HEURISTIC BIASES IN MATHEMATICAL REASONING

Matthew Inglis & Adrian Simpson

Institute of Education, University of Warwick

In this paper we briefly describe the dual process account of reasoning, and explain the role of heuristic biases in human thought. Concentrating on the so-called matching bias effect, we describe a piece of research that indicates a correlation between success at advanced level mathematics and an ability to override innate and misleading heuristic biases. Implications for the importance of teaching and learning mathematics are discussed.

Over the last thirty years psychologists have been very interested in the question of rationality. Do human beings reason according to normative logical rules? Most research answers this question in the negative, and the goal has shifted to explaining why. Recently, dual process theory has attempted to answer this question by positing the existence of heuristic biases: fundamental features of the brain that steer human cognition towards certain computational strategies (Stanovich, 2003).

The goal of this paper is to compare how matching bias, one particular heuristic bias, affects successful mathematicians compared to its effect on the general well-educated population.

DUAL PROCESS ACCOUNTS OF REASONING

In recent years, cognitive psychologists have brought together several theories that attempt to explain reasoning under one heading: dual process theory. In essence, the theory puts forward the idea that there are two quite separate parts to the human brain that perform different tasks in different ways (e.g., Evans, 2003; Stanovich, 2004; Stanovich & West, 2000).

One part, System 1, operates in a quick and automatic manner, and, very roughly speaking, corresponds with intuitive thought. System 1 is thought to be a large collection of autonomous subsystems, most of which are innate to all humans, but some of which may have been acquired through experience. The subsystems' processes are subconscious in nature and only the results can be actively reflected upon.

System 2, in contrast, is slow, sequential and conscious. It is unique to humans and is believed to have evolved relatively recently. It is this part of the brain that allows complex hypothetical thought, including abstract logic. System 2 is also involved in expressing the output of System 1, and it can monitor and possibly override these intuitive responses, although, as we shall see, this does not always happen.

The evidence for the dual process account comes from two different research programs. Firstly, some persuasive evidence has come from cognitive neuroscience.

2005. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, pp. 177-184. Melbourne: PME.

Both Goel & Dolan (2003) and Houdé, Zago, Mellet, Moutier, Pineau, Mazoyer & Tzourio-Mazoyer (2000) have used fMRI scans to reveal that people do actually use different parts of their brain when exhibiting responses that are traditionally attributed to System 1 or System 2. In particular, they found that extra activity in the right prefrontal cortex appears to be critical to detecting and overriding mistakes from System 1 reasoning.

A second source of evidence has come from mainstream psychological research into human reasoning. It has been found that many people respond in an apparently irrational, non-normative fashion when given straightforward logical reasoning tasks. For example, experimenters have found that people are much more likely to endorse logical arguments as valid if the conclusions are believable. Conversely, it is much harder to correctly evaluate logically valid arguments when the conclusion is unbelievable (Evans, Barston & Pollard, 1983).

This result makes sense when analysed from a dual process perspective. System 1 has a built in 'belief bias' heuristic, that causes the reasoner to be more uncritically accepting of arguments with conclusions that they believe to be true. Evolutionarily, belief bias makes sense: why spend time and resources evaluating the validity of an argument that leads to things that you already know are true? It is only if System 2 is effectively monitoring and overriding the automatic System 1 heuristics that invalid arguments might be detected.

Stanovich (2003) reviews a number of different heuristic biases which have been identified by psychologists. These include:

- the tendency to socialise abstract problems (Tversky & Kahnemann, 1983).
- the tendency to infer intention in random events (Levinson, 1995).
- the tendency to focus attention on items that *appear* relevant, but may not actually *be* relevant (Wilson & Sperber, 2004).[1]

It should be noted here that the use of the word 'bias' here is not to be equated with 'error'. Heuristic biases need not result in processing errors, and in fact in the vast majority of cases they don't – their relative success for small resource cost is what ensures their evolutionary survival. The word 'bias', in this paper at least, should be understood in the sense of a default cognitive tendency as opposed to a processing error.

MATCHING BIAS ON THE WASON SELECTION TASK

One of the most puzzling heuristic biases found in System 1 is the so-called matching bias effect (Evans & Lynch, 1973; Evans, 1998) – roughly speaking, the effect of disproportionately concentrating on the items directly referred to in a situation. This effect was first found using the famous Wason Selection Task (Wason, 1968), but has since been extended into other arenas.

