ZDM Mathematics Education
DOI 10.1007/s11858-012-0440-z

ORIGINAL ARTICLE

A pedagogic appraisal of the Priority Heuristic

Dave Pratt · Ralph Levinson · Phillip Kent · Cristina Yogui · Ramesh Kapadia

Accepted: 8 June 2012

Abstract We have explored how science and mathematics teachers made decisions when confronted with a dilemma in which a fictitious young woman, Deborah, may choose to have an operation that might address a painful spinal condition. We sought to explore the extent to which psychological heuristic models, in particular the Priority Heuristic, might successfully describe the decision-making process of these teachers and how an analysis of the role of personal and emotional factors in shaping the decision-making process might inform pedagogical design. A novel aspect of this study is that the setting in which the decision-making process is examined contrasts sharply with those used in psychological experiments. We found that to some extent, even in this contrasting setting, the Priority Heuristic could describe these teachers' decision-making. Further analysis of the transcripts yielded some insights into limitations on scope as well the richness and complexity in how personal factors were brought to bear. We see these limitations as design opportunities for educational intervention.

Keywords Risk · Decision making · Probability · Simulation

1 Preamble

This study starts to develop an approach towards the teaching of risk in schools, based on what is known about how people think about risk and using mathematical tools that may extend and support thinking. Psychological research reports a prevalence of heuristic thinking in coming to judgments about both chance and risk. Heuristic thinking refers to the intuitive approach people employ when reasoning about situations that are intractable either because: (1) theory itself is limited; (2) the reasoner has insufficient knowledge or awareness of the potential to apply that knowledge; (3) the reasoner is not able to analyse the situation through theoretical knowledge because of practical limitations such as the availability of time or relevant tools. What is mostly emphasised in the psychological research is the fallibility of human reasoning, which we see as analogous to the many studies in mathematics education research that report misconceptions in reasoning. Our interest is not in diagnosing misconceptions but in considering how thinking about risk can be improved by the teaching of appropriate conceptions and the use of mathematical tools. Gigerenzer et al. (1999) have argued that heuristics can provide ways of reasoning that are more effective than analytical methods.

Unlike the typical psychological enquiry, which sets out to identify underlying cognitive principles in human thinking, teachers and curriculum developers are interested in learning interventions which change thinking. The long-term pedagogic aim of our project is to develop tasks and tools that shape knowledge about risk. We note that psychological research is founded upon 'clinical' methods where the problem to which subjects respond is intentionally devoid of the many features and complexities that characterize authentic risk-based decision making. Such methods make little contribution to the nature of educational intervention (they do not claim to do so). We therefore ask whether heuristic thinking remains the prevalent mode of thought in the complex settings of classrooms.

D. Pratt (✉) · R. Levinson · P. Kent · C. Yogui · R. Kapadia
Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK
e-mail: d.pratt@ioe.ac.uk
URL: http://people.ioe.ac.uk/dave_pratt

Published online: 28 June 2012

However, the research described here does not start with students but with mathematics and science teachers. In such a new and unexplored area, we decided to find out first what teachers know, not only about risk but also about its teaching and learning. Do mathematics and science teachers deploy the heuristic thinking described in psychological research? How might they be sensitized to how they themselves think about risk and what would they see as the implications for teaching risk?

More specifically, we have somewhat unconventionally designed a computer-based microworld as a probe into current knowledge with the aim of making the reasoning process more visible for the researchers. We analyse whether the thinking of the teachers as expressed through this microworld reflects that predicted by psychologists and the extent to which heuristic thinking is robust in the face of contextual complexity and richness. At the same time, we search for design principles that might inform pedagogic theory.

2 Models for making judgements about risk

One of the earliest attempts to explain human rational reasoning, in the face of choices involving monetary gains, was given by the expected value theory (Friedman & Savage, 1952). This was initially explored by Daniel Bernoulli (1738) who suggested replacing objective amounts of money by expected utilities involving weights and sums. The expected utility theory (EUT) proposes that the decision maker can choose between different options by comparing their expected utility values (Mongin, 1998). For example, consider the situation in Table 1 where decision A has two possible outcomes with given probabilities whereas decision B has a single outcome:

Under EUT, decision A has an expected utility of: $-200 \times 0.05 + 0 \times 0.95 = -10$ whereas decision B has an expected utility of -5. EUT would argue that decision B is correct as it maximises expected utility. EUT allows a *trade-off* to be made between likelihood and probability in the way each is weighed within the overall formula. In order to account for differences in choices due to personal preferences, authors such as von Neumann and Morgenstern (1947) and Savage (1954) developed the idea of subjective expected utility that combines personal utilities and probabilities.

However, despite these modifications the EUT model of decision making cannot explain much human behaviour as identified through experiments (Brandstätter, Gigerenzer and Hertwig, 2006). Alternative frameworks are offered by models that assume heuristic thinking, in which the decision maker chooses an option that uses only part of the information, perhaps focussing only on likelihood or only on impact to avoid the need for a trade-off.

An example would be the use of lexicographic rules that order outcomes and/or their probabilities according to some criterion (or aspiration level) and stop the search once an option has been located that satisfies this criterion (Katsikopoulos, 2011). Priority Heuristics combine the priority and stopping rules and unlike EUT do not assume the need for exhaustive search (Katsikopoulos and Gigerenzer, 2008). As Katsikopoulos and Gigerenzer put it, "more generally, instead of being based on axioms, the Priority Heuristic models choice by incorporating psychological principles: relative evaluation, search stopped by aspiration levels, and avoiding trade-offs" (p. 17). Gigerenzer and Brighton (2009) argue that people are capable of making more efficient and robust decisions on the basis of simple heuristics than they might be if trying to use general-purpose strategies.

Despite the clear difference in their approaches, both EUT and the Priority Heuristic are fundamentally based on cognitive decisions. Sunstein (2003) argues that emotional outcomes can override consideration of probabilities. Loewenstein et al. (2001) and Bechara et al. (1997) call attention to the role of emotions in other aspects of decision making beyond perception of the situation and of the variables involved (impacts and probabilities), arguing that they should be treated as an important component that, working together with our cognitive evaluations, influences the way we think and process information.

3 Approach

We present two analyses, each based on the same body of data. In Analysis 1, we take one of the most developed heuristics and ask whether the model suitably predicts the decisions and the decision-making process. Evidence will be presented to show that, although this heuristic is partially successful, there is a need to examine how experiential and personal factors may have shaped the decision-making process—this approach is then presented in Analysis 2. First we describe the general approach and then we report, as a case study, Peter and Erica's activity with the software tool to illustrate the issues that emerged from the two analyses (all names of teachers are anonymised).

3.1 The Priority Heuristic

Recent research (Brandstätter et al., 2006; Katsikopoulos & Gigerenzer, 2008) has focussed on trying to specify what people attend to and the information-search procedures they use when making risk-based decisions. According to

Table 1 Parameters for making a choice between decisions A or B

A	-200 with $p = 0.05$
	0 with $p = 0.95$
B	-5 with $p = 1.00$

Brandstätter et al. (2006, p. 416), in the case of losses, the Priority Heuristic describes the decision-making process as below; to simplify the explanation, we will illustrate each step in parentheses using the example in Table 1.

