DOES STUDYING LOGIC IMPROVE LOGICAL REASONING?

Nina Attridge, Andrew Aberdein, Matthew Inglis

Loughborough University, Florida Institute of Technology, Loughborough University

There has long been debate over whether studying mathematics improves one's logical reasoning skills. In fact, it is even unclear whether studying logic improves one's logical reasoning skills. A previous study found no improvement in conditional reasoning behaviour in students taking a semester long course in logic. However, the reasoning task employed in that study has since been criticised, and may not be a valid measure of reasoning. Here, we investigated the development of abstract conditional reasoning skills in students taking a course in formal logic, using a more sophisticated measure. Students who had previous experience of logic improved significantly, while students with no previous experience did not improve. Our results suggest that it is possible to teach logical thinking, given a certain degree of exposure.

INTRODUCTION

Since the time of Plato (375B.C./2003) it has been assumed that people can be taught to think more logically, and in particular, that mathematics is a useful tool for doing so. This is known as the Theory of Formal Discipline (TFD) and is exemplified by the philosopher John Locke's suggestion that mathematics ought to be taught to "all those who have time and opportunity, not so much to make them mathematicians as to make them reasonable creatures" (Locke, 1706/1971, p.20). While there is some evidence that studying mathematics does indeed improve logical thinking skills, there is little evidence that studying logic itself improves one's logical thinking. The aim of the current study was to investigate the development of logical reasoning skills in undergraduate students taking a course in introductory formal logic. Before describing our study, we begin by reviewing the evidence that studying mathematics improves reasoning skills, then review previous investigations of whether studying logic can improve reasoning skills, along with the flaws in these investigations that we aimed to remedy.

The TFD was first tested systematically by Thorndike (1924), who measured children's general reasoning skills before and after one year of schooling. He reported that the subjects students studied had only a minimal influence on changes to their test scores. French, chemistry and trigonometry were associated with the largest, yet small, improvements, while other areas of mathematics (arithmetic, geometry and algebra) were associated with improvements close to zero.

However, Lehman and Nisbett (1990) found evidence that studying mathematics at university level was associated with improved conditional reasoning skills. Reasoning about conditional 'if…then' statements is a central component of logical reasoning (Inglis & Simpson, 2008), and fundamental to mathematics (Polya, 1954). Lehman and Nisbett tested US undergraduates in their first and fourth years of study on statistical

2016. In Csíkos, C., Rausch, A., & Szitányi, J. (Eds.). *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education,* Vol. 2, pp. 27–34. Szeged, Hungary: PME.

and methodological reasoning, conditional reasoning and verbal reasoning. Across their whole sample, which was formed of students from several majors, they found a correlation between number of mathematics courses taken and change in conditional reasoning behaviour (r = .31). The correlation was even stronger within the natural science majors (r = .66). This is a promising finding which suggests that conditional reasoning is an aspect of logical thinking that can be developed through mathematics education.

Conditional reasoning development was also investigated by Inglis and Simpson (2009), who compared scores in mathematics and non-mathematics undergraduates on entry to a UK university. They gave the undergraduates an abstract Conditional Inference Task which involved judging the validity of conclusions drawn from abstract conditional statements (e.g. If the letter is D then the number is 3; the number is not 3; conclusion: the letter is not D). Mathematics undergraduates performed significantly better than the comparison undergraduates, even after controlling for between-group differences in an intelligence measure. However, over the course of a year, the mathematics students improved by only 1.8%, which was not significant.

In a similar study, Attridge & Inglis (2013) investigated the development of conditional reasoning skills in mathematics and non-mathematics A-level students (A-levels are two-year post-compulsory courses in the UK, the results of which are used by universities to select incoming undergraduates). There was no difference between groups in conditional reasoning at the beginning of A-levels, but after one year the mathematics students' reasoning had significantly improved whereas the non-mathematics students' reasoning had not.

