TEACHING UNCERTAINTY AND RISK IN MATHEMATICS AND SCIENCE

Phillip Kent, Dave Pratt, Ralph Levinson, Cristina Yogui and Ramesh Kapadia
Institute of Education, University of London, United Kingdom
p.kent@ioe.ac.uk

We are investigating how secondary teachers make sense of the concept of risk, how it figures in their teaching, and what possibilities exist when a cross-curricular and technology-enhanced approach is taken. We have developed decision-making scenarios for socio-scientific topics that involve modelling with personal value systems alongside strictly quantifiable mathematical models. Precise models are limited and may be hedged around with judgements about authority and validity, whilst value judgements are generally weakly-quantifiable. Nevertheless coming to a decision requires the weighing of these diverse forms of information, each having some associated estimation (not necessarily numerical) of 'risk'. Going beyond the idea of risk in statistical theory, we are trying to understand how personal values and models influence thinking about risk and the process of decision-making, and the implications of this for classroom practice.

INTRODUCTION

In a current research project, 'Teacher's Understanding of Risk in Socio-scientific Issues'[1], we have been investigating how mathematics and science teachers make sense of the concept of risk, how the concept figures in their teaching (at upper secondary level, approximately ages 14 to 17), and what new possibilities exist for teaching where a cross-curricular and technology-enhanced approach is taken. Risk has become a growing issue for mathematics and science education in the UK, as curricula have come to include risk and its societal role, but risk is a difficult topic for teachers and the detailed questions about what and how to teach remain (for mathematics especially) largely unanswered.

This paper is concerned with how the idea of risk is used in decision-making, by considering theoretical ideas from the literature, and the results of empirical work with teachers and students. An initial assumption is that decision-making involves the coordination of different kinds of information, based on quantitative models and personal value systems and judgements. We will present a theoretically-focussed account of our ideas, in which we draw on a few empirical episodes to illustrate the argument.

MODELS OF DECISION-MAKING AND RISK, AND LIMITATIONS OF RATIONALITY

We begin by reviewing how economic and psychological theory has modelled the process of decision-making, and some ideas this work offers on the relationship between rational and intuitive thinking about risk.

Economic theory from the early 20th Century supposed the existence of a rational 'economic man', an omniscient individual that considers *all* the possible choices being faced, and systematically selects the option which maximises some specified benefit, or minimises some specified loss or harm. Research from the 1940s onwards (by Simon, Kahneman, Tversky, Slovic, etc.) questioned this supposition and proposed instead the idea of 'bounded rationality' (Simon, 1990, 1997), and 'satisficing man', the individual who chooses the option which is 'good enough', selected on the basis of a simplified (bounded) model of reality in which only a few choices and their associated factual information are relevant to the decision (Simon, 1997). The existence of this cognitive type was further demonstrated in terms of 'heuristics and biases' that have been observed consistently in hundreds of experiments (Kahneman, Slovic, & Tversky, 1982). Although satisficing is not always an effective way of making decisions, there are many everyday and professional situations in which it works well (Simon, 1997). Considered on the time scale of human evolution (1 or 2 million years), satisficing has been a very successful strategy for all but a tiny fraction of that time, as judged by the survival of the human species, faced with natural risks (hunger, extreme weather, predatory animals, etc.).

The concept of risk that is commonly used today in science and technology contexts also has its roots in the theory of decision-making (Edwards & Tversky, 1967), in the form of the 'subjective expected utility' (SEU) model: for every hazardous event there is both a likelihood of

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

the event happening, and a numerical utility (also called dis-utility) measure which expresses the impact that the hazard would have on an individual or organisation involved should it occur. The arithmetical product of the likelihood and utility is defined to be the 'risk' of the event[2]. Hazardous events usually consist of many different inter-connected hazards, and so there exists a 'total risk' for an event which is the sum of all the individual risks. The SEU model then supposes that for someone faced with a hazardous situation, it is possible for them to identify a comprehensive set of different possible courses of action, to analyse those into sets of hazard events, give values to the likelihood and dis-utility of every hazard, compute the total risk of each course of action, and conclude with an optimal decision by selecting the course of action with minimum 'total risk'.

The SEU model is based on 'economic man' and therefore, in Simon's analysis, 'the SEU model is a beautiful object deserving a prominent place in Plato's heaven of ideas. But vast difficulties make it impossible to employ it in any literal way in making actual human decisions' (Simon, 1990, p. 13). SEU models can only be used as highly abstract models of an 'over-simplified' reality, or as models for micro-problems that are tiny, very-carefully bounded parts of reality. There is also a fundamental problem of uncertainty about knowledge: in reality, it is not possible to know that you have all the knowledge that you need to know, or if it is reliable. So SEU-type models are not only limited in extent, but also subject to judgements about relevance, reliability, and authority of knowledge sources.

