STUDENTS' UNDERSTANDING OF A LOGICAL IMPLICATION AND ITS CONVERSE

Dietmar Küchemann & Celia Hoyles
School of Mathematics, Science and Technology
Institute of Education, University of London

In this paper we present an analysis of student responses to a written question involving logical implication and its converse in the context of simple number theory. Our sample consisted of relatively high attaining 13 year old students in randomly selected English schools. Most of the students regarded the statements representing the implication and its converse as essentially the same, and though nearly half were able to make a correct deduction on the basis of one of the statements being true, very few students used deductive reasoning as warrants for their conclusions about the truth of the statements. Students preferred to argue on the basis of empirical evidence, though many were not able to use the evidence in a mathematically appropriate way.

There is a considerable body of research over many years into argumentation in the mathematics classroom and how the processes of explanation, justification and proving can be fostered by teachers in order to develop a culture of mathematical interchange. Yackel (2001) for example took Toulmin's scheme (Toulmin, 1958) comprising conclusion, data, warrant and backing, as elaborated for mathematics education by Krummheuer (1995), to analyse interactive argumentation. In this paper, we report on student responses to a written question in which students are asked to evaluate the truth of a logical implication and its converse, and to justify their conclusions. Though a written question is far removed from the kind of interaction reported by Yackel, Toulmin's scheme provides a useful framework for discussing our results.

School students have a propensity to use inductive reasoning to validate conjectures in mathematics (e.g. Bell, 1976; Van Dormolen, 1977; Balacheff, 1988) rather than to prove them deductively. Even when students seem to understand the function of proof in the mathematics classroom (e.g. Hanna, 1989; De Villiers, 1990; Godino & Recio, 1997) and to recognise that proofs must be general, they still frequently fail to employ proof to secure beliefs in the truth of their warrants, preferring instead to rely on more data (e.g Fischbein, 1982; Vinner, 1983; Coe & Ruthven, 1994; Healy & Hoyles, 2000; Rodd, 2000; Simon, 2000). At the same time, there is some evidence from cross-sectional studies of logical thinking, that the use of 'child logic', whereby an implication and its converse are deemed as equivalent, decreases with age (O'Brien et al, 1971). In the 1970s there was considerable research on students' understanding of logical connectives, and recognition of the importance of domain knowledge (eg, Wason & Shapiro, 1971), but no systematic study of the development of logical reasoning over time in mathematical contexts, perhaps surprisingly given its importance for success in the subject.

How do students who are not taught explicitly about logical implication come to appreciate its structure from a basis (we surmise) of everyday or inductive reasoning? How do students exploit inductive reasoning (that provides a way of forming and testing conjectures, but not of proving them) to support their claims about the meaning, outcome and truth of “$p \rightarrow q$”, but also, in Toulmin’s language, begin to develop more general and explanatory warrants for their conclusions? In this paper, we begin to address these questions through an analysis of students' responses to a written question, L1, which we describe below.

Question L1 involved logical implication in the context of simple number theory. It involved the concepts odd, even, sum, and product, which would be familiar and meaningful to most of the relatively high attaining students that formed our sample. It was one of nine in a 50-minute written proof survey that was given to about 3000 students in June 2000 when they were approaching the end of Year 8 (age 13 years). A detailed scheme was devised to code the responses to all the questions, based partly on theoretical distinctions and partly on samples of pilot data. The students came from the highest attaining class (or classes) in 63 state funded secondary schools randomly selected within nine geographically diverse English regions. The same classes were also given a baseline mathematics test (using a broad selection of TIMSS items) about a month before taking the proof survey. We report here only on the students who were present on both occasions (N = 2663)[a]. All the students had followed the statutory National Curriculum in which they are encouraged to engage in informal argumentation about mathematical problems but where the nature of logical implication is not explicitly taught. Our assumption is therefore that analyses of cross-sectional responses will provide a picture of students’ understandings of the deductive process in a mathematical context, after having engaged in classroom experiences of justifying and explaining, either verbally or in writing.

RESULTS AND DISCUSSION

Question L1 was devised to investigate whether students could distinguish between a logical implication (Joe's statement) and its converse (Fred's statement). It was based on a classroom activity suggested by Watson (1995) for highlighting the directional nature of some proofs and the role of counter example. The introduction to the question is shown in Figure 1, below.