While there is evidence that studying mathematics at A-level (Attridge & Inglis, 2013) and undergraduate level (Lehman & Nisbett, 1990) is associated with improved conditional reasoning skills, there is less evidence that studying logic itself is associated with improvements in logical thinking. Next we review two studies that investigated this question and came to different conclusions.

Cheng, Holyoak, Nisbett and Oliver (1986) investigated the development of conditional reasoning skills in undergraduates taking a semester-long course in logic. The students completed four Wason Selection Tasks (with a mixture of conditional and biconditional statements and abstract and thematic content, see Figure 1 for an example) at the beginning and end of the course, which contained 40 hours of teaching, including the definition of the conditional. It seems reasonable to expect that after such training students should be fairly competent at dealing with conditional statements; it is difficult to imagine a more promising way to improve a student's logical thinking competency. Nonetheless, there was a non-significant decrease in errors of only 3%.

However, the lack of improvement that Cheng et al (1986) observed could be due to the measure they used. Since their study was conducted it has been suggested that Selection Tasks may not actually measure conditional reasoning skills, particularly

As part of your job as quality control inspector at a shirt factory, you have the task of checking fabric and washing instruction labels to make sure they are correctly paired. Fabric and washing instruction labels are sewn back to back. Your task is to make sure that all silk labels have the 'dry clean only' label on the other side.

Machine wash in warm water	Silk	Cotton	Dry clean only

You must only turn over those labels you need to check to make sure the labels are correct.

Figure 1. An example Wason Selection Task similar to those used by Cheng, Holyoak, Nisbett and Oliver (1986).

when the task is presented with a real-world context (Sperber, Cara & Girotto, 1995; Sperber & Girotto, 2002). Sperber et al. suggested that Selection Task performance is highly influenced by contextual judgments that pre-empt any reasoning. Their account, which implies that Selection Tasks do not actually measure reasoning processes, was supported across six studies (Sperber et al., 1995; Sperber & Girotto, 2002). Sperber and his colleagues showed that success rates in the task can be dramatically manipulated by altering the relevance of the content. Success in descriptive versions of the task can be increased to over 50%, in line with the success rates usually found with obligation-based contextual versions (Sperber et al., 1995).

Given that their measure may not actually reflect reasoning processes, Cheng et al's (1986) results are difficult to interpret. It may be that their participants did improve in logical reasoning, and this simply wasn't reflected in their measure. This interpretation is consistent with a similar study on teaching reasoning, in which White (1936) investigated the effect of logic training on 12-year-old boys' reasoning ability. One class spent an hour per week for three months being taught logic, including deduction, induction and syllogisms, while another class were not taught any logic. At the end of the three months the students were given a reasoning test that included, among other things, syllogism validity judgments. The class that had been taught logic scored significantly higher on the reasoning test than the control class. The authors concluded, conversely to Cheng et al. (1986), that logical thinking can be taught.

The difference in findings between White (1936) and Cheng et al (1986) may be due to the difference in the reasoning measure used, or it may be due to the difference in age between the participants in the two studies. Perhaps 12-year-olds' reasoning skills are more malleable than undergraduates' reasoning skills. To distinguish between these possibilities, we investigated reasoning development in undergraduates studying introductory formal logic, using a more sophisticated measure than a Selection Task. Undergraduate students completed an abstract conditional inference task (based on Attridge and Inglis' (2013) finding that studying mathematics was associated with improvement on an abstract conditional inference task) before and after being introduced to truth-functional logic.

METHOD

Design

The study followed a one-group pre-test/post-test design where the intervention was a course in logic and the pre-test and post-test was an abstract conditional reasoning task.

Participants

Participants (60 males, 19 females) were undergraduate students taking a course on logic at a medium-sized private research university in the South-Eastern United States. Students came from various majors, including computer science, software engineering, mechanical engineering, aerospace engineering, physics, and business. At Time 1, 79 participants completed the test and of these, 58 also completed it at Time 2.