Participants in the selection task are shown a set of cards, each of which has a letter on one side and a number on the other. Four cards are then placed on a table:

The participants are given the following instructions:

> Here is a rule: "if a card has a D on one side, then it has a 3 on the other". Your task is to select all those cards, but only those cards, which you would have to turn over in order to find out whether the rule is true or false.

When the rule is "if *P* then *Q*", the logically correct answer is to pick the *P* and *Q* (not *Q*) cards – D and 7 in this example – but across a wide range of published literature only around 10% of the general population make this selection. Instead most make the 'standard mistake' of picking the *P* and *Q* cards. Indeed, Wason (1968) suggested that about 65% incorrectly select the *Q* card.

A large and controversial literature (reviewed by Manktelow, 1999) has built up over the last thirty years that attempts to explain this range of answers. One of the most robust, and counterintuitive findings in this research program is associated with the 'matching bias' effect.

Evans & Lynch (1973) discovered that when the rule was changed to "if a card has a D on one side, then it *does not have* a 3 on the other" (our emphasis), performance is dramatically improved. They found that 61% selected the correct answer with this version. With the added negative, the *P* and *Q* cards are now D and 3. Intriguingly, the 'standard' mistake (that of selecting the *P* and *Q* cards) was made by virtually no one.

After testing various different versions of the rule by rotating the negatives (i.e. as well as D⇒3 and D⇒ 3, they used D⇒3 and D⇒ 3), Evans & Lynch concluded that participants were systematically biased in favour of selecting cards mentioned in the rule, regardless of the normatively correct answer. So, regardless of the where the negatives were in the rule, the D3 selection was always prominent. This effect has become known as matching bias, and has been found to be highly robust in abstract non-deontic contexts.

The dual process account suggests that matching bias is the result of a System 1 heuristic that appears to select salient features of the environment to spend processing time on. System 1 cues certain cards as relevant: namely the ones that match the lexical content of the rule, regardless of the presence of negations (Evans, 2003). System 2 then analyses the relevant cards and outputs the result.

In the D⇒3 version of the task, System 2 needs to successfully detect the logical mistake, something it seems to do rarely. Evans & Over (2004) note that, rather than finding this System 1 mistake, many respondents to the task simply rationalise and then output their intuitive first response: D and 3. In the D⇒ 3 version, however, the same rationalisation and output process leads to the correct answer: D and 3.

The current study examined whether those who are highly successful at mathematics exhibit the same heuristic biases as have been found in the general population.

One of the major historical reasons why mathematics is a compulsory subject at school level is that, in the UK at least, it has been assumed to help to develop 'thinking skills': the ability to think rationally and logically. This belief is all-pervasive in mathematical world. For example, in their subject benchmark standards for mathematics degrees, the QAA (the UK higher education regulator) write that:

[Mathematics] graduates are rightly seen as possessing considerable skill in abstract reasoning, logical deduction and problem solving, and for this reason they find employment in a great variety of careers and professions. (QAA, 2002).

But is this really true? Adopting a dual process framework, it is clear that one major sources of failure in abstract reasoning can be attributed to System 1 heuristics such as matching bias. So, if the QAA's claim were true, one might expect those successful in studying mathematics to be less affected by System 1 heuristic biases than the general population. This research design reported in the next section of this paper attempted to test this hypothesis.

METHODOLOGY

As our mathematical sample we used first year undergraduate students from a high ranking UK university mathematics department. All these students had been highly successful in their school mathematics career, having received top grades in their pre-university examinations. In order to have a 'general population' sample, we used trainee primary school teachers, again from a high ranking UK university. Whilst these trainee teachers came from a wide variety of subject backgrounds, and were not specialising in mathematics, it should be noted that they all had been educated to degree level, and thus cannot be said to be truly representative of the population at large. It is reasonable, however, to claim that they are representative of a general, well-educated population.

We asked half of each of our samples to complete a standard D⇒3 selection task, and the other half to complete a rotated negative version (D⇒ 3). The exact wording used was the same as that given above, and is identical to Wason's (1968). The experimental sample size was large, with a total of 293 people taking part.