The question had four parts. In part a) students were asked "Are Joe's and Fred's statements saying the same thing?". In b) students were told that the product of two whole numbers is 1271 and to suppose that Fred is right. They were then asked to select one of three conclusions, namely that you can be sure that the sum of the two numbers is i. even or ii. odd or that iii. you can't be sure whether the sum is odd or even until you know what the two numbers are. Students were then asked to state whether c) Joe's and d) Fred's statements were true and, in each case, to explain their answer.

L1	Joe and Fred are thinking about the pair of numbers 3 and 11.
	They notice that the SUM (3 + 11) is EVEN.
	They notice that the PRODUCT (3 × 11) is ODD.
	Joe says: If the SUM of two whole numbers is EVEN, their PRODUCT is ODD.
	Fred says: If the PRODUCT of two whole numbers is ODD, their SUM is EVEN.

Figure 1: Introduction to Question L1

In part a) 13 percent of students correctly stated from the outset that Joe's and Fred's statements are not saying the same thing, with a further 15 percent changing their answer from Yes to No at some stage. The vast majority of students, 71 percent, stated that they were saying the same thing.

In b), the large and rather obscure number 1271 was chosen to discourage students from trying to find the actual numbers (31 and 41) whose product this is. The aim of the item was to see whether students could play the mathematical game of supposing a statement is true, whether it is or not, and then making a correct deduction on this basis. Nearly half the students in the sample (47 percent) chose the correct option (that the sum must be even) with another 47 percent choosing the empirical option (you can't be sure until you know what the numbers are).

In c), 36 percent of students correctly stated that Joe's statement (Sum-even → Product-odd) is false *and* supported this with one or more appropriate counter example, as in this response: "No. Because if you do 2+4=6 (even) then 2×4=8 (even) so they are both even". Of this 36 percent, most (28 percent of the total sample) gave specific counter examples (usually just one, as in the above response), but some students (8 percent of the total sample) described the counter examples in general terms (namely, that both the initial numbers are even). A general explanation of this sort represents a shift from just looking at data to looking at structure: it goes beyond showing *that* the statement is false, though this is all that the item requires, by giving some insight into *why* it is false, and so provides some insight into the warrants students held to support their conclusion.

In d), 24 percent of students correctly stated that Fred's statement (Product-odd → Sum-even) is true *and* supported this with appropriate examples, as in this response: "Yes. 3×3=9. Product = odd. 3+3=6. Sum = even". Another 9 percent of students supported their correct evaluation of Fred's statement with a general explanation, of the sort "If the product is odd, then the numbers must be odd, so the sum is even".

In contrast to c), a general explanation is needed to answer d) satisfactorily as empirical examples can not show conclusively that Fred's statement is true.

The response frequencies for the four parts of question L1 are summarised in Tables 1 - 4, below.

L1a Response	YES (incorrect)	YES changed to NO (correct)	NO (correct)	Other
Frequency (%)	71	15	13	1

Table 1: Distribution of responses to L1a (N = 2663)

L1b Response	Sum is EVEN (correct, deduction)	Can't be sure (incorrect, empirical)	Sum is ODD and other incorrect
Frequency (%)	47	47	6

Table 2: Distribution of responses to L1b (N = 2663)

L1c Response	Correct, incorrect or no decision; no valid justification or incomplete	Correct decision; valid justification, specific	Correct decision; valid justification, general
Frequency (%)	64	28	8

Table 3: Distribution of responses to L1c (Joe: Sum-even → Product-odd) (N = 2663)

L1d Response	Correct, incorrect or no decision; no valid justification	Correct decision; incomplete justification, empirical	Correct decision; valid justification, general
Frequency (%)	67	24	9

Table 4: Distribution of responses to L1d (Fred: Product-odd → Sum-even) (N = 2663)

One striking feature of the frequencies in Tables 1 - 4 is that the proportion of students who seem to use general, deductive reasoning is far higher in part b) (47 percent) than in parts c) and d) (8 percent and 9 percent respectively). This difference might be partly due to the greater complexity of items c) and d) compared to b), although the fact that our multilevel statistical analysis showed a correct response to b) to be a significant predictor of overall proof score suggests that some appreciation of the structural nature of logical implication influences approach and success to questions involving deduction. However, it also suggests that many students who use an empirical rather than a more general, deductive approach in c) and d), do so out of choice or habit rather than necessity. Of course, the point has been made that a satisfactory mathematical explanation in c) does not require a deductive approach, though it does in d).