1. First, compare the minimum losses of the alternative decisions (0 and 5 in this example). If the difference between the two minima is at least 10 % of the maximum loss, choose the decision associated with the lesser of the two minimum losses. (In this example, this difference is <10 % of the maximum loss, 200, so move to Step 2.)
2. Otherwise, compare the probabilities of the minimum losses of the alternative decisions (0.95 and 1). If the two probabilities differ by more than 0.1, choose the decision associated with the higher probability of minimum loss. (In our example, these probabilities differ by <0.1, so we move to Step 3.)
3. Otherwise, compare the maximum losses of the alternative decisions (200 and 5) and choose the decision associated with the lower maximum loss. (So, option B is chosen, the same conclusion in this instance as given by EUT.)

(Our paraphrasing of the heuristic.)

In the literature, the Priority Heuristic has been illustrated through examples like the one in Table 1 where people are making decisions about different gambling situations with clearly specified profits or losses and probabilities. Note that there is an underlying assumption that the loss is the same as the impact or utility, which in practice may not be the case since losses could have minor or major consequences for any individual.

It is claimed that, under certain conditions, the Priority Heuristic predicts the decision most people will make *and* the decision-making process that they will undertake. One limitation, recognised by Brandstätter et al., is that the Priority Heuristic may not make the correct prediction when there is a clear difference between the expected values, $\sum_{\text{all hazards}} (\text{loss} \times \text{probability})$, of the different possible decisions. In our study, we arranged a few aspects of the scenario so that the decision appeared to us to be non-trivial. There is a recognition in the Brandstätter study that the problem representation may affect the Priority Heuristic's precision. In fact, the presentation of the scenario in our study is intentionally radically different from the above in terms of complexity and ambiguity; we return to this discussion in the conclusion.

The Priority Heuristic has received close scrutiny and several criticisms have emerged in the literature. Johnson et al. (2008) have suggested that, although the Priority Heuristic does capture some of characteristics of the decision-making process, it does not reflect observed frequent transitions between impacts and probabilities. They speculate that future models will in fact abandon the search for a single underlying heuristic. Several researchers have reported that their analyses do not support heuristic models that claim decisions are made on the basis of a trade-off (Birnbaum, 2008; Hilbig, 2008; Fiedler, 2010; Rieger & Wang, 2008).

In this paper, we explore this debate further but with a pedagogic, rather than psychological, perspective by using the Priority Heuristic to analyse the teachers' decision-making. As educationalists, our focus is on the effect of tools that might shape decision making. We seek to identify the scope of the Priority Heuristic to model a complex situation in which specially designed tools are made available. In particular, we test the Priority Heuristic in circumstances where:

1. The losses and likelihoods are not clearly specified and may even not be easily quantifiable.
2. Contradictory evidence needs to be resolved.
3. Rich descriptions of contextual matters are provided.
4. Decisions are made after careful consideration of the evidence and testing different possibilities through modelling the consequences of making different decisions.

In contrast to how the Priority Heuristic has been tested in the reported literature, we examine its relevance to risk-based decision making in a scenario which, though not authentic, shares many such characteristics. As a consequence, we hope to identify opportunities for educational intervention by observing the influence of experiential and emotional factors.

3.2 Deborah's Dilemma

Through an iterative design process, we developed a computer-based scenario, Deborah's Dilemma (DD). Mathematics and science teachers were invited to respond to the fictitious Deborah's difficulty in deciding whether to have an operation that could cure a painful spinal condition. The operation might result in a number of complications described through various, and at times deliberately conflicting, sources of information. Should Deborah choose not to have the operation, she would need to manage her pain level through changing her daily routines of work, domestic and leisure activity.

Information about Deborah's condition was set out within the software in a deliberately personal and rich way,

Table 2 A small sample of the information given to teachers through the software

On her condition	Because of the gradual decomposition of the discs one of my lower vertebrae has slipped away from the one above causing gradual narrowing of the spinal canal and increasing pain
On the effect upon her life and work	I work with a desktop computer but I can't sit for more than 5 min at a time
On the operation	This involves grafting a bone between the vertebrae which are out of alignment setting up a biological response which causes the graft to grow between the two damaged vertebrae
On the possible consequences	30 % of operations do not succeed…these were the risks if the operation did not succeed…Very serious complications including paralysis and incontinence—about 0.02 % (that is 2 in 10,000 cases)

to offer different perspectives with varying levels of authority. The information was accessed by reading text or by watching video of Deborah talking about her condition, how it affects her life and work, and what she had learned from her doctors and from personal research about the operation and its possible consequences. Table 2 gives an illustrative sample of the extensive information provided.

Two software tools accompanied the information about the condition. A probability simulator (Fig. 1) and a 'Painometer' (Fig. 2) allowed the teachers to model the possible complications to gain a sense of how often the operation might be successful, and how often complications of varying degrees of severity might occur. The teachers could draw on the information provided about Deborah and her condition in order to decide what they considered important in exploring what might happen if she chose to have the operation and what levels of likelihood should be assigned in each case. In Fig. 1, the teachers have decided on an overall success rate for the operation of 70 %; they have also chosen three possible consequential hazards and assigned probabilities based on the reports they have read about the operation. The 'patchwork' graph depicts as colour-coded squares how often, in this particular run of 1,000 operations, Deborah suffered operational failures and complications.

At the time of the data collection, the results were presented aggregated as a bar chart. As a result of observing teachers' decision-making we introduced the patchwork chart in Fig. 1 to enable the user to eyeball the data while at the same time get a sense of how the aggregated view is in fact made up of many individual cases, each of which is a future for Deborah.

Figure 1 includes only three complications as identified by the teachers but the user may enter up to 8. The underlying simulation model is quite simple though the teachers showed little interest in how it operated. If an operation is successful, there will be no complication. Complications can only occur during unsuccessful operations but an unsuccessful operation may not have a complication. The rate of complications occurring across all operations (including successful ones) is that entered by the user on the corresponding slider. (In the simulation, the program requires the probability of a complication given that the operation was not successful, which is calculated by dividing the overall probability of the complication by the probability of an unsuccessful operation. This probability is at least that entered by the user. In extreme cases,