The difficulty with risk in modern technological society is that it involves many situations where bounded rationality by itself is insufficient because 'technological risks' are beyond bounded thinking. Failures and difficulties in thinking are well-known: to take some simpler examples, most people's difficulty to clearly judge the relative risks of accidents in travel by aeroplane or by automobile (the former being many times safer than the latter for each passenger-kilometre, yet for many people it *feels* more risky), or the relative risks of giving a young child a vaccine (which may have rare serious side-affects) against not giving the vaccine (and placing the child and others at risk of suffering a life-threatening disease). There is a conflict of interpretation not only between bounded and heuristic thinking (and humans' limited mental analytical capacity) and the complexity of technological hazards, but also heuristic thinking is rather sensitive to the communication of information. Schneier (2008) considers the interaction between individuals' use of the 'availability heuristic' and the way that a technological society relies on media for the provision of public information: there is a tendency to give greater weight to incidents that are easily remembered (available) than incidents which are hard-to-remember, and also incidents which are vivid are more likely to be available in our memories. The media thrives on offering a diet of highly vivid information; the plane crash shown on TV news is a far more vivid memory than the thousands of plane journeys where there was no crash, or the numerous fatal car accidents every day which are are 'too routine' to be of media interest. More than this, there are of course parts of the media who seek to distort information for the sake of vividness and sensation.

So, it is evident that bounded thinking is inadequate, whilst the universal rationality of the SEU model is an unworkable ideal. What then is the middle ground in which we must operate: precise models are of limited extent and may be (especially in situations involving significant risks, where there is often a lack of comprehensive scientific knowledge) hedged around with judgements about authority and validity, whilst value judgements are generally weakly-quantifiable; nevertheless coming to a decision requires the balancing of these diverse forms of information[3]. Moving this argument into education: we think it is clear that students cannot be educated to think about risk only from a heuristic basis. There is a need for a systematic, quantified analysis of some kind, but what is the appropriate place for formal models? And as part of model-based thinking and decision-making, what is the appropriate role for formal ideas about uncertainty and probability, as encountered in the mathematics curriculum?

The path we take is to consider the place of personal models in decision-making, for two reasons: (1) because we take the (constructivist) view that learning involves the modification of pre-existing personal models, rather than learning being a process of replacing learners 'wrong' thinking with models for 'right' thinking; and (2) it is critical to respect personal models because personal values (as expressions of personal priorities and ethical positions) are inextricable from making decisions, as they determine the estimation of impacts by individual decision-makers.

Another issue is that the ‘classical’ theories of Simon, Kahneman and Tversky, etc. are largely based on situations where the only technology available is paper and pencil. We suppose that technology-enhanced tools have a dual influence: they may significantly change the nature of personal models and thinking about risk, and also they offer the potential for researchers to probe more deeply into how people think (Noss & Hoyles, 1996). In accord with the formal models, we continue to interpret risk as the combination of likelihood and impact of an event. However, we do not impose a single model for risk but instead try to understand how personal values and models influence thinking about risk and the process of decision-making.

RISK AS A CROSS-CURRICULAR CONCEPT FOR MATHEMATICS AND SCIENCE

Our research has involved working with a small group of secondary teachers, recruited in pairs of mathematics and science specialists from the same school, to co-design computer-based modelling tools in which they may explore and interrogate their own knowledge of risk, and how they deal with it in their classrooms. Our aim is not to provide teachers with ready-made tools or approaches to teaching risk, but to use their experiences of working on modelling tools and tasks to develop a deeper understanding of the pedagogy of risk, which in later research can develop into tools and activities aimed at classroom use. That said, we did hope that some students might be able to try out our prototype tools, and this proved possible in one of our teacher partner schools (as described below).

Science education in the UK has done more than mathematics education to engage with risk and the socio-scientific dimension of scientific ideas. Our findings from consultations with teachers suggest that the topic of risk in the science classroom is usually handled with a focus on the ‘social’ dimension of socio-scientific issues, such as how the popular media (mis)represents scientific knowledge and practice, and this is used to underline how scientific method deals with risk and uncertainty. What seems largely absent is any quantified approach involving numerical probabilistic information or mathematical modelling of risk, which we take to be essential for understanding risk. From the mathematics side, teachers report that probability and statistics is experienced as a dull topic for students, in which it is very difficult to get beyond artificial situations for learning. Risk was introduced explicitly into the National Curriculum of England for the first time in 2008, with an explicit aim of motivating realistic applications of mathematics, but there has yet to be any detailed guidance for teachers about how and what to teach, and the social nature of risk and the ‘fuzzy’ nature of personal values can be expected to be challenging for mathematics teachers.