Materials

To measure logical reasoning, we administered the abstract conditional inference task (Evans, Clibbens & Rood, 1995). In this task, participants are given a conditional rule (e.g. If the letter is M then the number is 5) along with a premise about that rule (e.g. The letter is M), followed by a conclusion derived from the rule and premise (e.g. The number is 5). The participant then deduces whether the inference to the conclusion is necessarily valid or invalid. The task contains 16 items of four inference types: Modus Ponens (MP; if *p* then *q*, *p*, therefore *q*), Denial of the Antecedent (DA; if *p* then *q*, not-*p*, therefore not-*q*), Affirmation of the Consequent (AC; if *p* then *q*, *q*, therefore *p*) and Modus Tollens (MT; if *p* then *q*, not-*q*, therefore not-*p*). The lexical content of the rules (letters and numbers) was generated randomly and the order of the problems was randomised by participant. The instructions were adapted from Evans et al.

Logic course

The course consisted of 37.5 hours of lectures over 15 weeks, covering traditional logic, symbolic logic and informal logic. The assessment consisted of 14 pop quizzes, two mid-term exams and a final exam. The participants were taught in three groups.

Procedure

Participants completed the tests in class at the beginning of the course in early January 2014, and again at the end of the course in late April 2014. Tests were completed using pen and paper under exam-style conditions.

RESULTS

Participants were split into two groups depending on whether or not they had previous experience with logic. This was determined on the basis of each participant's degree programme, and whether that programme usually involved a course with some degree of logic content prior to the one in question here. As such, this is only a proxy for prior logic experience. The prior logic group comprised students majoring in Computer Science, Software Engineering, or Electrical Engineering, all of whom should have taken Digital Logic or Discrete Mathematics before they take Logic. This allowed us to investigate the role of previous exposure to logic in any development found. One participant was removed on the basis of being an outlier in terms of change over time (scoring 16/16 at Time 1 and 5/16 at Time 2). The remaining 57 participants' data were subjected to a 2 (Time: pre-test, post-test) × 4 (Inference: MP, DA, AC, MT) × 2 (Prior Logic: yes, no) mixed ANOVA.

This revealed a main effect of inference, $F(3,165) = 41.31$, $p < .001$, $\eta_p^2 = .429$, where accuracy was higher on MP inferences than on all other inferences, all *p*s < .001, a main effect of Time $F(1,55) = 6.78$, $p = .012$, $\eta_p^2 = .110$, with higher accuracy at Time 2 ($M = 2.70$, $SD = 0.79$) than at Time 1 ($M = 2.54$, $SD = 0.63$), and no main effect of Prior Logic, $F(1,55) = 1.77$, $p = .188$, $\eta_p^2 = .031$. However, there was a significant interaction between Time and Prior Logic, $F(1,55) = 4.32$, $p = .042$, $\eta_p^2 = .073$ (see Figure 2). An independent samples t-test showed no difference in Time 1 scores between participants with ($M = 2.54$, $SD = 0.61$) and without ($M = 2.48$, $SD = 0.58$) prior logic experience, $t(55) = .37$, $p = .716$, $d = 0.1$. However, paired samples t-tests showed that in the students presumed to have studied logic previously, scores significantly improved between Time 1 ($M = 2.54$, $SD = 0.61$) and Time 2 ($M = 2.91$, $SD = 0.87$), $t(32) = 3.27$, $p = .003$, $d = 0.49$, while in the students who had not studied logic previously, scores did not significantly improve between Time 1 ($M = 2.48$, $SD = 0.58$) and Time 2 ($M = 2.52$, $SD = 0.63$), $t(23) = .41$, $p = .682$, $d = .07$. Despite this, the difference between groups at Time 2 was only marginally significant, $t(55) = 1.97$, $p = .054$, $d = 0.53$. All other interactions were non-significant.