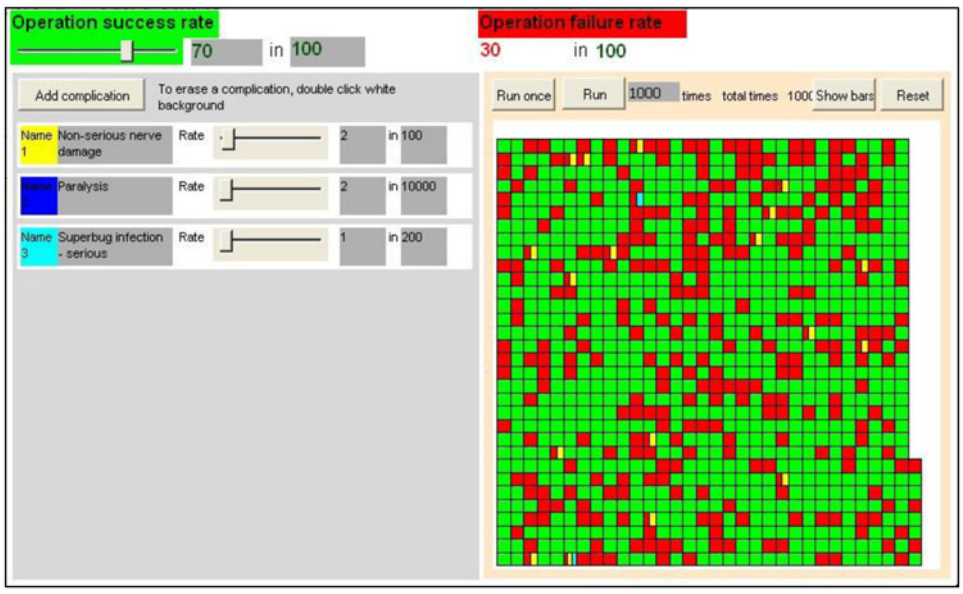

Fig. 1 The probability simulator used by the teachers to model what might happen if Deborah chose to have the operation

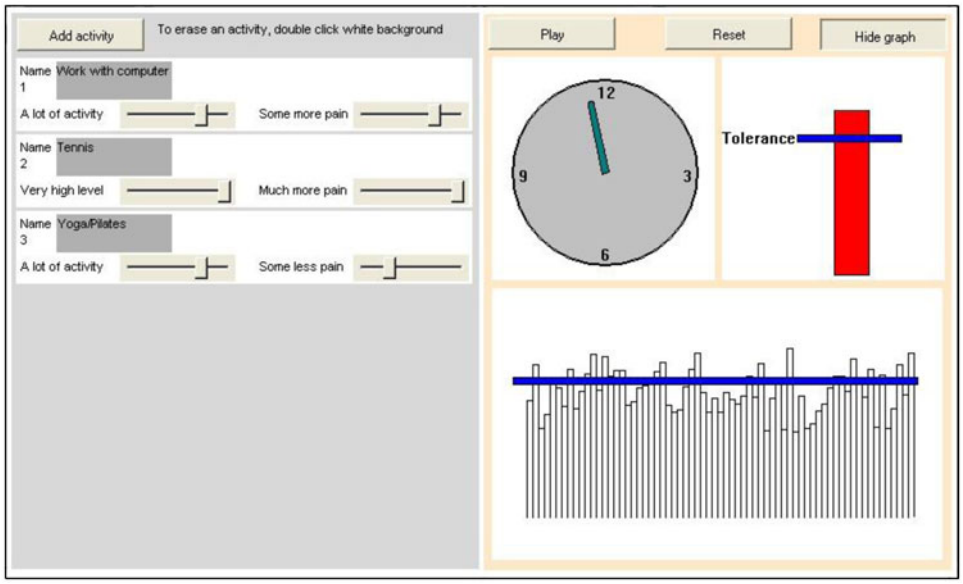

Fig. 2 The teachers used a 'Painometer' to model what might happen if Deborah did not have the operation

the probability could be >1, where the user has been inconsistent. This cannot happen with realistic figures but the program limits the probability to 1. Each instance of an unsuccessful operation may have one or more complications associated with it according to the probability calculations above, and each complication might take place independently of other complications. The outcome of each instance of an operation is independent of the outcomes of other instances of operations.)

Second, a 'Painometer' offered a quantified experience of Deborah's pain in relation to a "tolerable" level, as the pain was influenced by the activities in which Deborah engaged (Fig. 2).

In Fig. 2, the teachers have chosen three aspects of her life, which, as inferred from the descriptions, might affect Deborah's lifestyle, were she not to have the operation. The clock depicts time passing and the bar in the top right corner oscillates up and down as her pain varies as a result of these activities. (Although in everyday and scientific life, level of pain is regarded as highly subjective, in the simulation, we calculate the level of pain at each tick by adding random noise to the sum of positive and negative values associated with the fuzzy values given to each activity by the user.) The variation in pain can also be seen in the graph in the bottom right hand corner. The painometer changes height at each tick of the clock. The frequency of the activity and its consequence in terms of pain, as set by the user, are used as parameters within a stochastic model to determine the amount of pain in that tick of the clock. The teachers can also set Deborah's tolerance level as a reference point against which they can judge the amount of pain. The tolerance level has no impact on the actual level of the pain as output by the software and only functions as a visual guide, which nevertheless proved to be effective in stimulating discussion by users about pain tolerance as a personal characteristic.

3.3 Method

Three pairs of teachers (1 science and 1 mathematics from the same school in each pair) worked through DD and were asked to produce specific recommendation to Deborah; the time duration of this task was approximately 2 h (with group discussion at the end). The teachers were all experienced and taught 11–18 years olds in schools in London. We expected to hear more reflective views about their decision-making and how it relates to their teaching from experienced teachers than we might have witnessed with recently qualified teachers (in fact, this was apparent in the data). A researcher sat with each group to monitor discussion and only intervened to demonstrate relevant aspects of the software, to address any technical points and to ask questions for clarification. Under certain circumstances, such as when the teachers had reached a decision, the researcher might also have intervened in order to explore the basis of the decision in the teachers' thinking.

Data for the analysis consists of an audio transcript and a 'screen capture' video record of the teacher pair's interactions with the software. From the case study with Peter and Erica presented here, we are able to draw out issues about the validity and scope of the Priority Heuristic and the influence of emotional and contextual factors. This approach enables us to report the richness and complexity evident in the decision-making process. Of course, such an approach is limited in terms of its statistical generalisability but offers a meaningful narrative to the reader.

4 An illustrative case: Peter and Erica's activity with Deborah's Dilemma

We present in two stages the activity of Peter and Erica as they worked on task described in Sect. 3.2. The first stage represents the process by which Peter and Erica came to a decision based on their interpretation of the information given in DD. Later, we present their ongoing activity after an intervention from one of the researchers. Italicised comments refer to additions inserted into the protocol to clarify its meaning.

4.1 Stage 1: Before the researcher's intervention

Peter and Erica read the introductory information about the Dilemma and formed an initial reaction and then expressed some concern about the reliability of the data:

1 P: If I was Deborah I think I'd have the operation.
2 E: I agree—so we'll go for operation first.
3 P: One study, and does not say how many people, and just says 'reduced' pain.
4 E: We could assume the pain had gone completely. But we don't know how many people.
5 P: Now her own research…reliability, source.
6 E: Yes, that is questionable—one list from any old website you don't know, could be one person.

They went on to discuss the complications.

7 E: 1 in 1,000 of nerve root or spinal cord damage. 5 [*referring to the fifth listed complication*] is temporary and happens 1 % of the time…
8 P: Number 4 sounds scary [*referring to nerve root/spinal cord damage*]…it might mean a bit of tingling, pins and needles, which is a different level from being in a wheelchair. But you don't know, and of course they can't tell you because, when you're in the operating theatre, different things happen to what was expected.
9 E: Shall we put the other one—the 1 in 500, what was that? [*checks web page*] damage to trachea/oesophagus—only possibly permanent, but if you can't eat that is significant I would say—it would worry me! [*laughs*]
10 P: That sounds really horrible. Should happen <1 in 500 cases. Let's add that one [*returning to the software tool*]…call it 'trachea damage'.

