figure 1.

Contrasting semantic and syntactic reasoning.

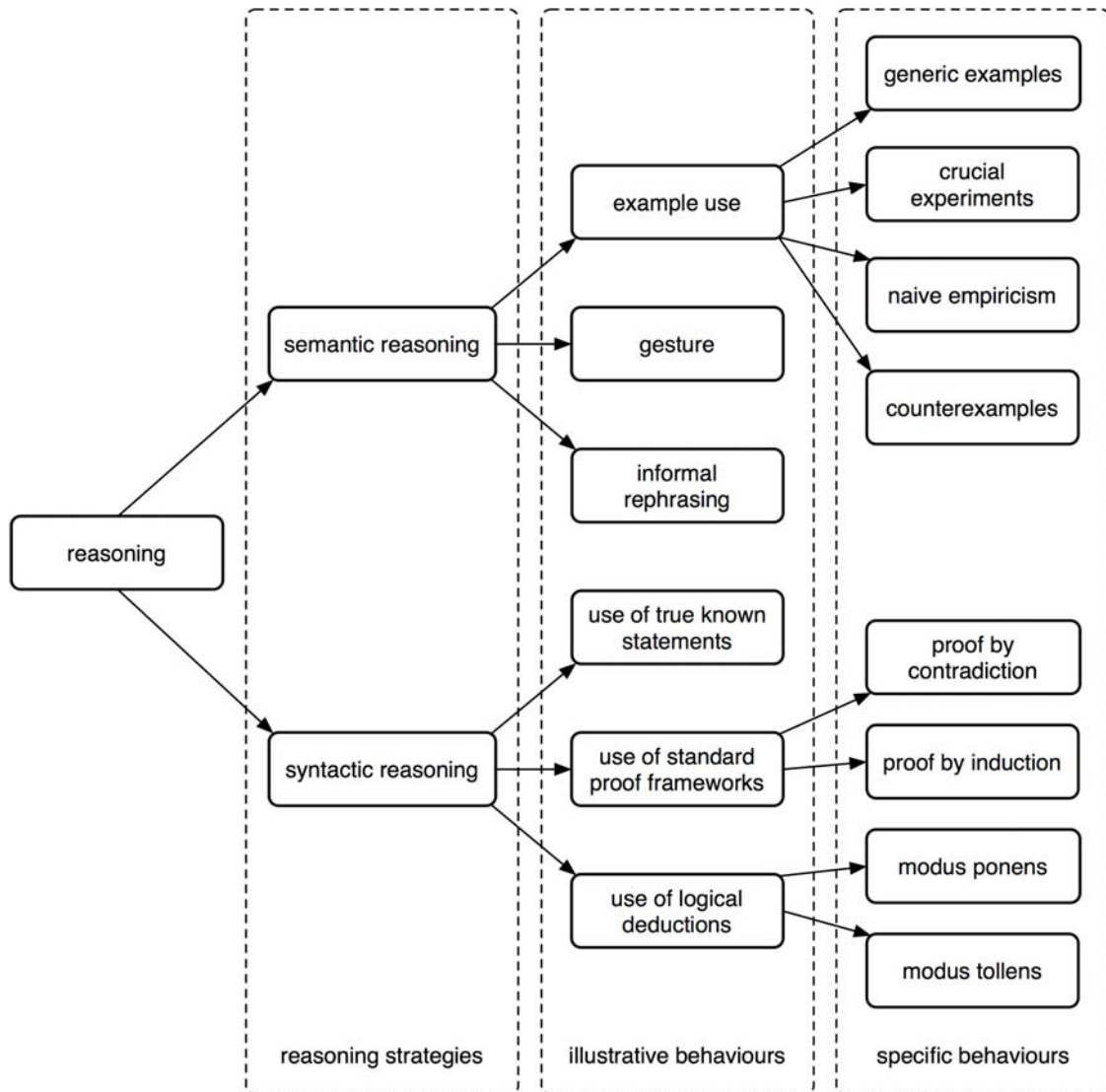


Figure 2

Semantic and syntactic reasoning strategies illustrated in terms of their use of the representation system of proof, the extra representation system and their products.