	Average	Time 1	Time 2	Absolute change	Percentage change
MP	3.85(0.35)	3.77 (0.63)	3.93 (0.26)	+0.16 (0.65)	+4.24
DA	2.26 (0.98)	2.12 (1.46)	2.40 (1.47)	+0.28 (1.36)	+13.21
AC	1.98 (1.43)	1.91 (1.61)	2.05 (1.57)	+0.14 (1.38)	+7.33
MT	2.42 (0.87)	2.25 (1.30)	2.60 (1.22)	+0.35 (1.84)	+15.56

Table 1. Mean Conditional Inference Scores split by Time and Inference. Standard deviations in parentheses.

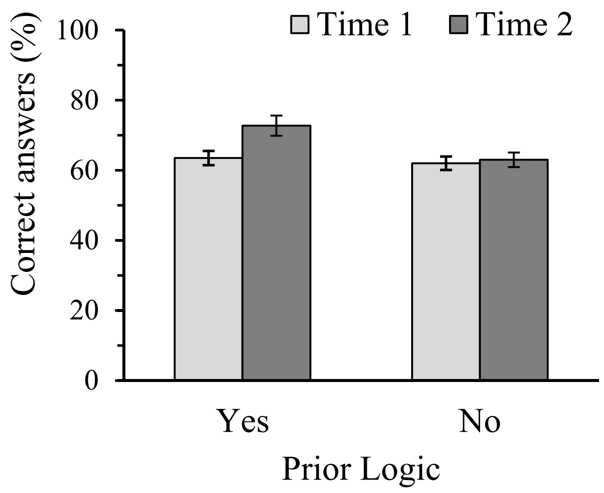

Figure 2. Interaction between Time and Prior Logic on Conditional Inference Scores. Error bars reflect ±1 standard error of the mean.

DISCUSSION

We investigated the development of conditional reasoning skills in undergraduates taking a course in logic. Overall, our results suggest that studying formal logic improves students' ability to deal with conditional statements, but only if they have had some experience with logic previously. While conditional inference scores did improve over time for the whole sample, when we examined the role of previous experience with logic, it became apparent that only those who had studied logic previously actually showed any gains in reasoning skills during the course. For those students who had not studied logic before, there was not a significant improvement in conditional inference scores over time. Interestingly, the students who had taken a logic course previously did not outperform those who had not at Time 1. This suggests that the amount of logic training the students had received previously was not sufficient to give them an advantage on our conditional inference task, but that it was sufficient to make the logic course in question more effective.

Our findings suggest that it is possible to teach logical thinking, but that a certain level of exposure may be necessary before students' skills begin to develop. We do not have data on the number of hours of previous study that participants had, but the fact that students without prior experience did not improve during the 37.5 hours of lectures involved in the current course suggests that a greater number of hours is required for development. Future research should systematically investigate the number of hours of exposure necessary for students' logical reasoning skills to improve.

It is interesting to note that the improvement we saw in conditional reasoning did not differ between the four inference types (MP, DA, AC and MT). Attridge and Inglis (2013) found that studying A level mathematics was associated with improved performance with the invalid inferences (DA and AC), and with *worse* performance on the MT inferences. In the present study, students improved to a similar extent on all of the inferences, which, perhaps unsurprisingly, suggests that teaching students logic is a more effective way to improve their logical thinking than teaching them mathematics. Nevertheless, if we compare the effect sizes for the increase in the number of correct responses, over all inference types, in each of these studies, the A level mathematics students' improvement ($d = .49$) was of a similar magnitude to that of the logic students who had some prior exposure to logic ($d = .49$) and larger than in the full logic class sample ($d = .34$). This was despite the mathematics students having no previous experience with logic and not receiving any explicit logic tuition during their A level. On the other hand, the mathematics course lasted for a full academic year, as opposed to one semester for the logic course. Although the two courses are not comparable in terms of length or student age and experience, the fact that learning mathematics appears to develop one's logical reasoning skills to a similar extent to studying formal logic is very promising for proponents of the TFD.