These findings suggest that the concept of risk should be a rich location for cross-curricular working where both mathematics and science teachers will have something to gain. We now present some of our initial experiments and results working in this area.

DEBORAH'S DILEMMA: A DECISION-MAKING SCENARIO

We have developed a prototype software environment called Deborah’s Dilemma, aimed primarily for use with teachers as a research tool to study teachers' thinking and models about risk. Users are invited to advise a fictitious person, Deborah, on whether to have an operation that could cure a spinal condition that is causing her considerable pain. The operation entails the possibility of both minor and major hazardous outcomes that need to be inferred from various sources of information that are provided in a descriptive scenario. Choosing not to have the operation entails making choices about ways of living in order to manage the ongoing pain resulting from the spinal condition, with the potential hazard in future years of spine degeneration and increasing pain.

Three software tools are offered for users to analyse the information presented, and to use their results to present a reasoned argument to Deborah. The first tool (‘Operation Outcomes’) is a probability simulator in which users model the possible consequences of having the operation. The probabilities for various complications (i.e., side effects of the surgery, ranging from minor to serious, such as ‘superbug’ infections, or death through general anaesthetic) can be derived from the given information. There are ambiguities in this information, including conflicting opinions from different experts, which are deliberately set up in order to provoke discussion and debate about trust and reliability of information. The second tool (‘Painometer’), is an attempt to give a quantified experience of Deborah's pain as she experiences it from day to day and how different activities may cause it to increase and decrease, relative to a ‘tolerable’ level. Since pain is an experience that defies objective measurement, the personal perception of pain is a potentially interesting context for probing people's personal models of risk. Users are required to decide what everyday and leisure activities Deborah should or should not do, and in what amounts, and to infer from the tool the effect on Deborah’s pain level of those activities, as expressed by the dynamically-varying level of a ‘pain-meter’.

Full discussion of these tools can be found in Pratt et al. (submitted). We focus here on the use of the third software tool, the ‘Risk mapping tool’ (Figure 1). This is a ‘coordination tool’ whose role is to support putting the two ‘sides’ of the argument together: the probabilistic, numerical results of the Operation simulator, and the more qualitative results that come from using the Painometer. This is intentional to the design of the scenario: we supposed that it is only by introducing some kind of measure of impact, and combining this with likelihoods, that the two sides of the argument can be brought together into a whole. The first version of the software – as used by the teachers in Evidence 1 below – did not have a coordination tool, and as expected we did observe struggles to coordinate information.

The use of a graphical map for decisions and their associated hazards was suggested by reading Fischhoff et al. (2006), who propose the use of ‘influence diagrams’ as a means of developing computable models out of narrative scenarios, ‘creating a computable model requires no more (and no less) than clear thinking about the precise issue that each node and link is meant to express’ (p. 137). They propose this as a way of bringing together professionals working in risk planning where there is a need for communication across many disciplines which is hard to achieve; in particular there tends to be a divide between the specialist mathematical modellers whose models (like the SEU model) give precise results about narrow problems, and other specialists who work with imprecise but broad-scale narrative scenarios. In developing a graphical mapping tool for our purposes, we decided to offer maps that are far simpler than the comprehensive language of influence diagrams, containing only ‘decision boxes’ and ‘hazard boxes’ (in which users can give any description, numerical or textual, of the likelihood and impact of the hazard). We also wondered if maps needed to be computable in order to give useful feedback to the user. We therefore decided to experiment with a simple form of map where the boxes on the screen are colour-coded according to their horizontal position within the map; users would be asked to think about this colour as a measure of risk, and so assign a ‘level of risk’ to each box simply by moving these to the appropriate part of the map. Whilst the mapping tool does enforce the association of impact and likelihood with each hazard, we did not enforce any model for how these relate to ‘level of risk’. It was exactly at this point where we hoped users would express their personal models for the situation, providing us with a window on their thinking about risk (Noss & Hoyles, 1996), as we illustrate in Evidence 2 below.