The model that Peter and Erica eventually developed is set out in Table 3; 35 min into the investigation, Peter and Erica ran their model, starting with one case then extending it to 10, 100 and 10,010 trials. With a large number of trials, the percentage of failures will tend to reflect the probabilities inserted into the model. Nevertheless, this often triggered discussion in which the number of failures seemed to be more persuasive than the ratios.

Table 3 Complications in Peter and Erica's model for the decision to have the operation

Outcome	Likelihood as entered	Likelihood as *n* in 10,000
Operation successful	450 in 500	9,000
Nerve damage	1 in 1,000	10
Trachea/oesophagus soreness	1 in 500	20
General anaesthetic	1 in 1,000	10
Superbug infection	1 in 400	25

The first two columns are exactly as entered by Peter and Erica. The third column has been created in order to compare more easily the different likelihoods entered. The model did not necessarily include all eventualities, just those that were seen as more significant by Peter and Erica. Although 'death' is not specifically mentioned, it is seen as possibly associated with one or more of the outcomes listed as discussed in line 15

11 E: That's not bad, 17 [*failures*] out of 10,010. I'd take those odds.

Peter and Erica continued to modify and run their model, adding complications related to anaesthetic and infection from a superbug.

12 E: 9,010 successful. And anaesthetic is 1, which doesn't necessarily mean dying [*laughter*]. And 33 superbugged—slightly horrible, but they should have gone to a better hospital. So that was the biggest. Nerve damage was pretty low [*referring to the number of cases of nerve damage out of 1,000, which was 8*].
13 P: But you don't know how severe that is; it could be anything from a sore throat up to no eating.
14 E: I have to overcome my fixation with this [*laughs*]. Only one had a problem with the general anaesthetic.
15 P: I think that probably means death, or severe brain damage, something pretty awful.
16 E: And superbug can be awful. But again, out of 990, which have failed, only 49 people, which is 50 in 1,000, which is tiny.
17 P: And the rest just had the pain they had before.
18 E: If you had the operation without success, you had the uncomfortable experience, but you haven't lost much else apart from time. At least you haven't gone backwards. I think she should have it.
19 P: If I was the doctor I would still say, if she warrants it, because there are exercises and stuff—there are things like the special neck brace.
20 E: Who wants to walk around with that? No.

21 P: I might go for the exercises.
22 E: I definitely wouldn't.
23 P: The surgery might be very painful, the recovery, I bet you'd have to wear a brace, and you couldn't eat food. I think you'd be in hospital a long time.
24 E: But would you, how do you know? It may sound silly, but the scar is a consideration, especially for people like me.

Erica confirms in line 18 their original intuition (line 1) that Deborah should have the operation. At this point (49 min into the investigation), the researcher proposed that they begin to consider Deborah's lifestyle through the Painometer tool. They discussed which activities to include and what levels to set the amount of activity and consequent pain incurred.

25 E: I don't believe her, that she does that much sport. She didn't seem that upset about not being able to do sport.
26 P: It did impact on her life though.
27 E: Shall we leave it as a fair bit?
28 P: May not be higher than that, just 'more pain'.

After 56 min, Peter and Erica ran their model of Deborah's lifestyle.

29 E: Look! [*Laughter*] She's always above the tolerance, apart from once in a blue moon.
30 P: Yes…oh look though, it's painful to look at isn't it?
31 P: If it was like that you would stop doing your sport.

The researcher proposed that they experiment with different settings.

32 P: OK, take her sport, like I can't do all these things.
33 E: But even if she does, she's still above it [Sport slider is moved to zero]—she can do lots and lots of work and no sport and she is OK [pain level always under the tolerance level]
34 P: You'd want the operation immediately.

Having added the sporting activity, Peter and Erica felt even more confident that Deborah should have the operation. They continued by adding further activities to model Deborah's lifestyle, and remained convinced that having the operation was the better option.

35 P: As soon as you look at this one it would make children think she should have the operation, with the impact on your life.
36 E: We were undecided until we started looking at the pain.
37 P: Yes, because then you are thinking about what it does to your life. Every day it always hurts, and when she does sport, it always hurts when she shops. The risky bits of the surgery might not happen to her, but she knows every day 'when I go shopping it's going to hurt me'. With the surgery lot of things are short-term, even if you got worse for a while then you know the end point is going to better than you were in the first place.
38 E: When we were looking at the surgery, successful outcome, we did not really, it wasn't conclusive until we looked at the pain threshold.

Peter and Erica often referred to the problem through the eyes of their students in school (see line 35). Perhaps this is a natural way of thinking for experienced teachers, constantly looking for resources to support their teaching, or perhaps it is because the teachers knew about our longer-term pedagogic goal. Either way, we valued these small insights into their thinking about teaching and learning. Lines 37–38 give some indication of what influences Peter and Erica in coming to this conclusion as do their later comment (lines 39–40):

39 P: It was very good as a tool for getting the idea of what the perceived risk is, from the surgery point of view, which was very clinical, these are numbers, studies have shown, research has shown, there is not much to say about it, but then when you looked at her real life, how the condition affects her, the impact of that is massive.
40 E: Yeah, you forget about all the numbers and think, "Bloody hell!"

4.2 Stage 2: After the researcher's intervention

After 85 min of the investigation, the researcher wished to probe Peter and Erica's basis for their position—wondering to what extent it was sensitive to the parameters in the problem. Peter and Erica seemed to have reached a secure position and could be challenged in this way (as described in Sect. 3.3). He asked Peter and Erica to consider how far the probabilities would need to change for them to reverse their decision that Deborah should have the operation. As a result, Peter and Erica reviewed the complications by comparing again their entries into the model with the original text and, after running the model 10,000 times, found complications on 50 occasions, i.e. a worse position than prior to the operation on 1 in 200 occasions.

Noting that Peter and Erica were now aggregating the complications as 'being in a worse situation', the researcher pressed by asking what they would think if all 50 occasions involved death or impairment.

41 E: I'm slightly—if I was Deborah, and there was me dying, that would be better than being still alive with something horrible—do you know what I mean, it sounds stupid, but if I'm dead I don't care,

but if I'm alive and feeling pain, obviously it depends—another thing is what her family situation was, if she's got young children, with a 1 in 200 chance, you'd rather be there for your kids, whereas being by yourself, you know you might have a slightly…I think I probably wouldn't, I dunno, I'm a bit of a…

After 100 min, Peter and Erica reviewed their position.

42 E: Oh, but she's still got to live with that pain every day, I'd still go for…I don't know if I'd change my mind…

43 P: She's got a 1 in 200 chance of being worse off.

44 E: But she's in pain for most of her life.

45 P: That is partly under her control; she could stop sport for example.

46 E: Yeah, I think I'm changing my mind, but she couldn't stop her work; she could stop driving, but she wouldn't be able to carry things. Oh, we should just have stopped when we were happy! [*laughter*]

47 P: She could change her job. Probably she's been through some of those thoughts already. She didn't go straight to the doctor. She's lived with it quite a long time.