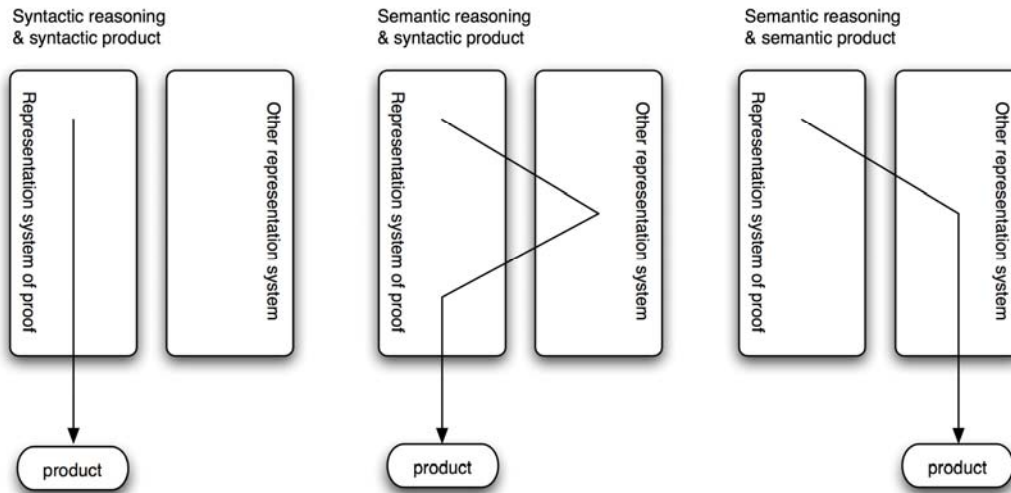


Figure 3.

The instructions given to participants.

All the numbers below should be assumed to be positive integers.

Definition. An abundant number is an integer n whose divisors add up to more than $2n$.

Definition. A perfect number is an integer n whose divisors add up to exactly $2n$.

Definition. A deficient number is an integer n whose divisors add up to less than $2n$.

Example. 12 is an abundant number, because $1 + 2 + 3 + 4 + 6 + 12 = 28$ and $28 > 2 \times 12$. However, 14 is a deficient number, because $1 + 2 + 7 + 14 = 24$, and $24 < 2 \times 14$.

Your task is to consider the following conjectures and determine, with proofs, whether they are true or false.

Figure 4.

The conjectures used in the interviews.

Conjecture 1. A number is abundant if and only if it is a multiple of 6.

Conjecture 2. If n is perfect, then kn is abundant for any $k \in \mathbf{N}$.

Conjecture 3. If p_1 and p_2 are primes, then p_1p_2 is abundant.

Conjecture 3b. If p_1 and p_2 are primes, then p_1p_2 is not abundant.³

Conjecture 4. If n is deficient, then every divisor of n is deficient.

Conjecture 5. If n and m are abundant, then $n + m$ is abundant.

Conjecture 6. If n and m are abundant, then nm is abundant.

Conjecture 7. If n is abundant, then n is not of the form p^m for some natural m and prime p .

³ Once participants had successfully solved Conjecture 3, the interviewer asked them verbally to tackle a modified version: Conjecture 3b.

Figure 5.

The examples given to participants.

Examples.

The first few abundant numbers are: 12, 18, 20, 24, 30, 36, 40, ...

The first few perfect numbers are: 6, 28, 496, 8128, ...

Table 1.

The number of examples used by each participant in each conjecture.

Conjecture	True/False	Number of examples used by Chris	Number of examples used by David
1	False	7	2
2	False	2	1
3	False	2	0
3b	True	6	0
4	True	2	1
5	False	4	2
6	True	5	2
7	True	not attempted	1
Total		28	9

Appendix.

Conjecture 1. A number is abundant iff it is a multiple of 6.

False. 20 is a counterexample to the ‘only if’ statement, it is abundant, but not a multiple of 6. 6 is a counterexample to the ‘if’ statement, it is a multiple of 6 and is perfect, not abundant. However, all multiples of 6 apart from 6 are abundant. Take $n = 6k$ for some $k \neq 1$. Then n has, at least, 1, k , $2k$, $3k$, $6k$ as distinct divisors which sum to greater than $12k = 2n$.

Conjecture 2. If n is perfect, then kn is abundant for any $k \in \mathbf{N}$.

False. If $k=1$ then kn is clearly not abundant. The statement is true provided $k \neq 1$. Suppose n is perfect with divisors d_1, d_2, \dots, d_r (i.e. $2n = d_1 + \dots + d_r$). Then kn has amongst its divisors $1, kd_1, kd_2, \dots, kd_r$, and these sum to $2kn + 1 > 2kn$. So kn is abundant.

Conjecture 3a. If p_1 and p_2 are primes, then $p_1 p_2$ is abundant.