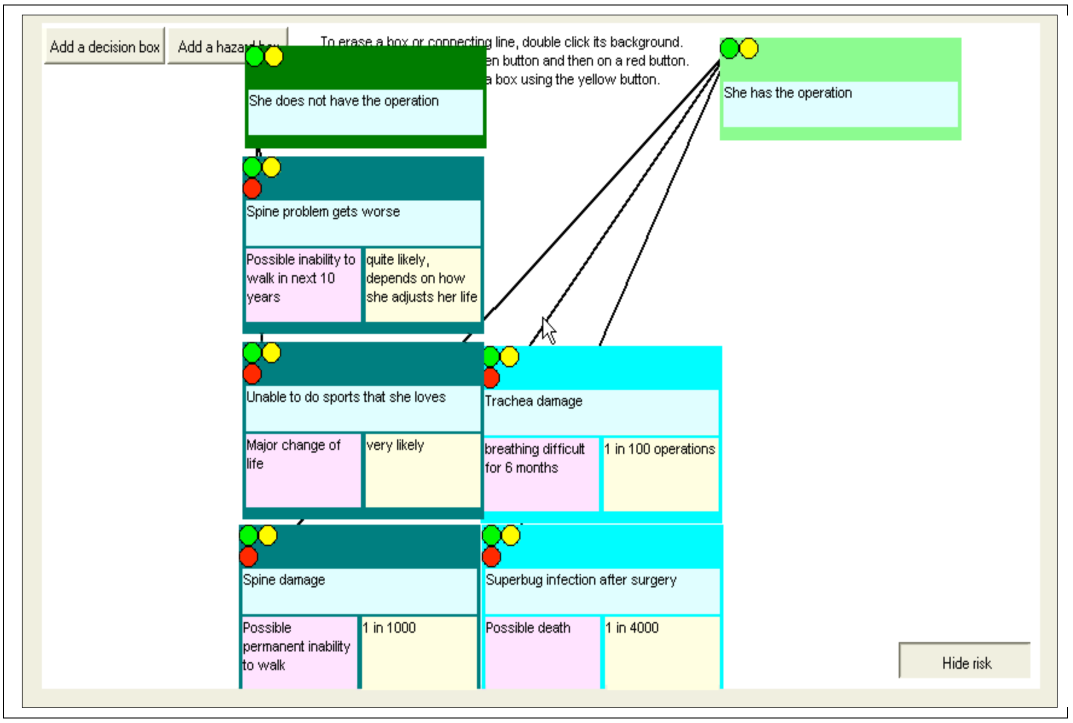
EVIDENCE OF THINKING ABOUT RISK AND PERSONAL MODELS

We present two pieces of evidence from our experiments. These are intended as highlights to illustrate the overall argument of this paper, and are not presented as typical of what teachers or students may do in Deborah’s Dilemma; large-scale evaluation of the software tools is not yet done, at the time of writing.

Evidence 1: Deborah’s Dilemma without the risk mapping tool

In an initial trial of the software, three pairs of teachers worked for about 2 hours on Deborah’s Dilemma. There was a clear and basic problem of keeping all the information in view, and we think this reinforced a tendency to flip between deciding for or against the surgery, depending on which tool was in the foreground of their thinking, and this does interfere with making a clear interpretation of their thinking. Nevertheless, a range of personal models were expressed. We also noted how the interactions between the teachers led to shifts in thinking, for example one mathematics teacher put weight initially on the numerical evidence of the probabilistic model for the surgery, but his view was modified through discussion with a science teacher who was sceptical about ‘the numbers’ and needed to take a broader view of the scenario (see Pratt et al, submitted). The data are not substantive to see how personal models of risk may have been changed though working in the scenario, however teachers did report changes in their thinking about probability more generally: for example, a biology teacher said that for him as a teacher probability meant no more than ‘something we do in teaching genetics’, so the idea of bringing numerical probability into consideration for a socio-scientific scenario was new and interesting.

It is significant to note that none of the teachers mentioned anything like the SEU model of risk. Their estimations of impact were not quantified numerically, but in terms of severity (mild/medium/severe, of operation side-affects, and of everyday pain experience). There was evidence of coordinating likelihood and impact in that whilst the impacts associated with not having surgery and living with the pain were generally of less severity than the impacts of the surgical hazards, the high probability of the former made them comparable with the low probability of the latter.



Boxes are colour-coded to show ‘risk level’ (position at left-edge = dark frame = maximum risk, at right-edge = light frame = minimum risk). Green boxes (‘She [has/does not have] the operation’) represent decisions, and Blue boxes hazards that are associated with each decision box (users can add lines to connect the boxes).