Each group were given approximately five minutes at the start of a lecture to tackle a questionnaire with two problems, one of which was the selection task.[2]

RESULTS

The results are shown in table 1. The first result to comment upon is that the mathematics students came out with a significantly different set of answers to the general population on the D⇒3 task (χ^2=14.1, df=4, p<0.01). More mathematicians selected the correct answer (13% v 4%) and fewer selected the 'standard mistake' (24% v 45%) than the general population. These findings are in line with previous work that has looked at mathematicians' performance on the selection task (Inglis & Simpson, 2004).

	Maths Students				General Population			
Selection	D⇒3		D⇒ 3		D⇒3		D⇒ 3	
P	25	36%	19	17%	12	23%	5	9%
P,Q	17	24%	0	0%	24	45%	0	0%
P, Q *	9	13%	80	71%	2	4%	45	78%
P,Q, P, Q	5	7%	5	4%	10	19%	3	2%
other	14	20%	8	7%	5	9%	5	9%
n	70		112		53		58	

Table 1: The results (*correct answer).

However, whereas Inglis & Simpson found that 28% of mathematics undergraduates selected the correct answer, only 13% of the current sample made this choice. This can probably be attributed to the differing experimental settings, possibly to the amount of time available to participants. Inglis & Simpson used an internet based methodology where participants were given unlimited time to make their selection, whereas in the current study thinking time was limited to at most a few minutes. It may be that a tacit understanding of the problem as a 'two-minute' or a 'ten-minute' one affects the likelihood that System 2 may find and correct a System 1 error – an interesting question still to be answered. Certainly care needs to be taken in comparing results *across* studies which use even slightly different methods.

However, since all the results in the current study were conducted with exactly the same methods, comparisons between our two groups are valid in this case.

On the D⇒3 version the mathematicians performed significantly better than the general population. However, on the D⇒ 3 version, the performances for both groups were dramatically higher, and there was no significant difference between the groups. On this version, 71% of mathematicians and 78% of the general population made the correct selection. Remarkably, nobody in either group made the 'standard mistake' of selecting the *P* and *Q* cards on the D⇒ 3 version.

To determine what effect the group that each individual came from had upon their selections, the data were analysed using a general linear models analysis, employing a saturated logistic model with a binomial error distribution. The χ^2 and p values reported are derived from this analysis.

The test version that the individual tackled was, as expected, highly significant in determining whether they selected the correct answer (χ^2=139.9, df=1, p<0.001). The group main effect was not significant (χ^2=0.47, df=1). However, the group × test interaction effect *was* significant (χ^2=5.94, df=1, p<0.02). This indicates that the effect of changing the test had a significantly *different* effect for the two groups.

Looking at the raw percentages reveals what this significant difference means. Whilst changing the test from the D⇒3 version to the D⇒ 3 version increased the success rate for the general population from 4% to 78%, it had a significantly smaller effect on the success rate for the mathematicians, which increased from 13% to 71%. That is, the mathematics sample was significantly less affected by matching bias than the general population.

DISCUSSION

Thus those who have studied mathematics to an advanced level appear to be less affected by the inbuilt System 1 heuristic known as matching bias than the general population. Note that although some researchers suggest that there is a correlation between measures of 'general intelligence' and the ability to inhibit System 1 responses (Stanovich & West, 1998), this effect cannot explain our results. There is no reason to suppose that the mathematics undergraduates are substantially higher in *general* intelligence than successful graduates from other academic disciplines. The only difference that we can be certain of is that the mathematics students have been educated in mathematics to a higher level than the trainee teachers.

What can explain this result? Does an advanced education in mathematics *cause* a reduction in the effect of matching bias, or are those who are naturally less prone to be affected by it more likely to be filtered into a mathematics degree? Additional research, in the form of a longitudinal study, would be needed to provide a satisfactorily answer to this question.

The results presented here raise a number of important questions, including:

> If an advanced education in mathematics did cause this effect, what parts of the syllabus are most important?
>
> To what level of education in mathematics must a student reach before the effect becomes significant?

Overriding misleading System 1 heuristics is a vital life skill. It is not sufficient to reject problems such as the selection task as being irrelevant laboratory experiments that have nothing to do with 'real life'. Stanovich (2003, p.53) writes:

> The issue is that, ironically, the argument that the laboratory tasks and tests are not like "real life" is becoming less and less true. "Life," in fact, is becoming more like the tests!

Those who struggle with abstract reasoning are vulnerable. They can find complex electricity bills problematic; they can struggle to follow convoluted instructions associated with tax returns or funding applications; and they can be misled by advertising. In short, they can easily fall foul of the many highly abstract rules and regulations that govern modern society. Whilst System 1 heuristics may be well adapted to the environment in which early man lived, they do not always result in maximisation of individual utility in the modern era.