The intervention apparently led to Peter and Erica being less confident about what decision Deborah should make. They finally wrote:

48 P/E: [*writing into the computer*]…she can to a certain degree control the pain by not doing certain activities like sport but this lowers her quality of life. If she has the operation, there is a 1 in 200 chance of her having horrible complications plus there are other alternatives with the exercises and the neck brace. Her personal home life would also be a significant factor, depending on children, etc. or if she is a carer…

5 Analysis 1: The effectiveness of the Priority Heuristic as a predictive tool

Brandstätter et al. (2006) claim that the Priority Heuristic not only predicts the decision but also describes the decision-making process. We examine both of these claims.

5.1 Method of Analysis 1

In Analysis 1, we examined the entries the teachers made when using the probability and painometer modelling tools in order to establish which hazards they regarded as significant and what likelihood of occurrence they attributed to them. We consulted the transcripts in order to check that the entries reflected the ongoing corresponding discussion. The values that we extracted from the data were independently verified by a second researcher from the team.

5.2 Applying the Priority Heuristic to Peter and Erica's activity

Table 3 shows the possible outcomes and their likelihoods as entered by Peter and Erica into their model of the decision to have the operation.

In modelling Deborah's lifestyle, Peter and Erica entered sport, work and shopping as three activities that impacted on her pain level. They formed the view that, with these activities in place, Deborah would suffer almost constant above-threshold pain.

Following the Priority Heuristic, as set out above, the minimum loss that could be incurred is zero for a successful operation and constant above-threshold pain if no operation. The lower loss is zero and so the Priority Heuristic predicts an initial decision that Deborah should have the operation. However, it is unclear whether the difference between zero loss and constant above-threshold pain is more than 10 % of the loss incurred by death or paralysis. If not, then the Priority Heuristic may predict a different decision, based on the probabilities of the minimum losses, 0.9 (for the successful operation) and 1, or perhaps slightly below 1, for the constant ongoing pain. The difference between these two probabilities is <0.1 and so the Priority Heuristic refers next to the maximum loss. This is presumably death or paralysis versus constant above-threshold pain and so the prediction is a decision not to have the operation.

In conclusion, the Priority Heuristic makes the correct prediction for Stage 1 provided constant above-threshold pain is regarded as an order of magnitude below the loss associated with death or paralysis, which is perhaps reasonable, though not entirely clear, given Erica's view in line 41.

After the researcher's intervention, Peter and Erica wavered towards a decision not to have the operation. An explanation can be sought by examining the transcript, and at the same time it is also possible to evaluate the claim that the Priority Heuristic describes the decision-making process (as well as predicting the actual decision).

Although Peter and Erica took an early view that Deborah should have the operation (lines 1–2), they had not yet assimilated all of the information generated by the software, so we take continued activity as part of the process of reaching a point where a decision might be made (activity such as making sense of the likelihoods and judging the

severity of the complications in lines 8–10, and such as deciding which activities in Deborah's life were significant and how they should be measured in lines 25–28).

The heuristic indicates that the decision-making process will place first priority on perceived losses. Certainly there is substantial focus on the harms that might result from complications arising out of the operation (lines 8–10; 12–18) and in Deborah's lifestyle (lines 25–28). However, there is also considerable discussion about likelihoods and these often take place alongside discussion of losses. More pertinently, Peter and Erica make specific reference to losses when describing how they were making their decision. Thus, in line 18, Erica explicitly articulates how her focus is on the possible zero loss outcome of having the operation and uses that fact to argue for Deborah having the operation. Similarly in lines 35–38, they clarify that the operation became increasingly the better option when they considered the constant pain of not having the operation, presumably in comparison to the possible zero loss when having the operation. Perhaps line 39 captures the sentiment when Erica refers to forgetting about the numbers or probabilities. These articulations strengthen the notion that in the end, after assimilating all of the information through reading, discussion and modelling, Peter and Erica did in Stage 1 seek to minimise the minimum loss, in accordance with the Priority Heuristic.

After the intervention, Peter and Erica were encouraged to consider that extreme complications were rather likely. According to the Priority Heuristic, the decision should remain the same since the minimum losses were not affected by this re-evaluation. So, why might Peter and Erica show signs of changing their mind? One interpretation is that the intervention focussed thinking on likelihoods. Another is that in fact the intervention called attention to the maximum losses (death or paralysis) and so triggered a range of emotions relating to possible scenarios such as Deborah having children or being a carer. Either way, it seems the intervention had the effect of artificially pushing Peter and Erica into the later steps of the Priority Heuristic, resulting directly or indirectly in the consideration of maximum losses and a decision not to have the operation.

5.3 Applying the Priority Heuristic to the activity of the other teachers

We report below briefly on the two other pairs of teachers.

Table 4 shows the entries made by Linda and Adrian. When discussing not having the operation, they envisaged a manageable level of pain but advised, "If the condition worsens then surgery is not a prohibitively dangerous option".

Linda and Adrian decided even before the modelling activity took place that, if deciding for themselves, they would have the operation. This was confirmed after modelling. However, they took a different view if they were making a decision on Deborah's behalf. In these circumstances, they decided not to have the operation, "playing safe; trusting the spine specialist".

Manageable pain might be seen as rather <10 % compared to the impact of death and so the Priority Heuristic would predict having the operation, which is consistent with Linda and Adrian. However, when making a decision on behalf of Deborah, Linda and Adrian switched to say they would not advise to have the operation, consistent with not stopping until Step 3 in the Priority Heuristic.

How might there be a difference in the operation of the Priority Heuristic in these two situations: making a decision for yourself versus giving advice to Deborah? Perhaps they are different. If we assume that all decisions are ultimately made selfishly (and this is a controversial assumption), the impact that should be entered is not the impact on Deborah but the impact on the decision maker. The decision maker may suffer extreme anxiety as the result of giving advice that leads to severe consequences but this impact is presumably less than death. The difference between the impacts of making either decision for the decision maker on behalf of another may be seen as somewhat less than the impact on Deborah herself and could potentially result in a different decision.

The third pair of teachers was Tim and Neil (Table 5).

Table 4 Complications in Linda and Adrian's model for the decision to have the operation

Outcome	Likelihood as entered	Likelihood as *n* in 10,000
Operation successful	95 in 100	9,500
Nerve damage	1 in 1,000	10
Tube damage	1 in 500	20
Anaesthetic complication	1 in 1,000	10
Superbug infection	1 in 4,000	2.5
Temporary hoarseness	1 in 100	100
Infection/fluid leak	1 in 100	100

Table 5 Complications in Tim and Neil's model for the decision to have the operation

Outcome	Likelihood as entered	Likelihood as n in 10,000
Operation successful	95 out of 100	9,500
Trachea damage	1 in 500	20
Hoarseness	1 in 100	100
Nerve root	1 in 1,000	10
Infection	1 in 100	100
Anaesthetic	1 in 10,000	1

Tim and Neil were very clear from the outset that they would have the operation, a view that scarcely wavered throughout the investigation. Only in extreme circumstances would they have considered not having the operation. The Priority Heuristic would predict a decision not to have the operation and so in this case fails to make the correct prediction.