Figure 1. Risk mapping tool for Deborah’s Dilemma

Much of the richness of the teachers’ thinking came through in terms of things other than probability and impact. For example, identity was a major factor: firstly, in the sense of ‘who is Deborah?’–is doing high-impact sports central to her life? What are the other aspects of her life that are not specified in the description–does she have dependent children, elderly parents? Secondly, in the sense of ‘who are we that are making the decision?’–this led to thinking about validity of the given information, the responsibility that is implicit in making a decision (e.g., are we friends of Deborah or not?), and who would have the ability to make a reliable judgement.

Evidence 2: Deborah’s Dilemma considered with a risk mapping tool

In an evaluation with a group of twenty 15-year old students, we allowed them to work independently in pairs for 1.5 hours so as to develop their own recommendation to Deborah. Unsurprisingly, their recommendations did not have the depth of thought which the teachers showed. We then presented the mapping tool to the students as a final plenary discussion activity led by ourselves. We prepared a map of the overall decision to be made (Deborah has the operation, or does not) and inserted several hazards connected with each decision (Figure 1). The boxes were positioned arbitrarily around the screen; we turned on the ‘show risk’ button which activated the colour of the frame of each box, and asked the students whether the boxes should be moved and why. About a third of the students spoke up in response, and each showed a concern for balancing impact and likelihood.

The students seemed able to coordinate impact and likelihood into a single entity, risk. They did not use a strong quantification but expressed a sense of impact and likelihood, each having high, low and 'in between' values, and that these combined to make risk. They could verbalise it (with a little steering from the researcher), making several comments like: the box must be 'in the middle, because it is quite risky, but not serious'. They were insistent in response to questioning that the box 'Unable to do sports' should be more to the left than 'Superbug infection', and that 'Superbug' should be at the same position as 'Trachea damage', in spite of Superbug having 'possible death' as an impact, but also having a much lower likelihood than the other hazards. It is interesting to speculate what would have happened if the mapping tool had required a strong quantification. Then death may have been measured as 'infinitely high' impact, preventing any further modelling (indeed, we observed that in experiments where we offered tools that forced likelihood and impact to be numerically quantified). By being looser about the quantification process, it became possible to consider hazards whose impact was less than death to be of equal risk because they were much more likely. This feels like appropriate 'common sense' thinking: if people really regarded death as having infinite impact, then no-one would ever have a surgical operation!

CONCLUSIONS

The highlights presented here from our empirical work so far offer some interesting glimpses of personal models at work, and the potential role of a 'risk mapping tool' for supporting teachers' thinking towards developing a suitable pedagogy. The interplay between science and mathematics teachers has been beneficial in this process. By the time of the conference we will be able to draw on more extensive empirical data to analyse the personal models for risk that we have observed, and to relate these to the variety of theoretical models for decision-making which we have introduced in this paper.

NOTES

1. Funded by the Wellcome Trust [www.wellcome.ac.uk], grant number WT084895MA.
2. In a surprising amount of professional and academic literature, and almost universally in popular discussion about risk, 'risk' is equated just to the likelihood of occurrence of a hazard. We do not consider here this concept of risk, but we do consider it elsewhere (Pratt et al., submitted).
3. Simon (1997, p. 88) writes about the need for theoretical analysis of 'the real world where human behavior is *intendedly* rational, but only *boundedly* so'.

REFERENCES

Edwards, W., & Tversky, A. (Eds.) (1967). *Decision Making: Selected Readings*. Penguin Books.

Fischhoff, B., de Bruin, W., Güvenç, Ã., Caruso, D., & Brilliant, L. (2006). Analyzing disaster risks and plans: An avian flu example. *Journal of Risk and Uncertainty, 33*(1), 131-149.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge University Press.

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings*. Kluwer Academic.

Pratt, D., Ainley, J., Kent, P., Levinson, R., Yogui, C., & Kapadia, R. (submitted). Informal inferential reasoning as a contextualised risk-based enterprise. Submitted to *Mathematical Thinking and Learning*. Online: www.RISKatIOE.org.

Schneier, B. (2008). The psychology of security [unpublished draft essay]. Accessed 31 October, 2009 from www.schneier.com/essay-155.html.

Simon, H. A. (1997). *Administrative Behavior: A study of decision-making processes in administrative organization* (Fourth ed.). New York: Free Press.

Simon, H. A. (1990). *Reason in Human Affairs*. Stanford, CA: Stanford University Press.