A well developed ability at using System 2 to monitor and override misleading System 1 output is important. This is exactly the sort of skill that, anecdotally, mathematics is supposed to develop (QAA, 2002). Perhaps strangely, however, whether or not it actually does, appears not to have been the focus of much empirical research.

Acknowledgements

We gratefully acknowledge the financial support of the Economic and Social Research Council.

References

Evans, J. St. B. T. (1995), Relevance and reasoning. In *Perspectives on Thinking and Reasoning*, Lawrence Erlbaum, Hove, UK, chapter 7, pp. 147–171.

Evans, J. St. B. T. (1998), ‘Matching bias in conditional reasoning: Do we understand it after 25 years?’, *Thinking and Reasoning* 4(1), 45–82.

Evans, J. St. B. T. (2003), ‘In two minds: dual-process accounts of reasoning’, *Trends in Cognitive Science* 7(10), 454–459.

Evans, J. St. B. T., Barston, J. L. & Pollard, P. (1983), ‘On the conflict between logic and belief in syllogistic reasoning’, *Memory & Cognition* 11, 295–306.

Evans, J. St. B. T. & Lynch, J. S. (1973), ‘Matching bias in the selection task’, *British Journal of Psychology* 64(3), 391–397.

Evans, J. St. B. T. & Over, D. E. (2004), *If*, OUP, Oxford.

Goel, V. & Dolan, R. J. (2003), ‘Explaining modulation of reasoning by belief’, *Cognition* 87, B11–B22.

Houdé, O., Zago, L., Mellet, E., Moutier, S., Pineau, A., Mazoyer, B. & Tzourio-Mazoyer, N. (2000), ‘Shifting from the perceptual brain to the logical brain: The neural impact of cognitive inhibition training’, *Journal of Cognitive Neuroscience* 12(5), 721–728.

Inglis, M. & Simpson, A. (2004), Mathematicians and the selection task, *in* M. Johnsen Hoines & A. B. Fuglestad, eds, ‘Proceedings of the 28th International Conference on the Psychology of Mathematics Education’, Vol. 3, Bergen, Norway, pp. 89–96.

Levinson, S. C. (1995), Interactional biases in human thinking, *in* E. Goody, ed., ‘Social Intelligence and Interaction’, CUP, pp. 221–260.

Manktelow, K. I. (1999), *Reasoning and Thinking*, Psychology Press, Hove, UK.

QAA (2002), *Mathematics, statistics and operational research subject benchmark standards*, http://www.qaa.ac.uk/crntwork/benchmark/phase2/mathematics.htm

Sperber, D., Cara, F. & Girotto, V. (1995), ‘Relevance theory explains the selection task’, *Cognition* 57, 31–95.

Stanovich, K. E. (2003), The fundamental computational biases of human cognition: Heuristics that (sometimes) impair decision making and problem solving, *in* J. E. Davidson & R. J. Sternberg, eds, 'The Psychology of Problem Solving', CUP, New York, pp. 291–342.

Stanovich, K. E. (2004), *The Robot's Rebellion: Finding Meaning in the age of Darwin*, Chicago University Press, Chicago.

Stanovich, K. E. & West, R. F. (1998), 'Cognitive ability and variation in selection task performance', *Thinking and Reasoning* 4(3), 193–230.

Stanovich, K. E. & West, R. F. (2000), 'Individual differences in reasoning: Implications for the rationality debate? ', *Behavioural and Brain Sciences* 23, 645–726.

Tversky, A. & Kahneman, D. (1983), 'Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgement', *Psychological Review* 90, 293–315.

Wason, P. (1968), 'Reasoning about a rule', *Quarterly Journal of Experimental Psychology* 20, 273–281.

Wilson, D. & Sperber, D. (2004), 'Relevance theory', *UCL Working Papers in Linguistics* 14, 249–287.

[1] It has been argued that the matching bias effect is a special case of a more general relevance heuristic (Evans, 1995). Indeed, some argue that the notion of relevance is fundamental to all cognition – meaning System 1 *and* System 2, cognition (Sperber, Cara & Girotto, 1995; Wilson & Sperber, 2004).

[2] The questionnaire consisted of two parts, each of roughly equal length. The other part is not related to this research, and is not reported here.