The above analysis has demonstrated that the Priority Heuristic correctly predicted the decisions and to some extent the decision-making process of two out of three pairs of teachers, though some issues were raised about the scope of the heuristic in both cases. In order to understand better the limitations of the Priority Heuristic we now examine how emotional and experiential factors might have influenced the teachers' decision making in ways not reconcilable by the Priority Heuristic.

6 Analysis 2: The influence of emotional and experiential factors

The above analysis provides some reasons to be concerned about the scope of the Priority Heuristic to model the decision-making process. In Analysis 2, we focus on the influence of emotional and experiential factors as evident in the data.

6.1 Method of Analysis 2

The transcripts were openly coded in relation to interactions with the software. By reading the transcripts and comparing them to the recorded screen activity, one researcher would identify initial themes. These themes were discussed by the research team, sometimes challenged, and potential new themes or sub-themes were proposed. These were then validated by another researcher and discrepancies resolved. Comparisons across thematised transcripts enabled the identification of common themes and contradictions (Flick, 2006). Subsequently phrases from the transcripts were grouped according to these themes and sub-themes (Wengraf, 2001). We report below on four themes.

6.2 Factors that shaped the decision-making process

Analysis 1 showed how the Priority Heuristic had limited success in predicting decisions and the decision-making process of the three pairs of teachers. In Analysis 2, we consider how emotional and experiential factors shaped the decision-making process and may not have been translated into the values used in the Priority Heuristic mechanisms.

The above themes have been reported elsewhere (Pratt et al., 2011; Levinson et al., 2011). We therefore offer in this text illustrative material only through Peter and Erica's account.

6.2.1 Trustworthiness and authority

The teachers often expressed a cautious response to the data provided. For example, in lines 5–6, Peter and Erica refer to an Internet study as problematic, since it is not clear how trustworthy is such information. Similar, often even more pronounced, concerns were expressed by all three pairs of teachers. Of course, in authentic decision making, there is almost always a concern about the reliability of the data. It is difficult to see how a mechanism such as the Priority Heuristic could ever reflect such concerns; it can capture judgments about probabilities and impacts but cannot manage uncertainties about the information on which estimations of probability and impact might be based.

6.2.2 Interpretation

We intentionally designed DD to incorporate partial and sometimes conflicting data since we felt this might trigger responses closer to those we might observe in natural settings. It was common for all three pairs of teachers to agonize over how to resolve conflicting evidence such as when trying to decide what probability to enter into the model given differing values from various doctors and consultants.

In lines 3–4, Peter and Erica express concern that the data is from one study only and that it is not clear how many people benefitted from reduced pain. They therefore question the reliability of the data. As above when discussing trustworthiness, it is difficult to see how concerns about reliability can be integrated into a model such as the Priority Heuristic.

A different type of interpretation issue is expressed by Peter in line 39, when he expresses how interpretation of frequency information is somehow different when carried out by a clinician compared to a personally involved subject. Erica's response in line 40 is memorable.

The overriding influence in decision making of impact over likelihood perhaps explains one aspect of the failure of EUT to reflect actual decision making and is reflected in the Priority Heuristic by prioritizing impact at Step 1 over likelihood at Step 2.

6.2.3 Impact versus probability

It might have been expected that when the probabilities of the consequences of the operation were so low, they would be ignored. In fact, all three pairs of teachers discussed probabilities throughout the decision-making process,

especially when modeling what might happen if deciding to have the operation. Nevertheless, probabilities seemed not to have much influence over the decision making. Rather than focusing on probabilities, Peter tends to consider preventative measures to avoid the operation (lines 19, 21, 31, 32, 47) and Erica looks to avoid pain or embarrassment (9, 38, 42, 44) and outcomes that might impact on appearance (24). This culminates in their explanation for their final decision (39) in terms of how the condition affects her and its massive impact on her life (40). After the researcher's intervention, probabilities seem to have more influence as shown in line 48. Perhaps the effect of the intervention was to reinforce consideration of probabilities resulting in a change of mind by Peter and Erica.

This experience is consistent with the Priority Heuristic, which places most emphasis on impacts in Step 1. The intervention seemed to artificially push Peter and Erica through Step 1 so that likelihoods became the focus as in Step 2.

6.2.4 Empathy and experience

One of the most striking features of all three transcripts was the degree of empathy that the teachers articulated either in support of Deborah or in imagining themselves in a similar scenario. This empathy would often trigger personal experiences, which seems to be important in attempting to weigh the issue especially in terms of its significance. The teachers would refer to anecdotal material they had read or seen on television. They would put themselves in Deborah's position and recognize the impact it could have on their own families and children.

There are several specific instances of empathizing by Peter and Erica. In line 14, Erica recognizes that she is perhaps overly concerned about how the operation might affect her eating; in line 20, she worries about wearing a neck brace and in line 24, the appearance of a scar is a concern for her.

The Dilemma, although contrived, seemed to spark reactions akin to how people might respond in authentic settings; we observed instances of the teachers imagining themselves in such a situation or wanting to know more information without which they felt unable to make the decision (lines 41 and 48).

Personal experience and empathy seem to be significant in attempting to judge the size of the impact. Perhaps though it is exactly this emotional effort that positions impact to the fore, ahead of the rather cold numerical figures expressed as probabilities or frequencies. This imbalance, as it might be seen through the eyes of the EUT model, is expressed in the Priority Heuristic by placing impacts in Step 1 ahead of likelihoods in Step 2.

7 Deliberations about the scope of the Priority Heuristic

Broadly speaking, much of the evidence from these three pairs of teachers is consistent with the Priority Heuristic. There is strong evidence that impact plays a more significant role than probabilities with the latter becoming emphasised in the featured case of one pair of teachers only through the effect of an intervention by the researcher. In another case, probabilities were strongly part of the discussion but seemed not to influence the actual decision. The superior role of impact is a key feature of the Priority Heuristic and is clearly one of the reasons that the Priority Heuristic predicted accurately the decisions of two pairs of teachers. We agree with Johnson et al. (2008) that the Priority Heuristic does seem to capture some important aspects of the decision-making process. Even the two complicating circumstances (the researcher's intervention and the distinction between making a decision for oneself or for Deborah) did not seem to undermine the Priority Heuristic.

The Priority Heuristic, as with all models, hides much of the process (Johnson et al., 2008), which is often centred on consideration of emotionally-charged issues. We conjecture that emotional responses are an essential part of the attempt to weigh the significance of impact and the effort involved may partially explain why impacts seem to take precedence over probabilities (Sunstein, 2003; Loewenstein et al., 2001).

In the (Brandstätter et al., 2006) study, impact and probabilities were given to the participants as numbers. The DD study therefore stands in contrast by accentuating features of authentic risk-based decision-making and seeks to promote through empathy exactly the emotive responses that may need to be made in order to 'measure' impact. If as a result the balance between impact and likelihood was further weighted towards impact, it is not surprising that the Priority Heuristic, with its accentuation on impact, accounted rather well for the decision making of two of the three pairs of teachers.