It is also striking that the corresponding frequencies in c) and d) are very similar. Care must be taken in interpreting this, however. It does not necessarily mean that the two items are similarly demanding. Indeed a good case can be made for expecting the frequencies in the last two columns of Table 4 to be substantially higher than the corresponding frequencies in Table 3[b]. That this is not the case could be due to the order of presentation of the two items, in that many students when they reach d) will be influenced by their response to c), given the widely held belief that Joe's and Fred's statements are essentially the same. In particular, this is likely to increase the number of students who say No in d), as roughly $1^1/_2$ times as many students say No in c) as say Yes.

If the raw frequencies in c) and d) can not easily be compared, the relative frequencies are still of interest. In particular, it is notable that the students who give a correct deductive reason in d) are greatly outnumbered (by nearly 3 : 1) by those who give a correct (as far as it goes) empirical reason. This fits with the findings of other studies (e.g. Bell, 1976; Van Dormolen, 1977; Balacheff, 1988) and supports the observation made earlier that students in our sample are for more likely to favour an empirical, data-driven approach to evaluating a mathematical statement than to consider mathematical structure and to argue deductively. However, it is of interest too that the corresponding ratio in c) is similar (just over 3 : 1), bearing in mind that a general, deductive reason is not necessary here.

An examination of individual students' scripts indicates that most students' explanations in parts c) and d) start with one or more empirical example. Usually, the explanation stays at this level, though occasionally, as the frequencies discussed earlier indicate, the empirical examples are followed by a more general statement which is an empirical generalisation of the examples or, as seems to be the case below, for which the examples serve as illustrations:

d) (Fred: Product-odd → Sum-even)

9=3×3 3+3=6 21=3×7 3+7=10
I believe that it is correct as only an odd times an odd can equal an odd and if you add together the two odds you always get an even.

Students who provide appropriate empirical examples (in support of a correct conclusion) commonly only provide one empirical example, or, where they provide several, the examples are often quite similar. This suggests that rather than starting with fairly random examples with which to explore Joe's or Fred's statement, the example is carefully chosen after 'inspecting' the statement in some way. Interestingly, too, the example chosen is often quite special, for example consisting of a pair of small numbers, often identical, such as 4 and 4 in c) or 5 and 5 in d). Such numbers would seem to be the antithesis of numbers that would provide a 'crucial experiment' (Balacheff, 1988), and nor would they seem well suited for a generic example; however, one should not discount the possibility that some students thought of the numbers in one of these two ways.

There seemed to be a tendency for students who generated a lot of data to go wrong, by making incorrect conclusions from the data, perhaps because they were overwhelmed by the data or perhaps because they lacked a clear idea of the nature of the data and how it could be used to validate a claim. We have tried to describe some of the warrants students seemed to use to explain why the data they presented supported (or not) their conclusions. This analysis suggests that students may not only fail to distinguish logical implication from its converse, but also fail to appreciate how data can properly be used to support a conclusion as to whether “p → q” is true or not. For example, some students would take “p → q” to be true if they found data to support it, even if they had other data that would serve as a counter example, as in the response below. This could arise because the counter example was ignored or in the belief that 'a statement is true if it is sometimes true' (perhaps by analogy with 'a statement is false if it is sometimes false').

c) (Joe: Sum-even → Product-odd)

Yes.	4+8=12.	2+4=6	1+15=16	3+15=18
	4×8=32.	2×4=8	1×15=15	3×15=45

Other students deemed “p → q” to be false if data could be found that falsified *both* propositions in the given statement, as illustrated below, suggesting that the statement was being seen as a conjunction rather than an implication.

c) (Joe: Sum-even → Product-odd)

No. Certain numbers added together will become ODD for example: 7+2=9.
The product of those two numbers is also now even: 7×2=14.

Interestingly, the student who gave the above response switched, in answer to part d), to seeing Fred's statement in the conventional way, as is shown below. Such inconsistency was not uncommon, which has led us to wonder how far a student’s expression of a warrant for their conclusions (either in writing or verbally during an interactive episode) can be assumed to be based on the appreciation that if a warrant is to be mathematical, it must be applied consistently (even if it might not be correct from the teacher’s point of view!).

d) (Fred: Product-odd → Sum-even)

Yes. If the product of the numbers is known to be odd the sum will be even:
7×3=21 7+3=10 5×5=25 5+5=10 9×5=45 9+5=14