Our results contradict those of Cheng et al (1986) who found that a semester long course in logic was not associated with any improvements in students' reasoning performance. We suggested that the measure Cheng et al used, four selection tasks, was not an appropriate measure of reasoning, and that this may be why they failed to find an effect of tuition. Our results support this interpretation: using a more sophisticated measure of conditional reasoning we found that a similar intervention resulted in significant improvement.

One limitation of our study is that we did not compare the logic students to a control group. This means that we cannot rule out the possibility that our participants would have improved even without taking the logic course. However, this alternative interpretation seems unlikely. First, the improvement was only seen in the subset of students with prior exposure to logic. If there were a general developmental trend in reasoning skills in the undergraduate population then we would expect to see this development across the whole sample. Second, Attridge (2013) did not observe any development in conditional reasoning skills in a sample of psychology undergraduates, and Inglis and Simpson (2009) did not observe any improvement in undergraduate mathematics students. Again, if the development we observed here were due to a general developmental trend, as opposed to the logic course, we would expect to have seen improvements in both of these groups.

Another limitation is that we did not directly measure prior logic experience; we used each participant's major as a proxy for whether or not they were likely to have taken a course with some logic content previously. This means that a few students in each group could have been miscategorised. Since we split participants by major, there is also the possibility that participants in the prior logic and non-prior logic groups may have varied on SAT scores or another unmeasured variable. However, there was an overall effect of time on conditional inference scores, averaging over both groups, so these issues should not be a major cause for concern. Rather, the effect of prior logic experience should be confirmed in future studies where potential confounding variables are controlled for, and if it transpires that a factor such as SAT scores is responsible for the group difference, then this in itself would be an interesting finding.

In conclusion, our findings suggest that, contrary to previous research, it is possible to improve students' logical reasoning through instruction. Nevertheless, the level of improvement we found was comparable to that seen in A level mathematics students, who received no explicit logic tuition. This is promising for proponents of the TFD, which suggests that teaching mathematics is an effective method for developing students' logical thinking skills.

References

Attridge, N. (2013). *Advanced Mathematics and Deductive Reasoning Skills: Testing the Theory of Formal Discipline*. PhD thesis, Loughborough University.

Attridge, N. & Inglis, M. (2013). Advanced Mathematical Study and the Development of Conditional Reasoning Skills. *PLoS ONE 8*(7): e69399.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E. & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18*, 293-328.

Evans, J. St. B. T., Clibbens, J. & Rood, B. (1995). Bias in conditional inference: implications for mental models and mental logic. *The Quarterly Journal Of Experimental Psychology, 48A*, 644-670.

Inglis, M. & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics, 67,* 187-204.

Inglis, M. & Simpson, A. (2009a). Conditional inference and advanced mathematical study: Further evidence. *Educational Studies in Mathematics, 72*, 185-198.

Locke, J. (1971/1706). *Conduct of the Understanding*. New York: Burt Franklin.

Lehman, D. R. & Nisbett, R. E. (1990). A longitudinal study of the effects of undergraduate training on reasoning. *Developmental Psychology, 26*, 952-960.

Plato. (2003/375B.C). *The Republic* (D. Lee, Ed.). London: Penguin.

Polya, G. (1954). *Induction and analogy in mathematics*. New Jersey: Princeton University Press.

Sperber, D., Cara, F. & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition, 57*, 31-96.

Sperber, D. & Girotto, V. (2002). Use or misuse of the selection task? Rejoinder to Fiddick, Cosmides, and Tooby. *Cognition, 85*, 277-290.

Thorndike, E. L. (1924). Mental discipline in high school studies. *The Journal of Educational Psychology, 15*, 1-22.

White, E. E. (1936). A study of the possibility of improving habits of thought in school children by a training in logic. *British Journal of Educational Psychology, 6*, 267-273.