Why might it have failed in the third case? We turn here to the first two themes emerging from Analysis 2, trustworthiness and interpretation. It was clear that the teachers did not find it easy to hold impact and likelihood in some sort of balance. Why should they not find such a trade-off difficult when impact especially is so difficult, if not impossible, to measure, necessitating perhaps some affective response, as conjectured above? One of the arguments for the existence of the Priority Heuristic is that trading-off demands complex conceptual activity. In fact, trading-off requires the co-ordination of two or more variables (here, at least impact and likelihood) as in conceptualising proportion and ratio, known to be cognitively challenging.

We suppose that in the face of such complexity, people do indeed look for simpler methods such as the Priority Heuristic. But it is very unlikely in our estimation that any one such model will describe how risk-based decision making takes place. Indeed, when the information itself is contested as in our scenario, the Priority Heuristic may be especially difficult to apply. We have given examples of how the information is seen by the teachers as problematical because of its dubious source, contradictory nature or incomplete data. The findings in the Brandstätter study were based on responses to situations that were already precisely quantified both in terms of loss (or gain) and likelihood. Issues about reliability or interpretation were minimised. We believe that the scope of the Priority Heuristic to deal with complicated though perhaps more common situations where the information itself is uncertain is limited. In such circumstances it is likely that the participants will find an alternative heuristic, which perhaps for example relies on the authority of the source, as in Tim and Neil's decision making.

8 Implications for educational intervention

There are various limitations in our approach. We used only three pairs of teachers, which makes statistical generalisation impossible, though, at this exploratory stage, we draw the benefit of a high sense of validity in the richness of the data. That validity is admittedly threatened by the artificiality and the specificity of the scenario. Since in practice, ultimately school children too will be faced with pedagogic contrivances, the relationship between our research tool, DD, and what might one day be used as a teaching aid, is perhaps relatively close.

The above analysis demonstrates that the teachers placed highest priority on losses (rather than probabilities) and that, in two out of three cases, this was consistent with the Priority Heuristic. There is some evidence that use of the Priority Heuristic is not robust when a simple intervention seemed to push one pair to focus on elements of the heuristic that would normally not have been triggered. Such a lack of robustness is not surprising when the teachers were dealing with a complex scenario with many aspects unquantified and in mutual conflict. Nevertheless, we believe that this uncertainty reflects common scenarios for personal decision-making. Perhaps this lack of precision in how well the Priority Heuristic models 'real' decision making is what concerns Cokely and Kelly (2009) when they claim from their experimental evidence that more precise process modelling of risk choices with the Priority Heuristic would require at least one parameter that creates variation in the search and stopping rules.

The lack of robustness materialises also in another, perhaps more profound, way. When faced with uncertain information, the basis for the Priority Heuristic may be undermined and in such situations we would expect the use of an alternative strategy. Indeed, our position is that authentic decision making is often based on appeal to authority and we saw some evidence of that in this study. In some ways, the lack of robustness of the Priority Heuristic might signal opportunities for the teacher, who wishes to perturb the student's thinking about risk. For example, a teacher might intervene to challenge the position a student is taking, triggering further reflection on the student's own decision making process.

In fact, the Priority Heuristic attempts to model how risk-based decision making operates but does not claim that this is how we would like things to be. As educationalists, we seek to go one step further, asking how should we go about teaching students so that their decision making might be more sensitive to the issues. Both the Priority Heuristic and our data stress that impact plays a highly significant role in decision making compared to probabilities. In contrast, EUT positions both with equal weighting. It seems reasonable to suggest that it would be important pedagogically to address this imbalance even if we would not want to suggest that EUT is realisable in individual decision making.

In designing the Dilemma, we massaged some of the medical data describing the consequences of the operation in order to present what seemed to us a non-trivial decision. In teaching contexts, it might also be important to choose scenarios where a clear decision is not easily made. A strength of DD in our view is that the design takes into account how emotion can play an important part in judging risk in the process of decision making. The decision maker is encouraged to reorganize information through the active use of the probability simulator and the painometer. Encouraging empathy not only piques curiosity but also brings to bear emotionally-charged reactions that give the scenario some degree of authenticity, though we recognise it remains contrived.

One consequence of utilising a situation intentionally to generate emotional responses is that the design investment is targeted at just one context. Thus another limitation of the study is the extent to which our conclusions would apply in other scenarios. The Priority Heuristic has been developed by researchers in settings where there was little or no emotional involvement and yet it was concluded that impact is typically given highest priority in decision making. If the Priority Heuristic settings and DD are seen as opposite ends of a dimension that measures the degree of emotional involvement of the participants, it is interesting that at both ends of this imaginary axis impact seems very significant.

In the version used with teachers, these emotions are realised almost entirely through empathy with the impact of the various hazards, complications of the operation and painful activities in Deborah's life. We are seeking better ways of representing the probabilities or rather Deborah's futures. The graph in Fig. 1 is a first attempt. Perhaps meaningful representations of how often different outcomes might happen will enable people to judge probabilities against impacts in a more balanced way.

We are not satisfied that people's risk-based decision making is dominated by strategies such as the Priority Heuristic or appealing to the source with most authority. We need to find tools that can list and order hazards by size of risk to support co-ordinated ways of thinking so that impact and likelihood can to some extent be traded off. After the experiment with teachers, we implemented such a tool and we offer it here as an example of the sort of intervention that might be made. This hazard-mapping tool is presented to users as a means of keeping an ongoing 'map' of their decisions by connecting boxes containing information they have entered about possible hazards (Fig. 3).

In Fig. 3, we envisage a user who has entered two decision boxes and three hazard boxes, associated with these decisions, as read from the information or observed through the videos. The user has entered a description of the chosen hazards. They might include information about impact and likelihood but in fact can enter any information, including ethical or moral concerns. The boxes can be connected to illustrate any conceptual links. The user may

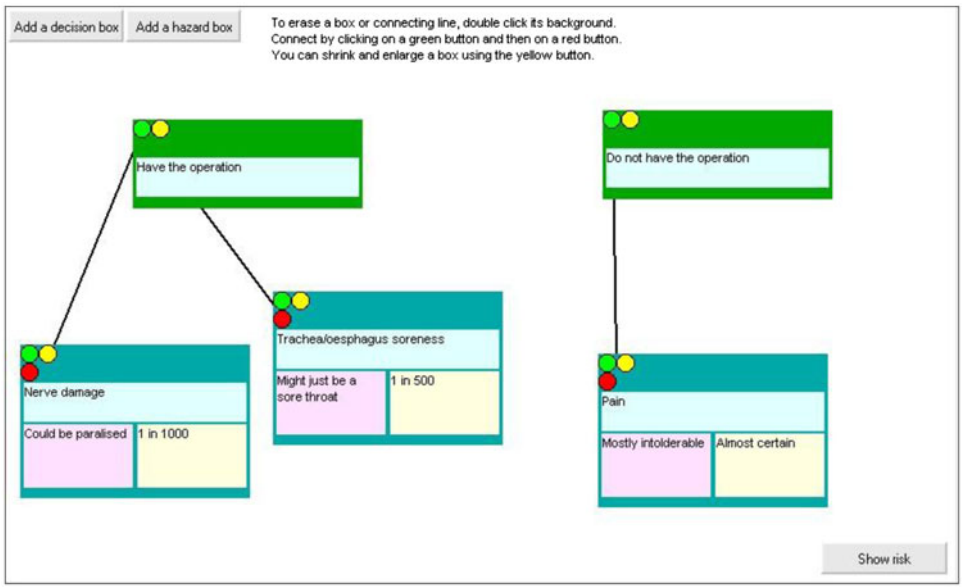

Fig. 3 Enhanced version of Deborah's Dilemma where users can map their analysis of the hazards