Some students were at what we tentatively suggest might be a transitional level[c]. Typically, they initially answered Yes to a) but changed this to No, presumably after reflecting on their answers to c) and d). Interestingly, those who changed their answer to No had a somewhat higher mean score on the other numerical/algebraic items in the survey and a higher baseline test score than those who stuck with No from the outset and than those who answered Yes (mean score of 6.7, 6.0 and 5.8 respectively for algebra and 16.6, 15.7 and 15.0 for baseline, compared with means of 6.0, out of a possible 13, and 15.3, out of a possible 22, in algebra and baseline for the sample as a whole). These students could cope with the suspension of disbelief and the deduction required in b) and were able to answer c) and d) competently using empirical data, and perhaps using large numbers as a crucial experiment to test their conclusion in d) more thoroughly, as in the response below:

d) (Fred: Product-odd → Sum-even)

Yes. Fred said the product of two whole numbers is odd, their sum is even. So:-
3×5=15 (odd) ✔ 3+5=8 (even) ✔
I tried 5 and 7, 9 and 17, and lots more, they all worked. Fred is right!

Students are more likely to adopt a general, deductive approach if they reflect carefully on the data they have generated or, indeed, if they reflect on the structure of a mathematical task before generating data. This is not something that is particularly emphasised in the English National Curriculum – perhaps even the opposite is the case, with students commonly being encouraged to formulate and test conjectures by first generating lots of data. It will be interesting to see, therefore, how the different types of warrant that students call upon to justify their conclusions, together with the consistency of their use, change over time, and to what extent the use of deductive reasoning to justify conclusions increases in our sample as the students get older.

ACKNOWLEDGEMENT

We gratefully acknowledge the support of the Economic and Social Research Council (ESRC), Project number R000237777.

---

[a] Virtually the same question was given to the same sample of students one year later in 2001.

[b] Consider, for example, someone who works on an empirical basis and whose first step is to find a pair of numbers that fits the first proposition in the statement under consideration. In d) (Fred's statement), the numbers will both be odd and so will of necessity fit the second proposition, thereby supporting the correct conclusion that Fred's statement is true. On the other hand, in c) (Joe's statement) the numbers that fit the first proposition will both be even or both be odd, so they might (both even) or might not (both odd) support the correct conclusion that Joe's statement is false.

[c] Our analysis of responses to the survey administered in 2001 will assist in this interpretation as we will be able to describe the trajectory of student responses over time.

REFERENCES

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, Teachers and Children* (pp 216-235). London: Hodder and Stoughton.

Bell, A.W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics*, 7, pp 23-40.

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal,* 20(1), pp 41-53.

De Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, pp 17-24.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics,* 3, 2, pp 9-18, 24.

Godino, J. D. and Recio, A. M. (1977). Meaning of proofs in mathematics education. *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education*, Vol 2, pp 313-320. Lahti, Finland.

Hanna, G. (1989). Proofs that prove and proofs that explain. In *Proceedings of the 13th Conference of the International Group for the Psychology of Mathematics Education*, pp 45-51. Paris.

Healy, L. & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31, 4, pp 396-428.

Krummheuer, G. (1995). The ethnology of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The Emergence of Mathematical Meaning: Interaction in Classroom Cultures*, pp 229-269. Hillsdale, NJ: Erlbaum.

O'Brien, T. C., Shapiro, B. J., and Reali, N. C. (1971). Logical thinking – language and context. *Educational Studies in Mathemtics*, 4, pp 201-219.

Rodd, M. (2000). On mathematical warrants. *Mathematical Thinking and Learning,* 3, pp 22-244.

Simon, M. A. (2000). Reconsidering mathematical validation in the classroom. In *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education*, Vol 4, pp 161-168. Hiroshima, Japan.

Toulmin, S. (1958). *The Uses of Argument.* Cambridge, UK: Cambridge University Press.

Van Dormolen, J. (1977). Learning to understand what a proof really means. *Educational Studies in Mathematics*, 8, pp 27-34.

Vinner, S. (1983). The notion of proof – some aspects of students' views at the senior high school level. *Proceedings of the 7th Conference of the International Group for the Psychology of Mathematics Education*, pp 289-294.

Wason, P.C. and Shapiro, D. (1971). Natural and contrived experience in a reasoning task. Quarterly J. Exp. Psych., 23, pp 63-71.

Watson, J. (1995). Why proof? *Mathematics in School*, 24, 3, pp 16-17.

Yackel, E. (2001). Explanation, justification and argumentation in mathematics classrooms. In *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education*, Vol 1, pp 9-23. Freudenthal Institute, Utrecht, Holland.