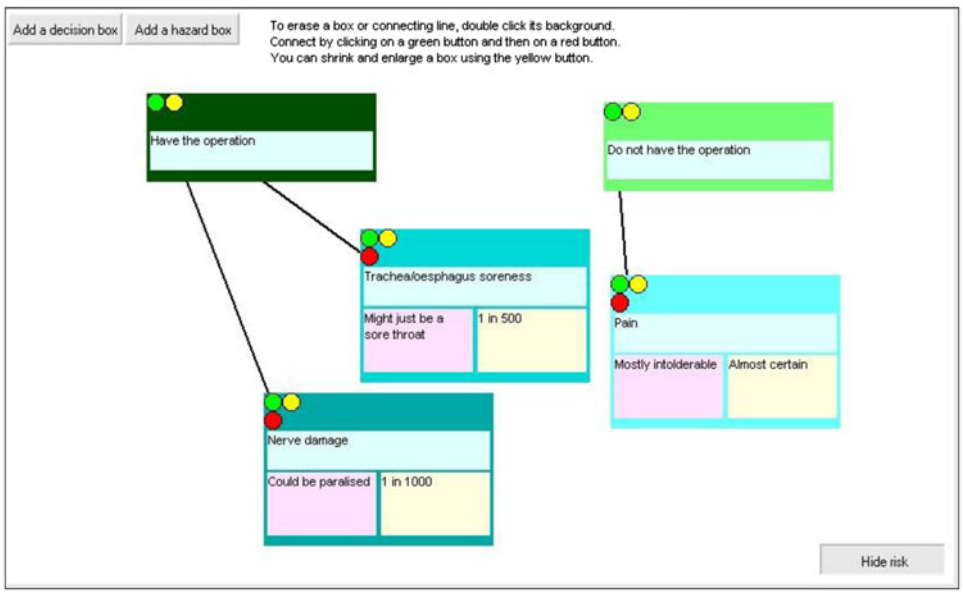

Fig. 4 The teachers would press the risk button (*bottom right corner*) and then drag the hazards so that the depth of color reflected the perceived risk of each hazard. Here, having the operation is perceived as riskier and so this decision and its associated hazards have been positioned further to the left giving them a darker color. Similarly nerve damage is seen as riskier than trachea soreness

at any time add more hazards, as they work on the Dilemma. The boxes can be dragged around the screen but the links will be maintained. At a later stage, the user would be encouraged to press the 'Show Risk' button in the bottom right hand corner. The boxes would change colour, becoming darker or lighter. Boxes towards the left of the screen would become darker while those to the right would become lighter on a continuous scale (Fig. 4). The user is told that the darker the hazard, the greater its risk. Inevitably, a user would now judge that some of the boxes were in the wrong position on the screen. They would be able to drag the boxes to what they would judge to be the correct relative position according to their estimation of the risks.

We conjecture that such tools could provide an educational intervention that would enable teachers and students to co-ordinate the dimensions of risk into a single construct in a process that includes the possibility that they might, under certain circumstances, use thinking about trade-offs rather than strategies, such as the Priority Heuristic, that avoid them.

Teaching about risk carries with it certain obligations. We see one pedagogic challenge as sensitising people to their own decision making, including their emotionally-charged heuristic thinking. It could be argued that this task falls most naturally to the teacher of social studies. We do not though think science and mathematics teachers can so easily escape their obligations. Insofar as science teachers are required to teach about socio-scientific issues and mathematics teachers about probability, the need to incorporate consideration of risk in this teaching is unavoidable. In this paper, we have set out some pertinent issues to inform the challenge of designing educational interventions in those classrooms.

References

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293–1295.

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica, 22*, 23–36 (Original work published 1738).

Birnbaum, M. H. (2008). Evaluation of the Priority Heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer and Hertwig (2006). *Psychological Review, 115*(1), 253–262.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*(2), 409–432.

Cokely, E., & Kelly, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgement and Decision Making, 4*(1), 20–33.

Fiedler, K. (2010). How to study cognitive decision algorithms: The case of the priority heuristic. *Judgment and Decision Making, 5*(1), 21–32.

Flick, U. (2006). *An introduction to qualitative research* (3rd ed.). London: Sage Publications.

Friedman, M., & Savage, L. (1952). The expected utility hypothesis and the measurability of utility. *Journal of Political Economy, 60*, 463–474.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristics: Why biased minds make better inferences. *Topics in Cognitive Science, 1*, 107–143.

Gigerenzer G., Odd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.

Hilbig, B. E. (2008). One-reason decision making in risky choice? A closer look at the priority heuristic. *Judgment and Decision Making, 3*(6), 4–462.

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. C. (2008). Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review, 115*(1), 263–273.

Katsikopoulos, K. (2011). Psychological heuristics for making inferences: Definition, performance, and the emerging theory and practice. *Decision Analysis, 8*(1), 10–29.

Katsikopoulos, K., & Gigerenzer, G. (2008). One-reason decision-making: Modeling violations of EUT. *Journal of Risk and Uncertainty, 37*(1), 35–56.

Levinson, R., Kent, P., Pratt, D., Kapadia, R., & Yogui, C. (2011). Developing a pedagogy of risk in socio-scientific issues. *Journal of Biological Education, 45*(3), 136–142.

Loewenstein, G., Weber, E., Hsee, C., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267–286.

Mongin, P. (1998). EUT. In J. Davis, D. Hands, & U. Mäki (Eds.), *The handbook of economic methodology* (pp. 342–350). Cheltenham: Edward Elgar.

Pratt, D., Ainley, J., Kent, P., Levinson, R., Yogui, C., & Kapadia, R. (2011). Role of context in risk-based reasoning. *Mathematical Thinking and Learning, 13*(4), 322–345.

Rieger, M. O., & Wang, M. (2008). What is behind the Priority Heuristic? A mathematical analysis. *Psychological Review, 115*(1), 274–280.

Savage, L. J. (1954). *The foundations of statistics* (2nd ed.). New York: Dover.

Sunstein, C. (2003). Terrorism and probability neglect. *Journal of Risk and Uncertainty, 26*(2–3), 121–136.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wengraf, T. (2001). *Qualitative research interviewing: Biographic narrative and semi-structured methods*. London: Sage.